Available online at www.sciencedirect.com
ScienceDirect
International Journal of Forecasting 27 (2011) 281–307
international journal of forecasting
www.elsevier.com/locate/ijforecast

Incorporating vintage differences and forecasts into Markov switching models

Jeremy J. Nalewaik*

Board of Governors of the Federal Reserve System, 5400 41st Street NW, 20015 Washington, DC, United States

Abstract

This paper incorporates vintage differences and forecasts into the Markov switching models described by Hamilton (1994). The vintage differences and forecasts induce parameter breaks close to the end of the sample, too close for standard maximum likelihood techniques to produce precise parameter estimates. A supplementary procedure estimates the statistical properties of the end-of-sample observations that behave differently from the rest, allowing inferred probabilities to reflect the breaks. Empirical results using real-time data show that these techniques improve the ability of a Markov switching model based on GDP and GDI to recognize the start of the 2001 recession.
Published by Elsevier B.V. on behalf of International Institute of Forecasters.

Keywords: Business cycles; Recession probabilities; Markov switching models; Real-time data analysis

1. Introduction

This paper discusses extensions of the standard Markov switching models which allow the estimated probabilities to reflect parameter breaks at or close to the end of the sample, too close for standard maximum likelihood techniques to produce precise parameter estimates. The basic technique is a supplementary estimation procedure, bringing additional information to bear on the estimation of the statistical properties of the end-of-sample observations which behave differently from the rest; the additional information required is the historical values of the end-of-sample observations over the entire length of the time series employed.

Table 1 illustrates. The second column contains a representation of the US GDP growth time series that was available after the June 2007 data release from the Bureau of Economic Analysis (BEA). Quarterly observations up to the end of 2005 have passed through a BEA annual revision, while the "current quarterly" estimates from 2006Q1 to 2007Q1 have not. Assume that these two types of estimates have different statistical properties, inducing a break in the time series at the known period T break point. The supplementary estimation procedure developed in this paper

* Tel.: +1 202 452 3792; fax: +1 202 872 4927.
E-mail address: jeremy.j.nalewaik@frb.gov.

0169-2070/$ - see front matter. Published by Elsevier B.V. on behalf of International Institute of Forecasters.
doi:10.1016/j.ijforecast.2010.05.017

Table 1
GDP growth time series as of June 2007, with a vintage break and a forecast appended.

Date	GDP growth	Data type	Procedure requires	
⋮	⋮	⋮	⋮	⋮
2005Q1	Δy_{T-3}	Annual revision	$\Delta y^{f_1}_{T-3}$	$\Delta y^{f_2}_{T-3}$
2005Q2	Δy_{T-2}	Annual revision	$\Delta y^{f_1}_{T-2}$	$\Delta y^{f_2}_{T-2}$
2005Q3	Δy_{T-1}	Annual revision	$\Delta y^{f_1}_{T-1}$	$\Delta y^{f_2}_{T-1}$
2005Q4	Δy_{T}	Annual revision	$\Delta y^{f_1}_{T}$	$\Delta y^{f_2}_{T}$
2006Q1	$\Delta y^{f_1}_{T+1}$	Current qtrly	$\Delta y^{f_1}_{T+1}$	$\Delta y^{f_2}_{T+1}$
2006Q2	$\Delta y^{f_1}_{T+2}$	Current qtrly	$\Delta y^{f_1}_{T+2}$	$\Delta y^{f_2}_{T+2}$
2006Q3	$\Delta y^{f_1}_{T+3}$	Current qtrly	$\Delta y^{f_1}_{T+3}$	$\Delta y^{f_2}_{T+3}$
2006Q4	$\Delta y^{f_1}_{T+4}$	Current qtrly	$\Delta y^{f_1}_{T+4}$	$\Delta y^{f_2}_{T+4}$
2007Q1	$\Delta y^{f_1}_{T+5}$	Current qtrly	$\Delta y^{f_1}_{T+5}$	$\Delta y^{f_2}_{T+5}$
2007Q2	$\Delta y^{f_2}_{T+6}$	Forecast		$\Delta y^{f_2}_{T+6}$

allows the post-break observations to be treated differently in the context of a Markov switching model, provided that a full history of the "current quarterly" estimates is available. While this may sound like an onerous requirement, databases such as the Federal Reserve Bank of Philadelphia's real-time dataset make it increasingly simple to meet.

In estimating the probability of a recession from the GDP growth time series available in June 2007, it would be standard practice to compute probabilities of recession up until 2007Q1 — i.e., the last quarter for which there is an official BEA estimate. However, by June 2007, a substantial amount of hard data relating to 2007Q2 is available, and thus a forecast based on these data may be quite informative. Table 1 appends such a forecast to the official GDP growth time series. The argument against doing this is clear: the forecast is likely to behave quite differently from the official BEA estimates.[1] Appending the forecast may indeed introduce a second break into this time series, but if a full time series of forecasts, constructed in the same manner as the 2007Q2 forecast, is available, this break may be treated in the same manner as the break introduced at time T by the vintage difference. The computation of a probability of recession for 2007Q2, incorporating the information from the forecast, then becomes feasible.

Table 1 appends a 1-quarter-ahead forecast to the official time series and stops there, but the techniques discussed in this paper allow k-period forecasts, an arbitrary distance ahead, to be appended to the series of interest. A large body of literature discusses forecasting binary indicators, usually NBER recessions, using financial variables; recently, Kauppi and Saikkonen (2008) discussed multi-period-ahead forecasting using lags of the binary indicator (see also Dueker, 2005). The procedures outlined here accomplish many of the same goals, allowing the incorporation of multi-period forecasts, with lagged states of the world influencing current and future states of the world through the Markov transition matrix. A key difference is that the states of the world are determined endogenously by the time series properties of the continuous variable of interest. In many cases, the state of the world over history is unobserved and must be determined by the model.

The supplementary estimation techniques developed here are computationally simple and highly intuitive. Many economists have computed the means and variances of variables in recessions and expansions,

[1] Interestingly, the same argument could be made against "appending" the "current quarterly" GDP growth estimates to the estimates that have passed through the annual revisions. Conceptually, the issue is exactly the same; the only difference is in the perceived severity of the breaks in the statistical properties of the time series.