Residential investment and recession predictability*

Knut Are Aastveit[†] André K. Anundsen[‡] Eyo I. Herstad[§]

October 4, 2018

Abstract

We assess the importance of residential investment in predicting economic recessions for an unbalanced panel of 12 OECD countries over the period 1960Q1–2014Q4. Our approach is to estimate various probit models with different leading indicators and evaluate their relative prediction accuracy using the area under the receiver operating characteristic curve as our metric of forecasting performance. We document that residential investment contains information useful in predicting recessions both in-sample and out-of-sample. This result is robust to adding typical leading indicators, such as the term spread, stock prices, consumer confidence surveys and oil prices. It is shown that residential investment is particularly useful in predicting recessions for countries with high home-ownership rates. Finally, in a separate exercise for the US, we show that the predictive ability of residential investment is – in a broad sense – robust to employing real-time data.

Keywords: *Recession predictability; Housing; Leading indicators; Real-time data; Panel data*

JEL classification: *C33; C53; E32; E37*

*This paper should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank. The authors gratefully acknowledge comments and suggestions from the editor, associate editor, two anonymous referees, Gunnar Bårdsen, Joe Gyourko, Nina Larsson Midthjell, Kjersti Næss Torstensen, Stuart Rosenthal, Phil Rothman, Albert Saiz, Jim Stock, Herman van Dijk and Carl Walsh as well as seminar and conference participants at Norges Bank, the SNDE in Tuscaloosa, the 2017 Annual Meeting of the Norwegian Association of Economists and the CFE 2017 in London.

[†]Norges Bank and Centre for Applied Macro and Petroleum Economics at BI Norwegian Business School, knut-are.aastveit@norges-bank.no.

[‡]Norges Bank, andre-kallak.anundsen@norges-bank.no.

[§]University of Chicago, eyoherstad@gmail.com.

1 Introduction

Economic policy decisions depend on whether the economy is in an expansion or a recession. Accurate predictions of business cycle turning points, and impending economic recessions in particular, are therefore of great importance to central banks and other policy institutions. A vast amount of research has shown that a variety of economic and financial variables contain predictive information about future recessions, see e.g. Marcellino (2006) and Liu and Moench (2016). In particular, Estrella and Hardouvelis (1991) and Estrella and Mishkin (1998) have documented that the slope of the term structure has strong predictive power for US recessions.[1] Several other variables have also been regarded as leading recession indicators, including stock prices (Estrella and Mishkin (1998) and Stock and Watson (2003)), the index of leading economic indicators (Berge and Jordà (2011) and Stock and Watson (1989)), oil prices (Hamilton (1983, 1996) and Ravazzolo and Rothman (2013, 2016)) and survey data (Hansson et al. (2005), Claveria et al. (2007) and Aastveit et al. (2016)).[2]

While the Great Recession has triggered a new search for methods and leading indicators that may be useful for calling recessions in real time, it has also highlighted the importance of understanding the linkages between the housing market, macroeconomic activity and financial stability, see e.g. Iacoviello and Neri (2010), Claessens et al. (2012), Mian et al. (2013), Mian and Sufi (2014), Jordà et al. (2015), Leamer (2007, 2015) and Kydland et al. (2016). More specifically, Leamer (2007) argues that real estate markets were grossly understudied by macroeconomists and policymakers interested in understanding business cycle dynamics. Considering the cumulative detrended contribution of each aggregate demand component to GDP growth before, during and after US recessions, he documents that residential investment offers by far the best early warning indicator of oncoming recessions. He concludes that a fall in the contribution of residen-

[1]Interestingly, Rudebusch and Williams (2009) also document that the term spread consistently outperforms professional forecasters in predicting recessions.

[2]As highlighted by, e.g., Evans (2005), Giannone et al. (2008), and Aastveit et al. (2014), an advantage of surveys and financial market data is that they are timely available, with few revisions.

tial investment to GDP growth is a reliable harbinger of a future recession.[3] Despite this focus, studies that use housing-related variables to forecast recessions are scarce. In this paper, we aim to fill this gap in the literature.

We evaluate both the in- and out-of-sample performance of residential investment growth in predicting economic recessions for a panel of 12 OECD countries over the period 1960Q1–2014Q4. Our main question is whether residential investment contains predictive information about future recessions over and above other typical leading indicators, such as the term spread, stock prices, consumer confidence surveys and oil prices. For the US, we use the business cycle dating chronology provided by the National Bureau of Economic Research (NBER) to date business cycle turning points. For the other countries, we use the business cycle chronology provided by the Economic Cycle Research Institute (ECRI). These recession indicators are binary variables, and we follow custom (see e.g. Estrella and Hardouvelis (1991) and Liu and Moench (2016)) and estimate non-linear probit models to map the changes in predictor variables into recession forecasts.[4]

When assessing the accuracy of a given probit model in predicting recessions, we follow Berge and Jordà (2011) and Liu and Moench (2016) and calculate the receiver operating characteristic (ROC) curve. To summarize the forecast performance implied by each ROC curve, we integrate the area under the curve (AUROC).

Recession predictability is evaluated both in-sample and in a quasi real-time out-sample forecasting exercise, where we recursively predict recessions for the 1990Q1-2014Q4 sample. In a final exercise, we assess the importance of data revisions for recession predictability, using real-time data for the US obtained from the ALFRED (ArchivaL Federal Reserve Economic Data) database maintained by the Federal Reserve Bank of St. Louis.

[3]Ghent and Owyang (2010), however, do not find similar results when studying the relationship between housing and business cycles using data for US metro areas. They find no consistent evidence that housing-related variables influence the city's business cycle. Instead they suggest the possibility that housing is merely a proxy for other consumption or wealth indicators.

[4]Two alternative parametric approaches that allow for recession forecasts are the threshold autoregressive model, see e.g. Potter (1995), Tommaso (1998), Ferrara and Guégan (2005), and Billio et al. (2013), and the Markov-switching model, see e.g. Hamilton (1989), Chauvet (1998) and Chauvet and Piger (2008).

We have three main findings. First, residential investment is useful in predicting recessions, both in-sample and out-of-sample. This holds true also when we include the term spread, stock prices, consumer confidence surveys and oil prices. In all cases, AUROC increases significantly when residential investment is added to the model. This result holds irrespective of forecasting horizon ($h = 1, \ldots, 6$) and whether we estimate models for each country separately or together as a panel. Our results are robust to excluding the Great Recession from the forecasting evaluation period.

Second, we find that residential investment improves predictability the most for countries with high home-ownership rates. For instance, we find substantial forecast improvements by including residential investment as a recession predictor for Spain, Norway and the US – all of which have a high home-ownership rate. For countries with low home-ownership rates, such as Japan and South Korea, residential investment does not contain predictive information about future recessions above other leading indicators.

Third, the real-time data exercise shows that the predictive ability of residential investment is – in a broad sense – robust to data revisions. However, as expected, the forecast accuracy is somewhat lower when we account for data revisions relative to the case where real-time data is not considered. Thus, with even more accurate data on residential construction activity, e.g. the number of cranes at work at each point in time, real-time recession forecasts could be improved even further.

Our paper contributes to the literature that estimate and predict business cycle turning points (see e.g., Anas et al. (2008), Darné and Ferrara (2011) and Billio et al. (2012) for applications to the euro area, Chauvet (1998), Chauvet and Piger (2008), Harding and Pagan (2002, 2006), Hamilton (2011) and Stock and Watson (2014) for applications to the US) by documenting that residential investment contains information over and above other typical leading indicators. We therefore corroborate and extend the findings in Leamer (2007, 2015) by showing that residential investment is important for predicting recessions also in a cross-country setting – especially in countries with a high home-ownership rate.

The rest of the paper is structured as follows. Section 2 discusses the empirical methodology used to predict recession probabilities and evaluate the classification of future recessions. Section 3 provides a description of the international panel data and the US real-time data used in our analysis. Section 4 presents results from the panel data exercise and the real-time analysis on US data. The final section concludes the paper.

2 Empirical methodology

2.1 A probit approach for predicting recessions

The state of the business cycle is a binary variable, taking the value one during recessions and zero otherwise. Since our dependent variable is binary, we follow custom in the literature on predicting recessions, see e.g., Estrella and Hardouvelis (1991) and Liu and Moench (2016), and estimate a probit model to predict the probability of a recession. Our probit model can be written as:

$$y_{i,t} = \begin{cases} 1 & \text{if there is a recession in quarter t} \\ 0 & \text{if there is not a recession in quarter t} \end{cases} \tag{1}$$

$$E(y_t|y_t^*) = P(y_t|y_t^*) = f(y_t^*) \tag{2}$$

$$y_{i,t}^* = \alpha_i + \sum_{j=1}^{k} \boldsymbol{X}_{i,t-j}\ \boldsymbol{\beta}_j + \sum_{j=1}^{k} \gamma_j IP_{t-j} + \varepsilon_{i,t} \tag{3}$$

where $f(y^*)$ is the cdf of the normal distribution, and $\boldsymbol{X}$ is a vector of explanatory variables used to forecast recession probabilities. IP is the industrial production index for advanced economies, used as a proxy for the global business cycle, whereas α_i is a country-specific intercept. We set the lag length, k, to 4 in all estimations.

We consider three different ways of estimating the model. First, we estimate the model using all available data, referred to as the full-sample estimate. Second, we estimate the model recursively using at each point in time only observations that are available at that particular point in time. This exercise is conducted using the final data vintage. Lastly, we estimate the model recursively using real-time vintage data for each period. Using real-time data, we get a recession probability estimate that is based on data that were available at the forecast origin. Due to lack of data availability, the real-time estimation is only done for the US.

2.2 Forecast evaluation

For a given model, m, a recession signal is issued whenever the estimated probability of a recession from that model, $\hat{p}_m$, exceeds some threshold level, τ. There are two types of errors that can be made: the model fails to predict a recession (Type I error), or the model issues a false recession signal (Type II error).

Let the true positive rate ($TPR_m(\tau)$) denote the share of all recessions in which a correct signal is issued, i.e. one minus the share of Type I errors. Further, let the false positive rate ($FPR_m(\tau)$) be the fraction of all non-recession events in which a false signal is issued (the share of Type II errors). Lowering the value of the threshold parameter will in general imply that the model issues more signals. While this increases the share of correctly predicted recessions, it comes at the cost of issuing more false alarms. The opposite is true if the value of the threshold parameter is increased.[5] Determining the optimal threshold requires knowledge of the policymaker's preferences regarding the trade-off between Type I and Type II errors, which depends (among other things) on the relative cost of the different outcomes, as well as the frequency at which recessions occur. One way of formalizing this trade-off is by formulating a loss function. For model m, a

[5]Thus, $\underset{\tau\to 0}{TPR}(\tau) = \underset{\tau\to 0}{FPR}(\tau) = 1$ and that $\underset{\tau\to 1}{TPR}(\tau) = \underset{\tau\to 1}{FPR}(\tau) = 0$. A perfect model never issues any false signals ($FPR = 0$), while it always correctly predicts all recession episodes ($TPR = 1$). For any $\tau \in (0, 1)$, an informative model should deliver a $TPR(\tau) > FPR(\tau)$.

linear loss function could take the following form (see e.g. Sarlin (2013)):

$$L_m(\theta, \tau) = \theta p(1 - TPR_m(\tau)) + (1 - \theta)(1 - p)FPR_m(\tau) \tag{4}$$

where p is the unconditional probability of a recession, or the frequency of recessions in the sample under consideration, whereas θ is the relative weight that the policymaker attaches to missing a recession.

A complementary tool that has been used to compare the classification abilities of early warning models is the *Receiver Operating Characteristic* (ROC), which plots the full mapping of the false positive rate, $FPR_m(\tau)$, and the true positive rate, $TPR_m(\tau) = TPR_m(FPR_m(\tau))$, across different values of the threshold parameter τ (see e.g. Jordà and Taylor (2011), Berge and Jordà (2011), Jordà and Taylor (2012) and Anundsen et al. (2016) for further details). To assess the recession classification abilities of various leading indicators, we follow Berge and Jordà (2011) and Liu and Moench (2016) and calculate the *Area Under Receiver Operating Characteristic* (AUROC), which takes into account every point on the ROC curve. More formally, the AUROC is defined as:

$$AUROC_m = \int_{\tau=0}^{1} TPR_m(FPR_m(\tau))FPR'_m(\tau)d\tau \tag{5}$$

The advantage of AUROC is that it is independent of the policymaker's preferences and it covers all possible preference parameters (see Elliott and Lieli (2013)). When comparing the performance of model m relative to model c, model m is preferred to model c if $AUROC_m > AUROC_c$, i.e. model m has a higher TPR for a given FPR than model c on average.[6]

As shown in DeLong et al. (1988), a numerical estimate of AUROC for model m can be achieved by considering a discrete time version of (5):

[6]A perfect model has $AUROC = 1$, while a completely uninformative model has $AUROC = 0.5$.

$$AUROC_m = \frac{1}{qr}\sum_{j=1}^{r}\sum_{i=1}^{q}\psi(\hat{p}_{m,i}^{\text{Rec.}}, \hat{p}_{m,j}^{\text{No rec.}}) \tag{6}$$

where $\hat{p}_m^{\text{Rec.}}$ are the q implied probabilities of a recession from model m in recessionary states, whereas $\hat{p}_m^{\text{No rec.}}$ denotes the r implied probabilities of a recession from the same model in non-recessionary states. The kernel, ψ, is given as:

$$\psi(\hat{p}_m^{\text{Rec.}}, \hat{p}_m^{\text{No rec.}}) = \begin{cases} 1 \text{ if } \hat{p}_m^{\text{No rec.}} < \hat{p}_m^{\text{Rec.}} \\ \frac{1}{2} \text{ if } \hat{p}_m^{\text{No rec.}} = \hat{p}_m^{\text{Rec.}} \\ 0 \text{ if } \hat{p}_m^{\text{No rec.}} > \hat{p}_m^{\text{Rec.}} \end{cases} \tag{7}$$

DeLong et al. (1988) suggest the following Wald-type test statistic to compare AUROC values from model m to model c (see Berge and Jordà (2011) and Anundsen et al. (2016) for economic applications):

$$W_{\text{AUROC}} = \frac{\text{AUROC}_m - \text{AUROC}_c}{se\,(\text{AUROC}_m - \text{AUROC}_c)}$$

W_{AUROC} follows a standard normal distribution under the null hypothesis of no difference and variances and covariances can be calculated as shown in DeLong et al. (1988).[7] Thus, when formally testing whether model m is preferred to model c, we compare W_{AUROC} to the relevant critical value from a standard normal distribution. We report results for two-sided tests.

3 Data

3.1 International data

We use quarterly data for 12 OECD countries; Australia, Canada, France, Italy, Japan, South Korea, New Zealand, Norway, Spain, Sweden, the UK and the US. For each country

[7] See Section A in the supplementary Appendix for an explanation of how the covariance matrix is calculated.

the dependent variable is a binary recession indicator, which takes the value one during recessions and zero otherwise. For the US, we use the business cycle dating chronology provided by the National Bureau of Economic Research (NBER) to date business cycle turning points. For the other countries, we use the business cycle chronology provided by the Economic Cycle Research Institute (ECRI).[8] However, since ECRI does not provide a business cycle chronology for Norway, we use the business cycle turning point dates from Aastveit et al. (2016) as the reference cycle for Norway.

We have an unbalanced panel, which at most covers the period 1960Q1 to 2014Q4, see Table B.1 in section B in the supplementary Appendix for details on data coverage for each country and each variable. For each of the countries, we consider residential investment, the term spread, stock prices, consumer confidence survey, oil price and industrial production index for advanced economies as explanatory variables.[9] The data on residential investment and the series for industrial production for advanced economies are taken from the OECD Economic Outlook. Most other series are collected from the Global Financial Database (GFD).[10] The exceptions are the interest rate series for South Korea and the stock price series for Canada and France, which are taken from the GVAR database.[11] The West Texas Intermediate (WTI) oil price is provided by the Federal Reserve Economic Data (FRED), maintained by the Federal Reserve Bank of St. Louis.

3.2 US real time data

Accounting for data revisions may be relevant when assessing the importance of various leading indicators for out-of-sample predictability of economic recessions. Unfortunately, a real-time database for the variables we use for the 12 countries in the panel data set does not exist. However, in order to shed some light on the importance of real-time

[8]See `https://www.businesscycle.com/ecri-business-cycles` for more details about the business cycle dates provided by ECRI.

[9]We calculate the term spread as the difference between the 10-year and 3-month government bond (treasury) yield. For the consumer confidence survey we use the quarterly difference, while for all the other predictors we transform the series using quarterly log differences.

[10]`https://www.globalfinancialdata.com/`

[11]`https://sites.google.com/site/gvarmodelling/data`

data for out-of-sample predictability of economic recessions, we use real-time data for the US obtained from the ALFRED (ArchivaL Federal Reserve Economic Data) database maintained by the Federal Reserve Bank of St. Louis. This database consists of collections of real-time vintages of data for each of the variables that we use, and the first vintage starts in 1966Q1. Vintages vary across time as either new data are released or existing data are revised by the relevant statistical agency. Using data from this database ensures that we only use data that were available at the date of the forecast origin. For the importance of using real-time data for macroeconomic forecasting, see Croushore and Stark (2001) and Croushore (2006).

4 Empirical Analysis

In this section, we describe our results from comparing various probit model specifications at different horizons. In doing so, we use an unbalanced panel for the 12 OECD countries described in Section 3. For some countries, such as the US, Australia and Japan, residential investment, the term spread and stock prices are available for the full sample, covering the period 1960Q1-2014Q4. For other countries, such as Spain, Norway and Sweden, some of these variables are first available in the 1970s or early 1980s. On the other hand, consumer confidence indicators are available only for a shorter sample for all countries except the US. While for most countries they are available sometime during the 1970s or early 1980s, for Sweden and South Korea they are first available in 1995 and 1998, respectively.

4.1 Forecasting exercise

We carry out both an in-sample and an out-of-sample analysis. In both exercises, we consider forecasts for horizons $h = 1, \ldots, 6$, where $h = 1$ corresponds to a nowcast. Our approach is therefore to forecast the probability of a recession h-quarters ahead. At each horizon, we begin by estimating baseline probit models with four lags, including two explanatory variables; the industrial production index for advanced economies and one of the typical leading indicators considered in the literature.[12] Then, since our main interest is to assess the importance of residential investment for in- and out-of-sample predictability of economic recession, we augment each baseline specification with four lags of residential investment growth.[13] Following this procedure, we can formally test whether residential investment contains predictive information over and above what is contained in other explanatory variables traditionally considered in the recession forecasting literature. We use AUROC to evaluate the performance of each model and also to compare the forecasting ability of each model with the baseline models. For each specification, we report the standard error of the AUROC. We can therefore evaluate the uncertainty around the estimates and assess whether adding residential investment to the models including standard recession predictors leads to a significant increase in recession forecasting accuracy.

4.2 Panel results

We report AUROC values for the different baseline probit models from our in-sample and out-of-sample exercise in Table 1 and Table 2, respectively. In both tables, the upper panel reports results for the sample 1990Q1-2014Q4 and the lower panel reports results for the sample 1985Q1-2005Q4, excluding the Great Recession.

For the out-of sample evaluation, we perform the following exercise for the 1990Q1-2014Q4 sample: First, each probit model is estimated using all data until 1989Q4. Then, the estimated parameters are used to predict recession probabilities over the next six

[12] Our results are robust to using two alternative measures of the global business cycle; the index of global real economic activity in industrial commodity markets (see Kilian (2009)) and growth in total GDP for OECD countries. Results are reported in Tables C.1-C.4 in section C in the supplementary Appendix.

[13] In a previous version of this paper, we followed Leamer (2007, 2015) more closely and used the contribution of residential investment to GDP growth as our baseline measure. However, as the contribution of residential investment to GDP growth depend not only on residential investment but also on other spending that varies with the business cycle, we use residential investment growth as our baseline measure. We are thankful to one of our referees for pointing this out. All results in the paper are, however, robust to using the contribution of residential investment to GDP growth as the baseline measure.

quarters: 1990Q1 ($h = 1$ corresponding to the nowcast) to 1991Q2 ($h = 6$). Next, we re-estimate the models, now using data until 1990Q1, and use the new estimated parameters to predict recessions for 1990Q2-1991Q3. We continue like this until we have forecasted recession probabilities for the end of the sample. A similar approach is used for the 1985Q1-2005Q4 sample.

Focusing on the upper panels of Table 1 and Table 2, we have two interesting results. First, for all models, the forecasting performance is better the shorter is the forecasting horizon. Second, among the various specifications, the models including residential investment growth achieve the highest AUROC values at most horizons.[14]

By comparing the upper and lower panels in the two tables, it is evident that residential investment is highly comparable to the other leading indicators also when excluding the Great Recession from the forecasting evaluation period. In fact, the term spread is the only leading indicator that achieve higher AUROC values for the pre Great Recession sample.

Next, we proceed to study whether or not residential investment contains predictive information about future recessions over and above the standard leading indicators considered in the literature, i.e., the term spread, stock prices, consumer confidence survey and oil prices. We therefore augment each baseline specification with four lags of residential investment growth.[15] For this exercise, we only report out-of-sample results.[16] The achieved gains (increase in AUROC) for both the 1990Q1-2014Q4 and the 1985Q1-2005Q4 sample are reported in panel A of Table 3. For the 1990Q1-2014Q4 sample, our findings suggest that the gain in AUROC is statistically significant for all specifications

[14] Note that the data samples for the various leading indicators differ, see Table B.1 in section B in the supplementary Appendix for details. The results reported in the various rows are therefore based on different estimation samples. Thus, the results for the relative performance based on the various leading indicators are therefore not strictly comparable. Results where the same estimation sample is used for each country and for all the various leading indicators are similar, please see Tables C.5-C.6 in section C in the supplementary Appendix.

[15] We have also explored whether residential investment add predictive information to alternative baseline models. In these alternative baseline models, we include three explanatory variables; the industrial production index for advanced economies and different combinations of two of the standard leading indicators. Our results suggest that including residential investment in these models also leads to improved forecast accuracy. Detailed results are given in Table C.7 in Section C of the supplementary Appendix.

[16] The in-sample results are similar.

Table 1: In-sample AUROC values for different leading indicators – Panel results.

Variable/Horizon	1-step	2-step	3-step	4-step	5-step	6-step
		Evaluation sample: 1990Q1-2014Q4				
Res. Invest	0.819	0.786	0.755	0.731	0.706	0.676
	(0.016)	*(0.018)*	*(0.019)*	*(0.019)*	*(0.020)*	*(0.020)*
Spread	0.732	0.714	0.706	0.711	0.709	0.707
	(0.018)	*(0.019)*	*(0.018)*	*(0.018)*	*(0.018)*	*(0.017)*
Stocks	0.784	0.751	0.714	0.675	0.646	0.629
	(0.016)	*(0.017)*	*(0.017)*	*(0.018)*	*(0.019)*	*(0.018)*
C. Conf	0.776	0.756	0.739	0.731	0.714	0.703
	(0.018)	*(0.019)*	*(0.019)*	*(0.020)*	*(0.020)*	*(0.020)*
Oil Price	0.718	0.680	0.649	0.641	0.633	0.625
	(0.019)	*(0.019)*	*(0.019)*	*(0.019)*	*(0.019)*	*(0.019)*
		Evaluation sample:1985Q1-2005Q4				
Res. Invest	0.813	0.791	0.769	0.743	0.711	0.684
	(0.020)	*(0.022)*	*(0.023)*	*(0.024)*	*(0.025)*	*(0.026)*
Spread	0.850	0.851	0.849	0.843	0.828	0.814
	(0.017)	*(0.017)*	*(0.017)*	*(0.016)*	*(0.017)*	*(0.017)*
Stocks	0.780	0.758	0.736	0.715	0.691	0.676
	(0.019)	*(0.020)*	*(0.020)*	*(0.021)*	*(0.023)*	*(0.023)*
C. Conf	0.801	0.767	0.744	0.736	0.726	0.724
	(0.028)	*(0.031)*	*(0.034)*	*(0.034)*	*(0.035)*	*(0.034)*
Oil Price	0.716	0.690	0.669	0.674	0.684	0.684
	(0.023)	*(0.023)*	*(0.024)*	*(0.023)*	*(0.022)*	*(0.023)*

Note: This table shows in-sample AUROC values (with standard errors in parentheses) for probit models using various leading indicators at different forecasting horizons. All the models are evaluated based on two different time periods, 1990Q1 to 2014Q4 and 1985Q1 to 2005Q4, respectively.

and at all horizons when residential investment is added to the model. Our results therefore suggest that residential investment contains important predictive information about future recessions over and above other leading indicators. Again, our results are robust to excluding the Great Recession from the forecasting evaluation period. Although there are some quantitative differences, we still find that models with residential investment growth obtain the highest AUROC values for most specifications and at most horizons. The only exception is that residential investment only contains predictive information about future recessions over and above the term spread at the shorter horizons.

Finally, we do the reverse exercise, i.e., we study whether or not the term spread,

Table 2: Out-of-sample AUROC values for different leading indicators – Panel results.

Variable/Horizon	1-step	2-step	3-step	4-step	5-step	6-step
		Evaluation sample: 1990Q1-2014Q4				
Res. Invest	0.788	0.738	0.687	0.643	0.603	0.574
	(0.015)	*(0.016)*	*(0.017)*	*(0.018)*	*(0.018)*	*(0.019)*
Spread	0.692	0.671	0.654	0.652	0.647	0.642
	(0.020)	*(0.020)*	*(0.020)*	*(0.019)*	*(0.019)*	*(0.019)*
Stocks	0.746	0.701	0.651	0.594	0.552	0.536
	(0.017)	*(0.018)*	*(0.019)*	*(0.019)*	*(0.020)*	*(0.019)*
C. Conf	0.750	0.712	0.683	0.669	0.643	0.628
	(0.019)	*(0.020)*	*(0.021)*	*(0.021)*	*(0.021)*	*(0.022)*
Oil Price	0.660	0.616	0.577	0.562	0.552	0.548
	(0.020)	*(0.021)*	*(0.021)*	*(0.020)*	*(0.020)*	*(0.020)*
		Evaluation sample:1985Q1-2005Q4				
Res. Invest	0.763	0.728	0.692	0.657	0.625	0.604
	(0.018)	*(0.019)*	*(0.020)*	*(0.021)*	*(0.022)*	*(0.023)*
Spread	0.784	0.782	0.775	0.763	0.746	0.729
	(0.022)	*(0.022)*	*(0.022)*	*(0.021)*	*(0.021)*	*(0.022)*
Stocks	0.735	0.697	0.657	0.615	0.579	0.561
	(0.022)	*(0.023)*	*(0.023)*	*(0.024)*	*(0.025)*	*(0.025)*
C. Conf	0.750	0.695	0.651	0.636	0.620	0.620
	(0.031)	*(0.033)*	*(0.035)*	*(0.034)*	*(0.034)*	*(0.034)*
Oil Price	0.651	0.614	0.591	0.589	0.596	0.599
	(0.026)	*(0.026)*	*(0.027)*	*(0.027)*	*(0.026)*	*(0.026)*

Note: This table shows out-of-sample AUROC values (with standard errors in parentheses) for probit models using various leading indicators at different forecasting horizons. All the models are evaluated based on two different time periods, 1990Q1 to 2014Q4 and 1985Q1 to 2005Q4, respectively.

stock prices, consumer confidence survey and oil prices contain predictive information about future recessions over and above residential investment. Results for the 1990Q1-2014Q4 sample are reported in panel B of Table 3. In this exercise, we augment a probit model with four lags of residential investment and the industrial production index of for advanced economies, with four lags of each of the alternative leading indicators separately.

The out-of sample results suggest that there is no gain from adding oil prices to a model containing residential investment. At all horizons, there is a gain from adding consumer confidence surveys, whereas stock prices only add at shorter horizons. On the other hand, the term spread has predictive information about future recessions over

Table 3: Out-of-sample AUROC gains – Panel results.

Variable/Horizon	1-step	2-step	3-step	4-step	5-step	6-step
Panel A: AUROC gains from adding residential investment growth						
Evaluation sample: 1990Q1-2014Q4						
Spread	0.098***	0.086***	0.073***	0.051***	0.028***	0.014***
	(0.012)	*(0.010)*	*(0.009)*	*(0.006)*	*(0.004)*	*(0.003)*
Stocks	0.073***	0.076***	0.073***	0.065***	0.046***	0.021***
	(0.010)	*(0.009)*	*(0.008)*	*(0.008)*	*(0.006)*	*(0.005)*
C. Conf	0.070***	0.078***	0.079***	0.060***	0.042***	0.027***
	(0.014)	*(0.013)*	*(0.013)*	*(0.012)*	*(0.011)*	*(0.010)*
Oil Price	0.116***	0.107***	0.090***	0.068***	0.042***	0.016***
	(0.013)	*(0.011)*	*(0.009)*	*(0.008)*	*(0.006)*	*(0.005)*
Evaluation sample:1985Q1-2005Q4						
Spread	0.036***	0.024**	0.014	0.007	-0.005**	-0.009***
	(0.012)	*(0.010)*	*(0.009)*	*(0.006)*	*(0.002)*	*(0.003)*
Stocks	0.057***	0.057***	0.046***	0.030***	0.013***	-0.001
	(0.011)	*(0.010)*	*(0.008)*	*(0.006)*	*(0.004)*	*(0.004)*
C. Conf	0.072***	0.078***	0.066***	0.037**	-0.000	-0.018*
	(0.022)	*(0.019)*	*(0.018)*	*(0.016)*	*(0.011)*	*(0.009)*
Oil Price	0.093***	0.078***	0.053***	0.032***	0.012**	-0.003
	(0.017)	*(0.012)*	*(0.010)*	*(0.006)*	*(0.005)*	*(0.004)*
Panel B: AUROC gains from adding various leading indicators						
Evaluation sample: 1990Q1-2014Q4						
Spread	0.002	0.019	0.041***	0.059***	0.072***	0.082***
	(0.010)	*(0.013)*	*(0.015)*	*(0.017)*	*(0.017)*	*(0.017)*
Stocks	0.031***	0.040***	0.038***	0.016	-0.006	-0.017
	(0.009)	*(0.011)*	*(0.012)*	*(0.012)*	*(0.012)*	*(0.013)*
C. Conf	0.031**	0.049***	0.066***	0.074***	0.068***	0.059***
	(0.012)	*(0.013)*	*(0.015)*	*(0.017)*	*(0.019)*	*(0.020)*
Oil Price	-0.013***	-0.015**	-0.020*	-0.013	-0.009	-0.010
	(0.004)	*(0.007)*	*(0.011)*	*(0.011)*	*(0.008)*	*(0.006)*

Note: Panel A shows AUROC gains (with standard errors in parentheses) when augmenting each baseline probit specification with residential investment growth as an additional regressor. All the models are evaluated based on the time periods 1990Q1 to 2014Q4 and 1985Q1 to 2005Q4, respectively. Panel B shows AUROC gains (with standard errors in parentheses) when augmenting a probit model with four lags of residential investment and the industrial production index of for advanced economies, with four lags of each of the alternative leading indicators separately. For all the models in Panel B the evaluation period is 1990Q1 to 2014Q4. For both panels, differences in accuracy that are statistically different from zero at a 10%, 5% and 1% significance level are denoted by one, two and three asterisks, respectively.

and above residential investment at longer horizons. That said, the gains are mostly considerably smaller in this case than when we conduct the opposite exercise, i.e. adding residential investment to models containing these indicators.[17]

When testing for equal forecast accuracy, we use two-sided tests. Monte Carlo evidence in Clark and McCracken (2011, 2015) show that the Diebold-Mariano test for equal forecast accuracy in the finite sample tends to have a size that is lower than the nominal size of the test, when compared against normal critical values. To the extent that this result also translates to comparing forecasts from nested probit models using AUROC's as the measure of forecast accuracy, the same may apply in our case.

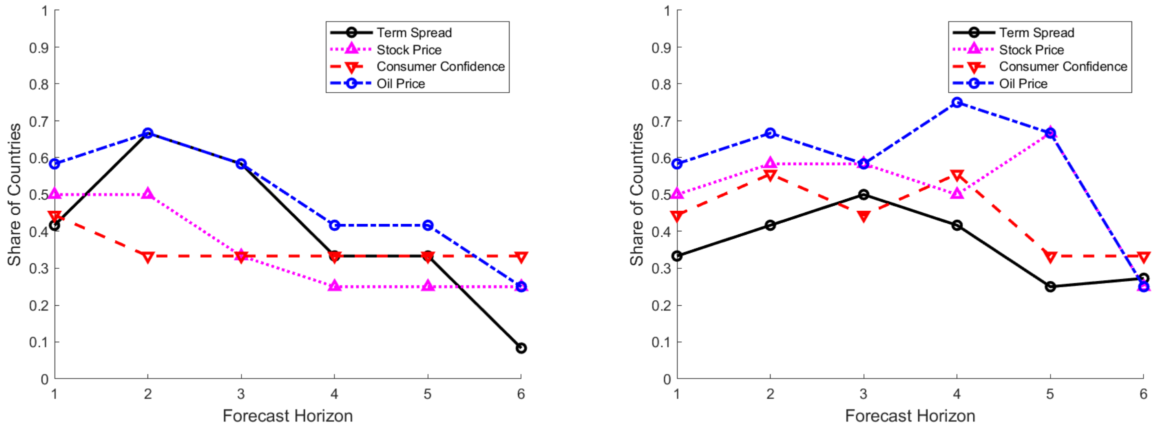
4.3 Country-specific results

Results from the panel analysis suggest a significant improvement in recession predictability when residential investment is added to the information set. In order to investigate the generality of this finding across countries, Figure 1 displays the share of countries where the AUROC increases significantly by augmenting the baseline specification with residential investment. The figure reports results for each baseline specification, both for the in-sample (Panel a) and out-of-sample (Panel b) exercise. For many of the countries in our sample, and in particular at the shorter forecasting horizons, including residential investment contributes to improving the forecasting accuracy of future recessions. There is, however, considerable heterogeneity across the different countries.

Tables D.1-D.12 in section D the supplementary Appendix highlight this heterogeneity in even more detail by reporting the AUROC values for all individual countries for both the in-sample and out-of-sample exercise. The results indicate a high degree of heterogeneity in the predictive ability of residential investment across countries. Residential investment seems to be particularly important in predicting future recessions for countries such as Spain, Norway and the US. These are all countries with relatively high home-ownership rates. In contrast, for countries such as South Korea and Japan, where home-ownership rates are relatively low, residential investment does not contain predic-

[17]Noteworthy exceptions are, however, consumer confidence surveys and the term spread at longer horizons.

Figure 1: Share of countries where residential investment adds significantly as a recession predictor relative to other leading indicators



(a) In-sample (b) Out-of-sample

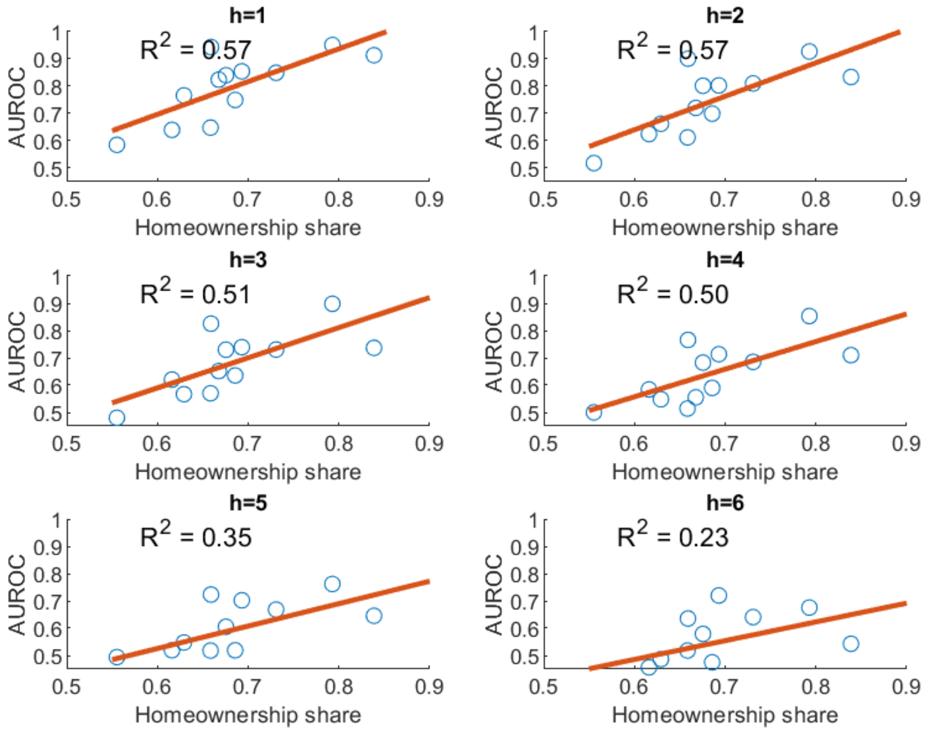
The figure shows the share of countries where we reject that the AUROC value based on predictions from the baseline model is equal to predictions from the baseline model augmented with residential investment. We report the result for both the in-sample and out-of-sample analysis. We use a significance level of 5%.

tive information about future recessions over and above the other leading indicators.

To shed some more light on this, Figure 2 shows a scatter plot of the AUROC from baseline models only containing residential investment at each forecast horizons against home-ownership rates. For forecast horizons, $h = 1 - 4$, there is a strong positive relationship between the AUROC and home-ownership rates. This suggests that residential investment is a stronger predictor of future recessions in countries with higher home-ownership rates.

Finally, Estrella and Hardouvelis (1991) and Estrella and Mishkin (1998) have documented that the slope of the term structure has strong predictive power for US recessions. Tables D.1-D.12 in section D the supplementary Appendix show that our results corroborate this finding for the US. There is, however, considerable heterogeneity in AUROC values across countries in the forecasting model that only includes the term spread. In particular, the term spread seems to be a good predictor of recessions in countries such as Australia, Canada, Norway, Sweden and the US. In contrast, for countries such as Italy, Japan, Spain and the UK, the term spread is not a very good predictor of recessions with

Figure 2: Scatter plot of AUROC values and home-ownership rates.



The figure shows scatter plots of AUROC and home-ownership rates, as well as a fitted line and its corresponding R^2 value for different forecast horizons.

AUROC values close to, or even below, 0.5 at several forecasting horizons.

4.4 Real-time results for the US

It is important to account for data revisions when assessing the importance of various leading indicators for out-of-sample predictability of economic recessions. While data for the term spread, stock prices and oil prices are not revised, data on residential investment are revised over time. It is therefore important to study whether the results in Section 4.2 and 4.3 are affected by accounting for data revisions.

Unfortunately, a real-time database for residential investment does not exist for the 12 countries in the panel. However, to shed light on whether data revisions affect the out-of-sample predictability of residential investment in forecasting economic recessions, we use real-time data for the US. In doing so, we conduct a recursive real-time out-of-sample forecasting exercise for the period 1966Q1-2014Q4, covering a total of seven recessions. Note that this is a considerably longer forecasting sample than the one used in the OECD cross-country panel study. Therefore, the relative performance of the various leading indicators differs somewhat from that obtained for the US in the panel study.

The main goal of this exercise is to investigate whether accounting for data revisions affects forecast performance, and not to compare the performance of the various leading indicators *per se*. The relevant comparison is therefore to check whether the AUROC gains and their statistical significance change between the quasi real-time exercise and the recursive real-time exercise. We therefore only report the AUROC gains in Table 4.

The results in Table 4 show that the AUROC gains from the real-time exercise (labeled RT in Table 4) are - in a broad sense - in line with the ones from the quasi real-time exercise (labeled as OOS in Table 4). As expected, for most specifications, the AUROC gains are somewhat smaller and standard errors somewhat larger when data revisions are taken into account. However, in most cases, the difference in AUROC gains is fairly small, and statistical significance is mostly unaffected. We take this as indicative evidence that data revisions are unlikely to materially affect the results from our panel analysis.

Finally, the relative performance of the different indicators for the quasi real-time exercise differs somewhat from that obtained for the US in the panel study in Section 4.3, see Table E.1 in the section E in the supplementary Appendix. This can – at least partly – be ascribed to two data-related differences. First, the real-time results for the US are based on a much smaller sample than the panel study. Second, both the estimation sample and the forecasting sample are different for the US real-time study than in the panel study. It is therefore worth noting that although the main finding that residential investment is a good predictor of US recessions is still maintained, residential investment does not contain out-of-sample predictive information about future recessions over and above the term spread at longer forecasting horizons.

Table 4: AUROC gains from including residential investment growth – US results. Estimated on 1953Q2 - 2014Q4, Evaluated on 1966Q1-2014Q4

Variable/Horizon		1	2	3	4	5	6
Spread	OOS	0.045	0.017	-0.030**	-0.034**	-0.033*	-0.036
		(0.028)	*(0.017)*	*(0.015)*	*(0.017)*	*(0.019)*	*(0.027)*
	RT	0.046*	-0.010	-0.052**	-0.048**	-0.043*	-0.053
		(0.028)	*(0.017)*	*(0.022)*	*(0.022)*	*(0.024)*	*(0.033)*
Stocks	OOS	0.049*	0.075**	0.080**	0.085**	0.059	0.062
		(0.027)	*(0.031)*	*(0.034)*	*(0.039)*	*(0.045)*	*(0.040)*
	RT	0.038	0.036	0.049	0.058	0.064	0.023
		(0.026)	*(0.026)*	*(0.036)*	*(0.048)*	*(0.049)*	*(0.048)*
C.Conf	OOS	0.043	0.050*	0.048*	0.064	0.120*	0.053
		(0.031)	*(0.030)*	*(0.027)*	*(0.041)*	*(0.061)*	*(0.051)*
	RT	0.022	0.045	0.019	0.043	0.096*	0.025
		(0.031)	*(0.033)*	*(0.038)*	*(0.051)*	*(0.055)*	*(0.051)*
Oil Price	OOS	0.286***	0.232***	0.147***	0.025	0.038	0.001
		(0.065)	*(0.055)*	*(0.049)*	*(0.040)*	*(0.043)*	*(0.045)*
	RT	0.263***	0.203***	0.120***	-0.010	-0.002	-0.048
		(0.060)	*(0.051)*	*(0.043)*	*(0.049)*	*(0.051)*	*(0.044)*

Note: This table shows AUROC gains (with standard errors in parentheses) when augmenting each probit specification with residential investment as an additional regressor. The models are evaluated based on the time period 1966Q1 to 2014Q4 and estimated on the sample 1953Q2 - 2014Q4. Differences in accuracy that are statistically different from zero at a 10%, 5% and 1% significance level are denoted by one, two and three asterisks, respectively.

5 Conclusion

In this paper, we have investigated whether residential investment growth helps to predict recessions over and above what is captured by standard leading indicators. Our results strongly suggest that recession predictability is improved, both in- and -out-of-sample, when residential investment is included. Conducting the reverse exercise, namely adding each of the alternative indicators to a model containing residential investment, we find less improvement in recession predictability. Moreover, our results suggest that residential investment is a particularly good indicator of the business cycle in countries where home-ownership rates are high. For the US, we test the robustness of these findings to accounting for data revisions. Our results are broadly robust to applying real-time data.

These results are important, since they suggest that the probability of timely recession detection by central banks and other policy institutions may be improved by taking into account the developments in residential investment – especially in countries with high home-ownership rates. Moreover, with even more accurate data on residential construction activity, e.g. satellite pictures of the number of cranes at work each quarter, real-time recession forecasts could be improved even further.

References

Aastveit, K., K. Gerdrup, A. Jore, and L. Thorsrud (2014). Nowcasting GDP in real time: A density combination approach. *Journal of Business and Economic Statistics 32*(1), 48–68.

Aastveit, K. A., A. S. Jore, and F. Ravazzolo (2016). Identification and real-time forecasting of Norwegian business cycles. *International Journal of Forecasting 32*(2), 283–292.

Anas, J., M. Billio, L. Ferrara, and G. L. Mazzi (2008). A System for Dating and Detecting Turning Points in the Euro Area. *The Manchester School 76*, 549–577.

Anundsen, A. K., K. Gerdrup, F. Hansen, and K. Kragh-Sørensen (2016). Bubbles and crises: The role of house prices and credit. *Journal of Applied Econometrics 31*(7), 1291–1311.

Berge, T. J. and O. Jordà (2011). Evaluating the classification of economic activity into recessions and expansions. *American Economic Journal: Macroeconomics 3*(2), 246–77.

Billio, M., R. Casarin, F. Ravazzolo, and H. van Dijk (2012). Combination schemes for turning point predictions. *Quarterly Review of Economics and Finance 52*(4), 402–412.

Billio, M., L. Ferrara, D. Guégan, and G. L. Mazzi (2013). Evaluation of Regime Switching Models for Real-Time Business Cycle Analysis of the Euro Area. *Journal of Forecasting 32*(7), 577–586.

Chauvet, M. (1998). An econometric characterization of business cycle dynamics with factor structure and regime switching. *International Economic Review 39*(4), 969–996.

Chauvet, M. and J. Piger (2008). Comparison of the real-time performance of business cycle dating methods. *Journal of Business and Economic Statistics 26*, 42–49.

Claessens, S., M. A. Kose, and M. E. Terrones (2012). How do business and financial cycles interact? *Journal of International Economics 87*(1), 178–190.

Clark, T. and M. W. McCracken (2011). Testing for unconditional predictive ability. In D. F. Hendry and M. Clements (Eds.), *Oxford Handbook on Economic Forecasting*, pp. 415–440. Oxford University Press.

Clark, T. and M. W. McCracken (2015). Nested forecast model comparisons: A new approach to testing equal accuracy. *Journal of Econometrics 186*(1), 160 – 177.

Claveria, O., E. Pons, and R. Ramos (2007). Business and consumer expectations and macroeconomic forecasts. *International Journal of Forecasting 23*(1), 47–69.

Croushore, D. (2006). *Forecasting with Real-Time Macroeconomic Data*, Volume 1 of *Handbook of Economic Forecasting*, Chapter 17, pp. 961–982. Elsevier.

Croushore, D. and T. Stark (2001). A real-time data set for macroeconomists. *Journal of Econometrics 105*(1), 111–130.

Darné, O. and L. Ferrara (2011). Identification of Slowdowns and Accelerations for the Euro Area Economy. *Oxford Bulletin of Economics and Statistics 73*(3), 335–364.

DeLong, E. R., D. M. DeLong, and D. L. Clarke-Pearson (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics 44*(3), 837–845.

Elliott, G. and R. P. Lieli (2013). Predicting binary outcomes. *Journal of Econometrics 174*(1), 15–26.

Estrella, A. and G. A. Hardouvelis (1991). The Term Structure as a Predictor of Real Economic Activity. *Journal of Finance 46*(2), 555–576.

Estrella, A. and F. S. Mishkin (1998). Predicting U.S. Recessions: Financial Variables As Leading Indicators. *Review of Economics and Statistics 80*(1), 45–61.

Evans, M. D. (2005). Where are we now? real-time estimates of the macro economy. *International Journal of Central Banking 1*(2), 127–175.

Ferrara, L. and D. Guégan (2005). Detection of the Industrial Business Cycle using SETAR Models. *Journal of Business Cycle Measurement and Analysis 2005*(3), 353–371.

Ghent, A. C. and M. T. Owyang (2010). Is housing the business cycle? Evidence from US cities. *Journal of Urban Economics 67*(3), 336–351.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics 55*(4), 665–676.

Hamilton, J. D. (1983). Oil and the Macroeconomy since World War II. *Journal of Political Economy 91*(2), 228–248.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica 57*, 357–384.

Hamilton, J. D. (1996). This is what happened to the oil price-macroeconomy relationship. *Journal of Monetary Economics 38*(2), 215–220.

Hamilton, J. D. (2011). Calling recessions in real time. *International Journal of Forecasting 27(4)*, 1006–1026.

Hansson, J., P. Jansson, and M. Löf (2005). Business survey data: Do they help in forecasting GDP growth? *International Journal of Forecasting 21*, 377–389.

Harding, D. and A. Pagan (2002). Dissecting the Cycle: A Methodological Investigation. *Journal of Monetary Economics 49*, 365–381.

Harding, D. and A. Pagan (2006). Synchronization of cycles. *Journal of Econometrics 132*(1), 59–79.

Iacoviello, M. and S. Neri (2010). Housing Market Spillovers: Evidence from an Estimated DSGE Model. *American Economic Journal: Macroeconomics 2*(2), 125–164.

Jordà, O., M. Schularick, and A. M. Taylor (2015). Betting the house. *Journal of International Economics 96*(S1), 2–18.

Jordà, O. and A. M. Taylor (2011). Performance Evaluation of Zero Net-Investment Strategies. NBER Working Papers 17150, National Bureau of Economic Research, Inc.

Jordà, O. and A. M. Taylor (2012). The carry trade and fundamentals: Nothing to fear but (FEER) itself. *Journal of International Economics 88*(1), 74–90.

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review 99*(3), 1053–1069.

Kydland, F. E., P. Rupert, and R. Sustek (2016). Housing Dynamics over the Business Cycle. *International Economic Review 57*(4), 1149–1177.

Leamer, E. E. (2007). Housing IS the business cycle. In *Housing, Housing Finance, and Monetary Policy*, pp. 149–233. Proceedings, Federal Reserve Bank of Kansas City.

Leamer, E. E. (2015). Housing really is the business cycle: What survives the lessons of 2008-09? *Journal of Money, Credit and Banking 47*(1), 43–50.

Liu, W. and E. Moench (2016). What predicts US recessions? *International Journal of Forecasting 32*(4), 1138–1150.

Marcellino, M. (2006). Leading indicators. In G. Elliott, C. W. J. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 1, pp. 879–960. Amsterdam: Elsevier.

Mian, A., K. Rao, and A. Sufi (2013). Household balance sheets, consumption, and the economic slump. *Quarterly Journal of Economics 128*(4), 1687–1726.

Mian, A. and A. Sufi (2014). What explains the 2007-2009 drop in employment? *Econometrica 82*(6), 2197–2223.

Potter, S. M. (1995). A Nonlinear Approach to U.S. GNP. *Journal of Applied Econometrics 10*, 109–125.

Ravazzolo, F. and P. Rothman (2013). Oil and U.S. GDP: A Real-Time Out-of-Sample Examination. *Journal of Money, Credit and Banking 45*(2-3), 449–463.

Ravazzolo, F. and P. Rothman (2016). Oil-price density forecasts of US GDP. *Studies in Nonlinear Dynamics & Econometrics 20*(4), 441–453.

Rudebusch, G. D. and J. C. Williams (2009). Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve. *Journal of Business & Economic Statistics 27*(4), 492–503.

Sarlin, P. (2013). On policymakers' loss functions and the evaluation of early warning systems. *Economics Letters 119*(1), 1–7.

Stock, J. H. and M. W. Watson (1989). New indexes of coincident and leading economic indicators. In *NBER Macroeconomics Annual 1989, Volume 4*, NBER Chapters, pp. 351–409. National Bureau of Economic Research, Inc.

Stock, J. H. and M. W. Watson (2003). Forecasting Output and Inflation: The Role of Asset Prices. *Journal of Economic Literature 41*(3), 788–829.

Stock, J. H. and M. W. Watson (2014). Estimating turning points using large data sets. *Journal of Econometrics 178*(P2), 368–381.

Tommaso, P. (1998). Characterizing Asymmetries in Business Cycles Using Smooth-Transition Structural Time-Series Models. *Studies in Nonlinear Dynamics & Econometrics 3*(3), 1–18.