Journal of Macroeconomics 47 (2016) 166–187

Contents lists available at ScienceDirect

Journal of Macroeconomics

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

Anticipating business-cycle turning points in real time using density forecasts from a VAR

Sven Schreiber[a,b,*], Natalia Soldatenkova[c]

[a] Macroeconomic Policy Institute (IMK), Hans-Böckler-Str. 39, 40476 Düsseldorf, Germany
[b] Free University of Berlin, Institute of Statistics and Econometrics, Germany
[c] University of Bonn, Bonn Graduate School of Economics (BGSE), Lennéstraße 35, 53113 Bonn, Germany

ARTICLE INFO

Article history:
Received 10 February 2015
Accepted 2 December 2015
Available online 23 December 2015

JEL Classification:
C53 (forecasting models)
E37 (cycle forecasting)

Keywords:
Density forecasts
Business-cycle turning points
Real-time data
Nowcasting
Bootstrap

ABSTRACT

For the timely detection of business-cycle turning points we suggest to use medium-sized linear systems (subset VARs with automated zero restrictions) to forecast monthly industrial production index publications one to several steps ahead, and to derive the probability of the turning point from the bootstrapped forecast density as the probability mass below (or above) a suitable threshold value. We show how this approach can be used in real time in the presence of data publication lags and how it can capture the part of the data revision process that is systematic. Out-of-sample evaluation exercises show that the method is competitive especially in the case of the US, while turning-point forecasts are in general more difficult in Germany.

1. Introduction

In this paper we suggest a linear system approach to the old problem of business-cycle turning-point prediction, taking into account the data availability and revision problems in real time.[1] This approach differs from the usual methods used to detect business-cycle turning points, which is usually done with non-linear models such as probit or Markov-switching methods.[2] The general idea is that we use a linear (system) model to predict the continuous output variable several steps ahead. We then use the estimated probability density function (pdf) of the forecast to calculate the probability of a realization below the previously defined recession threshold (or above a certain boom threshold). As described in more detail below, using monthly data we employ a threshold of a negative cumulated growth rate of -1% over a time span of six months to call a recession.[3]

* Corresponding author at: Macroeconomic Policy Institute (IMK), Hans-Böckler-Str. 39, 40476 Düsseldorf, Germany.
E-mail address: svetosch@gmx.net, Sven-Schreiber@boeckler.de (S. Schreiber).

[1] The subVAR applications were carried out with the *gretl* econometrics software (see Cottrell and Lucchetti, 2013, especially chapter 8 on real-time data), the model confidence set procedure was done with the MulCom package for Ox, and the comparison models were implemented in Matlab.

[2] While a subclass of Markov-switching models can actually be regarded as linear, the general case – for example with regime-dependent dynamics – yields non-linear models (see Krolzig, 1997).

[3] The averaging of consecutive months can be naturally interpreted according to the "triangle" approach by Harding and Pagan (2002), where episodes can be very short and intense or more drawn out and gradual to qualify as recessions.

http://dx.doi.org/10.1016/j.jmacro.2015.12.002
0164-0704/

Our approach has the following advantages: First, we can define the real-time variables in our multivariate system such that we also capture the revision process of consecutive data publications, by keeping (some of) the superseded data publications in the econometric system.[4] Secondly, compared to probit models there is reason to hope that the direct forecast of the continuous output variable is better able to exploit the information contained in the data. After all, in order to fit the probit model it is necessary to reduce the target variable to a binary regime variable, which discards quite a bit of information. Because of the linearity of the estimator our method is also computationally robust. Finally, as shown in the applications below, considering more than two regimes just means to partition the predictive density into more than two exogenously defined regions, which is straightforward.

There are also drawbacks of our approach which have to be acknowledged: In order to make use of a broad information set, we use relatively large VARs as the starting point for our forecasting model. These initial models are then reduced with automated coefficient restrictions following the general-to-specific method, but the initial models suffer from the curse of dimensionality, i.e. the combination of too many variables and lags may exceed the available degrees of freedom. In a scenario with only quarterly data and only short available revision data histories our approach may therefore not be the most suitable one. A more fundamental restriction is that our model presupposes linear time series processes. Thus if the DGP were actually non-linear, our forecasting models would only be approximations. On the other hand, the same variables are often analyzed with linear models in other macroeconomic contexts, and thus linear models seem to be perfectly reasonable.

Finally, our method is also affected by a turning-point recognition lag. If for example we receive in some publication period a recession signal based on a forecast h steps ahead (realistically assuming a moderate forecast horizon $h < 5 + p$, where p is the publication lag, i.e. the number of periods it takes before an initial data release happens), this means that the beginning of the recession actually happened in some reference period up to $-(h - p - 5)$ months ago, and so in reality the recession would likely be already underway. Although this may seem unfortunate, it is the logical consequence of the definition that a decline in economic activity must have a certain minimum duration to call it a recession.[5]

Related literature. An early example that linear prediction models can be applied to the problem of turning point determination with continuous target variables is given in Stock and Watson (1993). Österholm (2012) uses a similar approach as we do, in the sense of applying a linear model (Bayesian VAR, BVAR), and working with the predictive densities. We discuss and apply his approach in Section 2.

Non-linear turning point applications are an active area of research; in the domain of models with binary dependent variables the introduction of an explicitly dynamic probit setup by Kauppi and Saikkonen (2008) has spurred applied research, for some recent examples see Ng (2012), Nyberg (2010), and Proaño and Theobald (2014). The main alternative is Hamilton's Markov-switching approach, and some recent applications with real-time data are Hamilton (2011), Nalewaik (2012), or Theobald (2014).

Complete real-time datasets with various data vintages are not as readily available as standard final-release data. Therefore we start our study in the next section with a slightly longer pseudo real-time dataset of the USA for 1986–2013 where we take into account the data publication lags, but we do not include the revision information. With this simplified dataset we introduce our suggested approach and by means of out-of-sample forecasting we compare it with other established methods for the determination of turning points, namely a Markov-switching model, several dynamic probit specifications, and the BVAR mentioned above. After that section the principal workings of our method should be sufficiently clear and we proceed in Section 3 with the analysis of actual real-time data sets with revision information. The available samples in that case are 1993–2013 for the USA and 1995–2012 for Germany which are used for fully real-time applications of our approach in Section 4.

2. Pseudo real-time simulation: method and application to the USA

In Section 3 with the various actually published vintages of the data we explain more details of our proposed subset VAR (subVAR) approach. In contrast, in the present section we compare a simplified version of our approach with other methods that have been used in the literature to detect turning points. The simplification consists of using only final data releases (as of 2015) for all historical variables, not the vintages that were actually available in real time. However, the various econometric models still take into account the publication lags of the variables: for example, the posited information set for period t never includes the realization of industrial production in t, because a first release of that observation would only be available from $t + 1$ onwards. Our simplification in this section means that instead of using the first vintage published in $t + 1$ we use the revised value published much later (but still describing activity in period t, of course). We call this exercise a pseudo real-time evaluation, which for reasons of easy data availability we conduct in the sample 1986m2 through 2013m4, where the initial estimation sample ends in 1999m12. Throughout this section we work with variables that are either stationary or have been transformed to stationarity by forming (log) differences. See Section 3 for a discussion of how to use non-stationary levels in our proposed approach.

[4] See Corradi et al. (2009) for a discussion of the information content of data revisions.

[5] In probit or Markov-switching models a formal minimum-duration requirement is typically missing for the forecasts. Instead at estimation time a specification is chosen that somehow delivers reasonable regime classifications in sample, where "reasonable" usually also means that the regime episodes should not be too short.