

Kotchoni, Rachidi; Stevanovic, Dalibor

**Working Paper**

## Forecasting U.S. Recessions and Economic Activity

Document de travail, No. 2018-14

**Provided in Cooperation with:**

Department of Economics, School of Management Sciences (ESG UQAM), University of Quebec in Montreal

*Suggested Citation:* Kotchoni, Rachidi; Stevanovic, Dalibor (2018) : Forecasting U.S. Recessions and Economic Activity, Document de travail, No. 2018-14, Université du Québec à Montréal, École des sciences de la gestion (ESG UQAM), Département des sciences économiques, Montréal

This Version is available at:

<https://hdl.handle.net/10419/234769>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

**DOCUMENT DE TRAVAIL / WORKING PAPER**

**No. 2018-14**

**Forecasting U.S. Recessions and  
Economic Activity**

**Rachidi Kotchoni *et* Dalibor Stevanovic**

**Février 2018**

# Forecasting U.S. Recessions and Economic Activity

**Rachidi Kotchoni**, Economix-CNRS, Université Paris Nanterre, France  
**Dalibor Stevanovic**, Université du Québec à Montréal, Canada

**Document de travail No. 2018-14**

**Février 2018**

Département des Sciences Économiques  
Université du Québec à Montréal  
Case postale 8888,  
Succ. Centre-Ville  
Montréal, (Québec), H3C 3P8, Canada  
Courriel : [brisson.lorraine@uqam.ca](mailto:brisson.lorraine@uqam.ca)  
Site web : <http://economie.esg.uqam.ca>

Les documents de travail contiennent souvent des travaux préliminaires et/ou partiels. Ils sont publiés pour encourager et stimuler les discussions. Toute référence à ces documents de travail devrait tenir compte de leur caractère provisoire. Les opinions exprimées dans les documents de travail sont celles de leurs auteurs et elles ne reflètent pas nécessairement celles du Département des sciences économiques ou de l'ESG.

**Copyright (2018): Rachidi Kotchoni et Dalibor Stevanovic.** De courts extraits de texte peuvent être cités et reproduits sans permission explicite des auteurs à condition de référer au document de travail de manière appropriée.

# Forecasting U.S. Recessions and Economic Activity\*

Rachidi Kotchoni<sup>†</sup>

Dalibor Stevanovic<sup>‡</sup>

This version: February 13, 2018

## Abstract

This paper proposes a simple nonlinear framework to produce real-time multi-horizon forecasts of economic activity as well as conditional forecasts that depend on whether the horizon of interest belongs to a recessionary episode or not. Our forecasting models take the form of an autoregression that is augmented with either a probability of recession or an inverse Mills ratio. Our most parsimonious augmented autoregressive model delivers more accurate out-of-sample forecasts of GDP growth than the linear and nonlinear benchmark models considered, and this is particularly true during recessions. Our approach suits particularly well for the real-time prediction of final releases of economic series before they become available to policy makers. We use standard probit models to generate the Term Structure of recession probabilities at every period. Interestingly, the dynamic patterns of these Term Structures are informative about the business cycle turning points.

*JEL Classification: C35, C53, E27, E37*

*Keywords: Augmented Autoregressive Model, Conditional Forecasts, Economic Activity, Inverse Mills Ratio, Probit, Recession.*

---

\*We thank Frank Schorfheide for helpful comments and insightful guidance. Manuel Paquette-Dupuis has provided excellent research assistance. The second author acknowledges financial support from the Fonds de recherche sur la société et la culture (Québec) and the Social Sciences and Humanities Research Council.

<sup>†</sup>Economix-CNRS, Université Paris Nanterre. Email: rachidi.kotchoni@u-paris10.fr

<sup>‡</sup>Corresponding Author: Département des sciences économiques, Université du Québec à Montréal. 315, Ste-Catherine Est, Montréal, QC, H2X 3X2. Email: dstevanovic.econ@gmail.com

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Modeling the Economic Activity</b>	<b>4</b>
<b>3</b>	<b>Modeling the Probability of Recession</b>	<b>8</b>
<b>4</b>	<b>Predicting Turning Points in Real Time</b>	<b>10</b>
<b>5</b>	<b>Empirical Application</b>	<b>12</b>
5.1	Full-Sample Analysis . . . . .	12
5.1.1	Static vs. Dynamic Probit Models . . . . .	12
5.1.2	Term Structure of recession probabilities . . . . .	15
5.1.3	In-sample prediction of GDP growth . . . . .	15
5.2	Real-Time Out-of-Sample Analysis . . . . .	19
5.2.1	Predicting Turning Points in Real Time . . . . .	20
5.2.2	Forecasting GDP Growth . . . . .	22
5.3	Real-Time Analysis of the Latest Recessions of the U.S. Economy . . . . .	25
5.3.1	The Great Recession (2007-2009) . . . . .	26
5.3.2	The 2001 Recession . . . . .	30
<b>6</b>	<b>Conclusion</b>	<b>34</b>

# 1 Introduction

This paper proposes a framework to perform multi-horizon, real-time and nonlinear forecasting of macroeconomic activity variables (typically, the GDP growth). Our workhorse is an Augmented Autoregressive (AAR) model, which is a direct autoregressive (AR) model of order one augmented with probabilities of recession and/or Inverse Mills Ratios (IMR). The probability of a recession at a given horizon  $h$  is predicted conditionally on the information available (in real time) at period  $t$  using a Probit model. A forecast that is conditional on an expansion at the horizon of interest is termed optimistic and a forecast computed conditionally on a recession scenario is termed pessimistic. Methods are proposed to infer the business cycle turning points from the dynamics of the term structure of recession probabilities.

Optimistic and pessimistic forecasts can in principle be obtained by splitting the sample according to whether there is a recession or not, as done in an illustrative example presented in Hamilton (2011). Here, we follow an alternative approach that involves IMR corrections. Dueker (2005) and Dueker and Wesche (2005) propose a Qual VAR model, which is a VAR system that includes a latent variable that governs the occurrence of a binary outcome. This approach is not favored here because it does not naturally lead to state-dependent forecasts of economic activity in real time.<sup>1</sup> Our AAR model falls within the broad family of conditional forecasting models studied by Clark and McCracken (2013). This family includes all forecasting models that assume a particular policy path (e.g., announced inflation target) or a scenario for the future path of given macroeconomic variables (e.g., low inflation and high unemployment)<sup>2</sup>. Conditional forecasting models are used by major financial institutions and regulatory agencies worldwide to perform Stress Tests, see e.g. Grover and McCracken (2014). Typically, the goal of a Stress Testing exercise is to predict the impact of a more or less strong adverse shocks affecting one sector or the overall business environment on a particular outcome. The methodology developed in this paper can be useful in that context as

---

<sup>1</sup>Our paper assumes that recession dates are observed up to the most recent official turning point. Studies that attempt to predict the business cycle turning points include: [Chauvet (1998), Chauvet and Hamilton (2006), Chauvet and Piger (2008), Stock and Watson (2010), Berge and Jorda (2011) Stock and Watson (2012) or Ng (2014)]. Other studies attempt to identify the variables that lead future economic activity, e.g. [Stock and Watson (1989), Issler and Vahid (2006), Ng and Wright (2013)].

<sup>2</sup>For instance, Giannone et al. (2010) perform an inflation forecasting exercise conditional on pre-specified paths for oil price indicators. Schorfheide and Song (2013) produce inflation and growth forecasts conditional on forecasts that are obtained from judgmental sources. Other references on conditional forecasts include Sims (1982), Doan et al. (1984), Meyer and Zaman (2013) and Aastveit et al. (2014).

well. Indeed, our pessimistic forecast can serve as input for a wide range of Stress Testing models.

To implement our models empirically, we first need an operational definition of a recession. Obviously, a recession is a period running between a peak and the next through of economic activity while an expansion is a period between a through and the next peak.<sup>3</sup> We assume that the peaks and throughs of economic activity are observed (with a release lag) and that they coincide with the NBER dates.<sup>4</sup> Second, we need a model to predict the probability of a recession  $h$  periods ahead. Sophisticated models that account for structural breaks and state dependence in the dynamics of the probability of a recession could have been used, as for example in Chauvet and Potter (2002), Chauvet and Potter (2005). Here, we advocate the simplest Probit model that allows us to obtain parsimonious closed form expressions for our conditional forecasts.

Our empirical application starts with a set of *in-sample* performance evaluation exercises. First, we compare the in-sample performance of our static Probit model to that of three dynamic Probit models and two Markov Switching (MS) models. The results suggest that the dynamic specifications outperform the static Probit model only at horizons 1 and 2. Second, we estimate the static Probit model for the probability of recession using three conditioning information sets. The first information set contains the Term Spread (TS) only, the second contains the TS and the Federal Funds Rate (FFR) while the third further contains the Credit Spread (CS). We find that the Probit model that uses only TS as regressor compares favorably to those that use two or more regressors, especially at horizons  $h = 3$  and beyond. Third, we compare the AAR to a simple AR model, an Augmented Distributed Lag (ADL) model and a MS model in terms of their ability to forecast GDP growth a few quarters ahead. The ADL model is a version of the AAR model that implicitly assumes a linear structure for the probability of a recession. The AR model produces uninformative forecasts as soon as the horizon exceeds  $h = 3$ . The MS model does well at horizons 1 and 2 but its performance deteriorates fast as  $h$  increases. The ADL model is less and less resilient than the AAR model as the forecast horizon increases.

---

<sup>3</sup>The previous definition of a recession raises two practical issues. The first issue concerns the precise meaning of the expression “*economic activity*”. The Business Cycle Dating Committee of the National Bureau of Economic Research (NBER) does not provide a precise definition to this expression. Rather, it defines a recession as “*a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales.*” The second issue concerns the identification of the business cycle turning points (i.e., peaks and throughs of real economic activity) from the observed data. The Business Cycle Dating Committee does provide a precise response to the latter issue by regularly publishing recession dates with approximately one year lag.

<sup>4</sup>The business cycle dates can be found at <http://www.nber.org/cycles.html>.

We compare our model and the benchmarks above in a real-time out-of-sample forecasting exercise covering 1981Q1 - 2015Q4 period. Our AAR model outperforms the AR over the whole evaluation period and particularly during recessions. The performance is maximized at one-year horizon. The non-linearity of the Probit probability boosts the forecasting accuracy over the ADL model, but the improvement is tiny during NBER downturns. When compared to Markov switching models, our approach dominates uniformly over all forecast horizons during the whole evaluation period, and at horizons  $h = 3$  and  $h = 4$  quarters during recessions. Overall, the AAR model improves the forecast accuracy of GDP growth by up to 30% during recessions compared to the best nonlinear benchmark. Our method suits particularly well to produce real-time predictions of final releases of economic data before they become available, which is of great importance for policy makers.

Our finding that the predictability of economic series varies across the business cycle is not new in the literature. For instance, Rapach and Zhou (2010) and Gargano and Timmermann (2014) find that the predictability of the stock market and commodity prices is stronger during recessions. Leroux et al. (2017) document instabilities in the predictability of real activity variables, stock market returns, exchange rates and inflation growth. Chauvet and Potter (2013) use a methodology that is similar to ours and find an improvement in forecasting performance at short horizon. Our approach provides a simple and flexible nonlinear framework to forecast economic series in real time. Moreover, our first step Probit models captures recession signals quite well up to four quarters ahead. As a result, our AAR model improves the accuracy of GDP growth forecasts over the benchmark at longer horizons.

Finally, we use our model to conduct a real-time analysis of the last two recessions of the US economy.<sup>5</sup> Our results suggest that the AAR model delivers more accurate real-time forecasts than the benchmark models. We also find that the dynamics of the term structure of recession probabilities are quite informative about the business cycle turning points. Indeed, the shape of the term structure of recession probabilities switches from convex to concave before a recession and from concave to convex after a recession.

---

<sup>5</sup>First, we estimate a static Probit model for the probability of a recession and an AAR model for the GDP growth using a sample that stops at the latest official turning point before each recession. Second, the estimated parameters and the most recent release of GDP are plugged into the AAR model to obtain forecasts of the probability of a recession and of GDP growth rate at different horizons. Finally, benchmark models are estimated and their out-of-sample predictions compared to those of the AAR model.



The remainder of the paper is organized as follows. Section 2 details the construction of the AAR model. Section 3 motivates the static Probit model used for the probability of recessions and discusses alternative approaches. Section 4 presents our strategy to infer turning points from the term structure of the probability of recession in real time. Section 5 presents the empirical application, and section 6 concludes. A separate document contains supplementary material.

## 2 Modeling the Economic Activity

Let  $y_t$  denote an economic activity variable (e.g., GDP growth, unemployment rate, etc.),  $R_t \in \{0, 1\}$  the indicator of recession<sup>6</sup> at time  $t$  and  $X_t$  a set of potential predictors of recessions. Our main objective is to produce multi-horizon forecasts of the variables  $y_t$ . For that purpose, we consider using a family of Augmented AutoRegressive (AAR) models specified at a quarterly frequency.

The intended models normally take the form:

$$y_{t+h} = \rho_{h,0} + \rho_{h,1}y_t + \delta_h R_{t+h} + v_{t+h}, \quad (1)$$

for  $t = 1, \dots, T-h$ , where  $h \geq 1$  is the forecast horizon and  $v_{t+h} \sim N(0, \sigma_h^2)$  is a Gaussian error term. This error is assumed potentially correlated with  $R_{t+h}$  but uncorrelated with lagged realizations of  $y_t$ . Unfortunately, these models cannot be used for real time forecasting as the right hand side contains a regressor that is not yet observed at period  $t$ .

Taking the expectation of  $y_{t+h}$  conditional on the information available at time  $t$  yields:

$$E(y_{t+h}|y_t, X_t) = \rho_{h,0} + \rho_{h,1}y_t + \delta_h \Pr(R_{t+h} = 1|y_t, X_t).$$

Historical values of the economic activity variable  $y_t$  might have been used by the economists of the NBER to produce the series  $R_t$ . Therefore, there is a risk that a model that forecasts the probability of a recession at period  $t+h$  conditional on an information set that includes  $y_t$  be spuriously good in-sample and bad out-of-sample. To avoid this issue, we posit that the probability of a recession at period  $t+h$  depends on  $X_t$  only. Moreover, we advocate a functional form that

---

<sup>6</sup>That is,  $R_t = 1$  if the NBER dating committee designated period  $t$  as a recession time and  $R_t = 0$  otherwise.

leads to a Probit model:

$$\Pr(R_{t+h} = 1|y_t, X_t) = \Pr(R_{t+h} = 1|X_t) = \Phi(X_t\gamma_h), \quad (2)$$

where  $\Phi$  is the cumulative distribution function (CDF) of the standard normal random variable. Therefore, an equation that expresses the expected value of  $y_{t+h}$  in terms of quantities that depends on the information available at time  $t$  is given by:

$$E(y_{t+h}|y_t, X_t) = \rho_{h,0} + \rho_{h,1}y_t + \delta_h\Phi(X_t\gamma_h) \equiv \hat{y}_{t+h}, \quad (3)$$

Accordingly,  $y_{t+h}$  may be represented as an Augmented Autoregressive (AAR) process as follows:

$$y_{t+h} = \rho_{h,0} + \rho_{h,1}y_t + \delta_h\Phi(X_t\gamma_h) + \tilde{v}_{t+h}, \quad (4)$$

where  $\tilde{v}_{t+h} \equiv v_{t+h} + \delta_h(R_{t+h} - \Phi(X_t\gamma_h))$  is a zero mean error term.

Note that the forecasting formula (3) exploits the information content of  $X_t$  in a nonlinear manner. This suggests an alternative ADL model where  $X_t$  enters linearly in the right hand side. That is:

$$y_{t+h} = \rho_{h,0} + \rho_{h,1}y_t + X_t\beta_h + \varepsilon_{t+h}, \quad (5)$$

where  $\varepsilon_{t+h}$  is an error term. The model above implicitly assumes a linear structure for the probability of a recession. ADL models similar to this have been explored, among others, in Gilchrist et al. (2009), Chauvet and Potter (2013) and Ng and Wright (2013). The valued added of the AAR model (4) vis-à-vis the ADL model (5) is attributable to nonlinearity.

It is possible to go one step beyond the average forecast (3) by taking advantage of the observability of the recession dates ( $R_t$ ). Indeed, a forecast can be generated based on the pessimistic scenario that the economy will experience a recession at horizon  $t + h$ , i.e.:

$$E(y_{t+h}|y_t, X_t, R_{t+h} = 1) = \rho_{h,0} + \rho_{h,1}y_t + \bar{\delta}_h + \delta_{h,1}E(\tilde{v}_{t+h}|y_t, X_t, R_{t+h} = 1).$$

Under a joint normality assumption on  $v_{t+h}$  and the error term ( $u_{h,t}$ ) of the latent equation under-

lying the Probit (2), we obtain:

$$E(y_{t+h}|y_t, X_t, R_{t+h} = 1) = \rho_{h,0} + \rho_{h,1}y_t + \bar{\delta}_h + \delta_{h,1} \frac{\phi(X_t\gamma_h)}{\Phi(X_t\gamma_h)} = \underline{y}_{t+h}, \quad (6)$$

where  $\delta_{h,1} = Cov(u_{h,t}, v_{t+h}|R_{t+h} = 1)$ ,  $\phi$  is the probability distribution function (PDF) of the standard normal random variable,  $\bar{\delta}_h$  is a constant shift and  $\delta_{h,1} \frac{\phi(X_t\gamma_h)}{\Phi(X_t\gamma_h)}$  stems from a “break” in the structure of dependence between  $y_{t+h}$  and  $X_t$  due to the recession. Note that this break is absent when  $v_{t+h}$  is uncorrelated with recessions so that  $\delta_{h,1} = 0$ . The pessimistic forecast may be used to assess how severe a recession is expected to be if it were to effectively occur at the forecast horizon. This kind of formula can be used to perform a wide range of stress testing exercise in the banking sector, anticipate extreme losses on a portfolio or assess the fragility of the housing sector.

Likewise, another forecast based on the optimistic scenario of no recession at horizon  $t + h$  can be computed as:

$$E(y_{t+h}|y_t, X_t, R_{t+h} = 0) = \rho_{h,0} + \rho_{h,1}y_t + \delta_{h,0} \frac{-\phi(X_t\gamma_h)}{1 - \Phi(X_t\gamma_h)} = \bar{y}_{t+h}, \quad (7)$$

where  $\delta_{h,0} = Cov(u_{h,t}, v_{t+h}|R_{t+h} = 0)$  and  $\delta_{h,0} \frac{-\phi(X_t\gamma_h)}{1 - \Phi(X_t\gamma_h)}$  is a break that marks expansion periods. This optimistic forecast can be used to assess how favorable the economic conjuncture is expected to be if an expansion were to occur at the forecast horizon.

The variables  $\frac{\phi(X_t\gamma_h)}{\Phi(X_t\gamma_h)}$  and  $\frac{-\phi(X_t\gamma_h)}{1 - \Phi(X_t\gamma_h)}$  are the well-known IMRs. The parameters  $\bar{\delta}_h$ ,  $\delta_{h,0}$  and  $\delta_{h,1}$  are all expected to be negative if  $y_t$  is pro-cyclical (that is, if  $y_t$  increases during expansions and shrinks during recessions). In our framework, the terms  $\delta_{h,1} \frac{\phi(X_t\gamma_h)}{\Phi(X_t\gamma_h)}$  and  $\delta_{h,0} \frac{-\phi(X_t\gamma_h)}{1 - \Phi(X_t\gamma_h)}$  capture the combined effects of factors that are hard to measure such as supply and demand shocks, policy responses to these shocks, investors sentiments, consumer confidence, agents anticipations, etc.

Pooling the forecasting formulas (6) and (7) yields:

$$y_{t+h} = \rho_{h,0} + \rho_{h,1}y_t + \bar{\delta}_h R_{t+h} + \delta_{h,0} IMR_{t,h,0} + \delta_{h,1} IMR_{t,h,1} + \tilde{v}_{t+h}, \quad (8)$$

where  $\tilde{v}_{t+h}$  is a zero mean error term and:

$$\begin{aligned} IMR_{t,h,1} &= \begin{cases} \frac{\phi(X_t\gamma_h)}{\Phi(X_t\gamma_h)} & \text{if } R_{t+h} = 1, \\ 0 & \text{otherwise.} \end{cases}, \\ IMR_{t,h,0} &= \begin{cases} \frac{-\phi(X_t\gamma_h)}{1-\Phi(X_t\gamma_h)} & \text{if } R_{t+h} = 0, \\ 0 & \text{otherwise.} \end{cases}. \end{aligned}$$

To implement the AAR model empirically, we first estimate a Probit model for the probability of recessions to obtain  $\hat{\gamma}_h$ . This estimate is used to compute fitted values for the probability of recession  $\hat{P}_{t,h} = \Phi(X_t\hat{\gamma}_h)$  and for the IMRs  $\widehat{IMR}_{t,h,1}$  and  $\widehat{IMR}_{t,h,0}$ . The average forecasts are obtained as the fitted values of the following OLS regression:

$$y_{t+h} = \rho_{h,0} + \rho_{h,1}y_t + \delta_h\hat{P}_{t,h} + e_{t+h}, \quad (9)$$

where  $e_{t+h}$  is an error term. Finally, the parameters used to compute the optimistic and pessimistic forecasts are deduced from the following OLS regression:

$$y_{t+h} = \rho_{h,0} + \rho_{h,1}y_t + \bar{\delta}_h\hat{P}_{t,h} + \delta_{h,0}\widehat{IMR}_{t,h,0} + \delta_{h,1}\widehat{IMR}_{t,h,1} + \tilde{e}_{t+h}, \quad (10)$$

where  $\tilde{e}_{t+h}$  is an error term. Recall that  $R_{t+h}$  is replaced by  $\hat{P}_{t,h}$  above as a means to avoid endogeneity biases.

An alternative framework to produce state-dependent forecast of economic activity is provided by the Markov Switching (MS) model of Hamilton (1989). The simplest version of this model allows only the intercept to be state-dependent as follows:

$$y_{t+h} = \mu_{R_t} + \rho y_t + \varepsilon_{t+h} \quad (11)$$

where  $\varepsilon_{t+h} \sim N(0, \sigma_\varepsilon^2)$ . In a more flexible specification, the autoregressive root is allowed to be state-dependent as well:

$$y_{t+h} = \mu_{R_t} + \rho_{R_t}y_t + \varepsilon_{t+h} \quad (12)$$

It is further possible to let the variance of  $\varepsilon_{t+h}$  depend on  $R_t$ . However, we restrict (11) and (12)

to the homoskedastic case in our empirical applications.

Chauvet and Potter (2013) obtained a model that is similar to our AAR model by augmenting an autoregression with the probability of recession as predicted by a dynamic factor regime switching model. Our approach differs from the one of Chauvet and Potter (2013) in that we treat the recession indicator  $R_t$  as observed (up to a release lag) and we specify its probability as a function of the lags of other explanatory variables rather than the lags of  $R_t$  itself. Another important contribution of our approach is that it explicitly takes the endogeneity of recessions into account when computing the optimistic and pessimistic forecast. This endogeneity is reflected in the correlation between the latent variable of the Probit and the error term of the equation that is used to forecast the economic activity variables.

### 3 Modeling the Probability of Recession

In the previous section, we have chosen to model the probability of a recession using a static Probit for three reasons. First, this model has a structural flavor as it emerges naturally from assuming the existence of a latent lead indicator  $Z_{h,t}$  that takes the form:

$$Z_{h,t} = X_t \gamma_h + u_{h,t}, \text{ for all } t \text{ and } h, \quad (13)$$

with  $u_{h,t} \sim N(0, 1)$ , and which satisfies:

$$R_{t+h} = \begin{cases} 1 & \text{if } Z_{h,t} > 0, \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

Second, our optimistic and pessimistic forecasting formulas depends on the expressions of  $E(v_{t+h}|X_t, R_{t+h})$ . These expressions are easily calculated by assuming that  $(u_{h,t}, v_{t+h})$  are jointly Gaussian, for all  $h \geq 1$ .<sup>7</sup> The third argument in favor of the static Probit model is that it is transparent and easy to replicate.

The IMR terms resulting from the calculation of  $E(v_{t+h}|X_t, R_{t+h})$  *havetheusualinterpretationofHeckman(1976)* (14) *maybeviewedas an attempt to infer the behavior of the NBER dating committee from historical data. The AAR mo*

---

<sup>7</sup>Note that it is not clear how one would generate conditional forecasts analogue to (6) and (7) in the context of the ADL model.

The exercise which consists of predicting the probability of recessions is not new in the literature. Stock and Watson (1989) use a probabilistic framework to construct a coincident and a leading index of economic activity as well as a recession index. Estrella and Mishkin (1998) examine the individual performance of financial variables such as interest rates, spreads, stock prices and monetary aggregates at predicting the probability of a recession. They find that stock prices are good predictors of recessions at one to three quarters horizon while the slope of the yield curve is a better predictor beyond one quarter. The forecasting power of the yield curve is also documented in Rudebusch and Williams (2009), who find that professional forecasters do not properly incorporate the information from the yield spread. Nyberg (2010) advocate a dynamic Probit model and find that in addition to the TS, lagged values of stock returns and foreign spreads are important predictors of a recession. Anderson and Vahid (2001) apply nonlinear models to predict the probability of U.S. recession using the interest-rate spread and money stock (M2) growth. Wright (2006) estimates several Probit models and finds that adding the FFR to the TS outperforms the model of Estrella and Mishkin (1998) that used the TS only. Christiansen et al. (2013) find that sentiment variables have predictive power beyond standard financial series.

Kauppi and Saikkonen (2008) consider three different dynamics for the latent lead indicator  $Z_{h,t}$ . The first specification below leads to the simplest *Dynamic Probit* model:

$$Z_{h,t+h} = X_t\gamma_{1,h} + R_t\gamma_{2,h} + u_{h,t}, \text{ for all } t. \quad (15)$$

The second specification, the *Autoregressive Probit* model, is given by:

$$Z_{h,t+h} = X_t\gamma_{1,h} + Z_{h,t}\gamma_{2,h} + u_{h,t}, \text{ for all } t. \quad (16)$$

Finally, the third specification gives rise to the *Dynamic Autoregressive Probit* model:

$$Z_{h,t+h} = X_t\gamma_{1,h} + R_t\gamma_{2,h} + Z_{h,t}\gamma_{3,h} + u_{h,t}, \text{ for all } t. \quad (17)$$

These specifications incorporate the inertia of  $Z_{h,t}$  and  $R_t$  when forecasting future recessions.

Hao and Ng (2011) find that dynamic Probit models improve upon the static Probit, especially when predicting the duration of recessions. This result is not surprising, in particular at short

horizons and around turning points. However, the dynamic feature of these models makes them unsuitable for a real-time forecasting exercise as  $R_t$  is usually observed with at least one-year lag. A static Probit that uses financial predictors released at high frequency does not suffer from this shortcoming.

## 4 Predicting Turning Points in Real Time

There is a difference between the prediction of the probability of a recession and the prediction of the beginning and end of a recession. The latter exercise is slightly more difficult as it requires decision science tools in addition to a probabilistic model. This section discusses how to infer turning points from the predicted probabilities of the recession in real time.

At a quarterly frequency, the first release of GDP is available with one lag while the ‘final’ value is released with approximately one year lag.<sup>8</sup> The NBER turning points are released with at least four lags. These aspects may be ignored if we are interested only in assessing the *in-sample* performance of the Probit and AAR models based on historical data. However, a strategy to deal with release lags is needed if one wishes to conduct a real time analysis.

If the current period is  $t^*$  and the latest turning point occurred at period  $t^* - l$ , then the final releases of NBER recession dates are available in real time only up to period  $t^* - l$ . Therefore, we can use fully revised data covering the periods  $[1, t^* - l]$  to estimate the probability of recessions at any horizon  $h$ . We have:

$$\Pr(R_{t+h} = 1|X_t) = \Phi(X_t\gamma_h), \quad t \in [1, t^* - l - h].$$

The estimate  $\hat{\gamma}_h$  of  $\gamma_h$  obtained from above can be used to generate out-of-sample forecasts of the probability of recession. As we choose to include only high frequency financial variables in  $X_t$ , this out-of sample exercise does not suffer from release lag issues. We therefore can compute  $\hat{P}_{t,h} = \Phi(X_t\hat{\gamma}_h)$  as well as the variables  $\widehat{IMR}_{t,h,0}$  and  $\widehat{IMR}_{t,h,1}$  for periods  $t \in [1, t^*]$ . Note that the out-of-sample periods run from  $t^* - l - h + 1$  to  $t^*$  for this Probit.

The next step is to estimate the AAR model for an economic activity variable based on the

---

<sup>8</sup>In the realm of real time data, the “final value” of a variable is a release that is unlikely to be revised in the future. Strictly speaking, there is actually no final value.

available information. At a quarterly frequency, the first release of economic activity variables is generally available with only one lag. Nonetheless, we constrain the in-sample period to be the same as for the Probit model. That is, we estimate Equations (9) and (10) by OLS using the sample covering the periods  $t \in [1, t^* - l - h]$ . At time  $t^*$ , the most recent release of the GDP growth is for the period  $t^* - 1$ . Equations (3), (6) and (7) take this latest release and the estimates  $\hat{\gamma}_h$  as input to return nowcasts ( $h = 1$ ) and forecasts ( $h > 1$ ) of economic activity. A similar strategy is employed for the ADL and MS models.

Using our static Probit model, we can calculate the term structure of the probability of recession at a given period  $t$  as the mapping  $P_t : h \mapsto \Phi(X_t \hat{\gamma}_h)$ ,  $h \geq 1$ . As we move forward from period  $t$  to periods  $t + 1$ ,  $t + 2$ , etc., the term structure of recession probabilities is updated to  $P_{t+1}$ ,  $P_{t+2}$ , etc. Our empirical experiments show that the sequence  $P_t, t > 1$  is clustered into successive blocs of convex and concave curves. This suggests two possible strategies to identify turning points.

The first strategy relies on the upper envelope of the concave blocks and the lower envelope of the convex blocs. Suppose that at period  $t$  the term structure of recession probabilities  $P_t$  is concave. At that period, the next business cycle peak is predicted to occur at  $t + h_t$ , where  $h_t$  is the horizon where  $P_t$  is maximized. If  $P_t, P_{t+1}, \dots, P_{t+H}, H \geq 1$  is a block of concave term structure of recession probabilities, then we can compute an upper envelope curve for this bloc and predict the beginning of the next recession as the maximum of this curve. Business cycle troughs are predicted similarly. If  $P_t$  is convex, then the next business cycle through is expected to occur at  $t + \tau_t$ , where  $\tau_t$  is the horizon that minimizes  $P_t$ . Considering a bloc  $P_t, P_{t+1}, \dots, P_{t+L}, L \geq 1$  of convex term structure of recession probabilities, we can compute a lower envelope curve for this bloc and predict the end of the next recession as the minimum of this curve.

The second strategy relies on the timing of the changes in the shape of the term structure of recession probabilities. Indeed, a convex term structure curve of recession probabilities suggests that recession is less and less likely for some time. If this curve suddenly switches from convex to concave, this suggests that a new signal that raises the prospects of a recession just came in. One might therefore want to predict the beginning of the next recession as  $t + h_t$ , where  $h_t$  is the horizon that maximizes  $P_t$ , and  $P_t$  is the beginning of a concave block. Likewise, the end of a recession may be predicted as  $t + \tau_t$ , where  $\tau_t$  is the horizon that minimizes  $P_t$  and  $P_t$  is the beginning of a convex block. With these two complementary strategies to identify turning points in hand, we



analyze below the last two recessions that occurred in the US in order to see whether they were predictable.

## 5 Empirical Application

For this application, we use the quarterly NBER recession indicator available in the FRED2 database. Data on TS, CS and FFR are also obtained from the same source.<sup>9</sup> The real time vintages of GDP data are obtained from Federal Reserve Bank of Philadelphia real-time data sets for macroeconomists. The time span starts in 1959Q1 and ends in 2016Q4. We consider three different designs for  $X_t$ . In the first design,  $X_t$  is restricted to contain TS only. In the second design,  $X_t$  contains TS and CS. In the third design,  $X_t$  contains TS, CS as well as FFR. It is found that the addition of CS and FFR to TS generally adds little to the predictive power of our Probit models, especially at horizons beyond  $h = 2$  quarters. Therefore, most of the results presented here are for the case where  $X_t$  reduces to TS. Additional results are in the supplementary material.

### 5.1 Full-Sample Analysis

This section presents in-sample forecasts based on models that are estimated on the full sample. The sample used here consists of historical data. Accordingly, release lag issues are ignored.<sup>10</sup>

#### 5.1.1 Static vs. Dynamic Probit Models

Figure 1 compares the in-sample fits of the static Probit model (2) to those of the dynamic Probit models (15), (16) and (17). All models are estimated using the TS only. At horizons 1 and 2, the static Probit model predicts lower probabilities of recession during recession times and higher probabilities of recession during expansion times than the dynamic Probit models. At longer horizons, the advantage of the dynamic Probit models erodes so that there is no visible difference between the performances of all four models. These results supports that the dynamic Probit models provide a more reliable signal only when recession is imminent. This is expected because

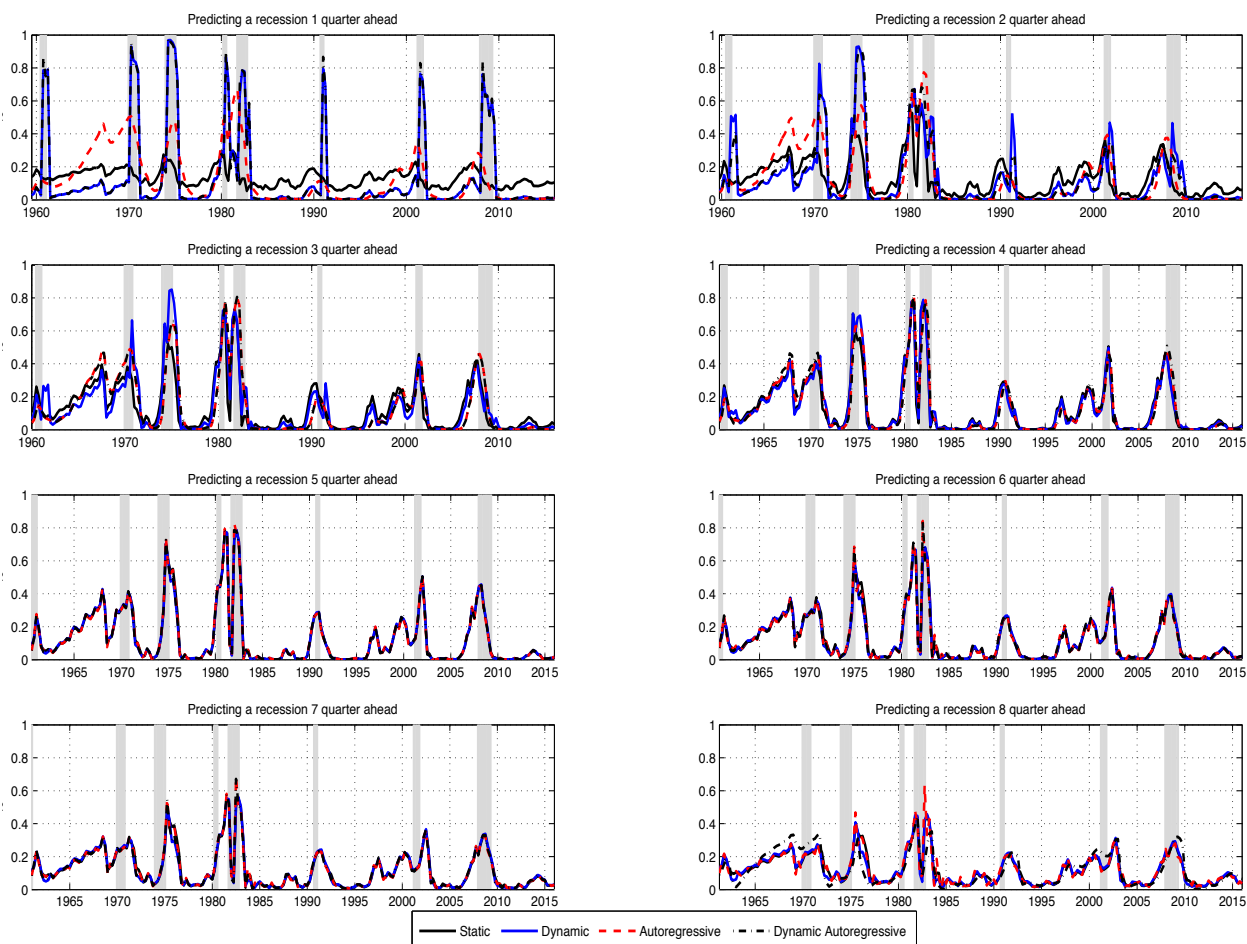
---

<sup>9</sup>The TS is defined as difference between a 10-Year Treasury Constant Maturity Rate (labelled GS10 in FRED2) and a 3-Month Treasury Bill: Secondary Market Rate (TB3MS). The credit spread (BAA10YM) is Moody's Seasoned Baa Corporate Bond Yield Relative to Yield on 10-Year Treasury Constant Maturity.

<sup>10</sup>Hence, the full sample ends on 2016Q4 and contains only the most recent values of GDP, as obtained from the 2017Q1 vintage.

recession episodes do not last enough to allow full exploitation of the potential of the dynamic Probit models at the longer forecast horizons.

Figure 1: Predicting the probability of a recession: In-sample performance of static and non-static Probit models

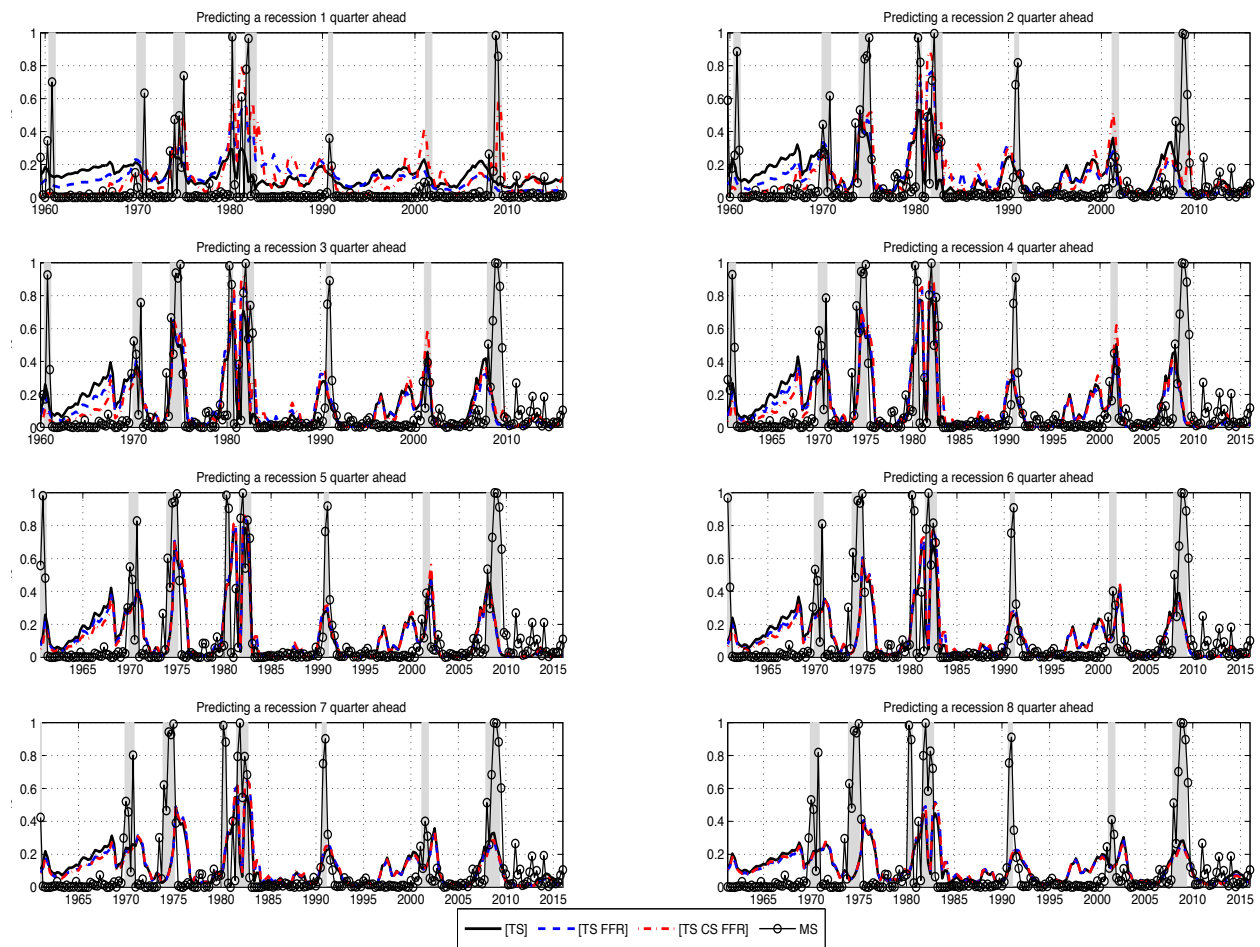


This figure shows in-sample predictions of the probability of a recession obtained from the static, dynamic, autoregressive and dynamic-autoregressive Probit models with the Term Spread only as regressor.

Figure 2 compares the performance of the static Probit model across different conditioning information sets. As said previously, the first Probit model is estimated using TS only as regressor. The second model is estimated using TS and FFR while the third model combines TS, CS and FFR. The figure also shows the filtered probabilities of recession predicted by the MS model with state-dependent intercept. We see that these filtered probabilities fit the data better than the Probit models at all horizons. However, this impressive in-sample capability of the MS model is deceptive about its out-of sample performance.

At short horizons, the addition of FFR and CS seems to improve the performance of the static Probit model. This is particularly visible around 1965 where there was no recession while the probability of recession predicted using TS only is higher than the ones predicted based on the two other information sets. However, the predictions of all three static Probit models become more and more similar as the forecasting horizon increase. If we abstract from the beginning of the sample, the differences between the predicted probabilities of the Probit models are indeed negligible at horizons  $h \geq 3$ .

Figure 2: Predicting the probability of a recession: In-sample performance of the static Probit model across different information sets



This figure shows in-sample predictions of the probability of a recession obtained from the static Probit model conditioned on three different information sets: [TS], [TS FFR] and [TS CS FFR]. MS stands for the filtered probabilities obtained from the Markov Switching model with state-dependent intercept.

Overall, the results shown in Figures 1 and 2 suggest that the static Probit model estimated with TS alone is a good benchmark, especially if our focus is on horizons 3 quarters and beyond.

The latter model is therefore favored because of its parsimony, its convenience for real time analysis and its robustness in out-of-sample analyses.

### 5.1.2 Term Structure of recession probabilities

Figure 3 shows a 3-D plot of the term structures of the probabilities of recession. The two horizontal axes are respectively the time stamp of the information set used to compute the term structure curves and the horizons to which the predicted probabilities of recessions belong. We see that this figure consists of clusters of concave and convex curves. Indeed, the term structure curves are concave in the neighbourhood of and during recessions while they are convex during expansions. The shape of the term structure curves conveys a more reliable signal than its level about the prospects of a recession. This claim is better illustrated by other 2-D plots that are shown in subsequent analyses.

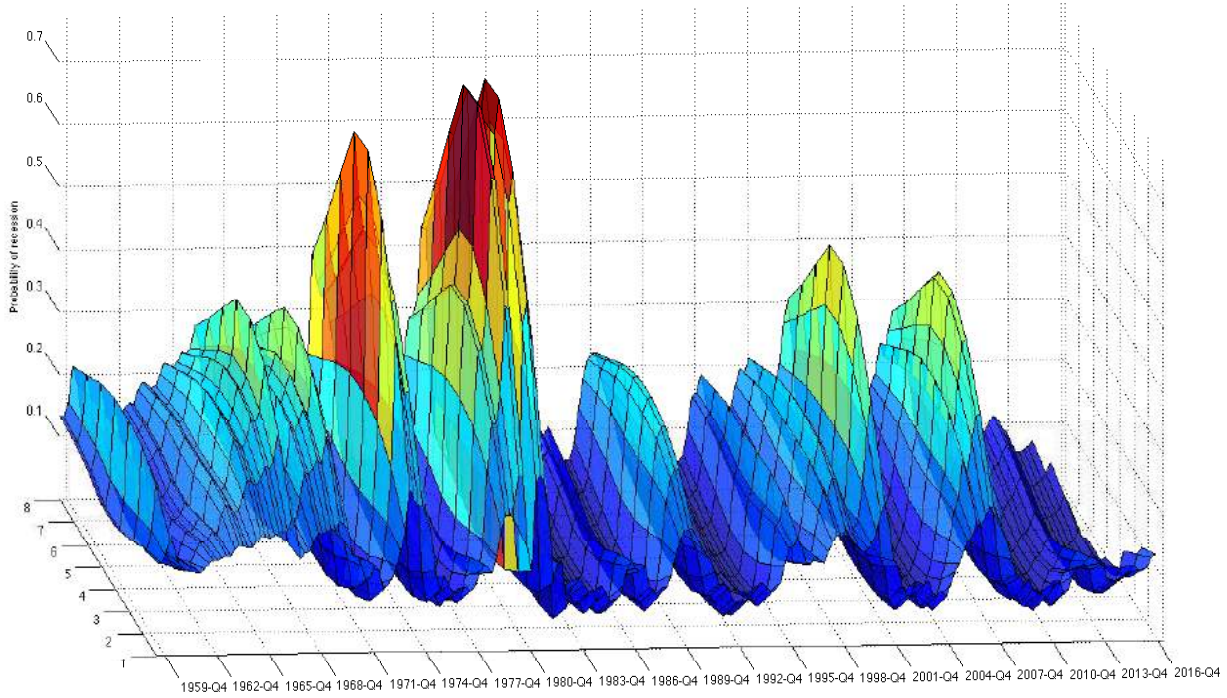
One possible approach to reduce the dimensionality of the information contained in this 3-D plot is to summarize each term structure curve into a single number measuring its shape: the Term Spread of recession probabilities. Here, we consider the spreads obtained by respectively taking the probabilities of recession three, four and five quarters ahead minus the probability of recession one quarter ahead. Figure 4 shows the results. We see that the beginning of each recession is immediately preceded by a large peak in the Term Spread curves.

A few peaks of the Term Spread of recession probabilities are observed at periods that have not been declared recessionary by the NBER. Such peaks may be indicating short episodes during which the economy underperformed or recessions that have been avoided due to timely and adequate policy responses. It is interesting to note that the Term Spread of recession probabilities exhibits no significant peak since 2010Q4.

### 5.1.3 In-sample prediction of GDP growth

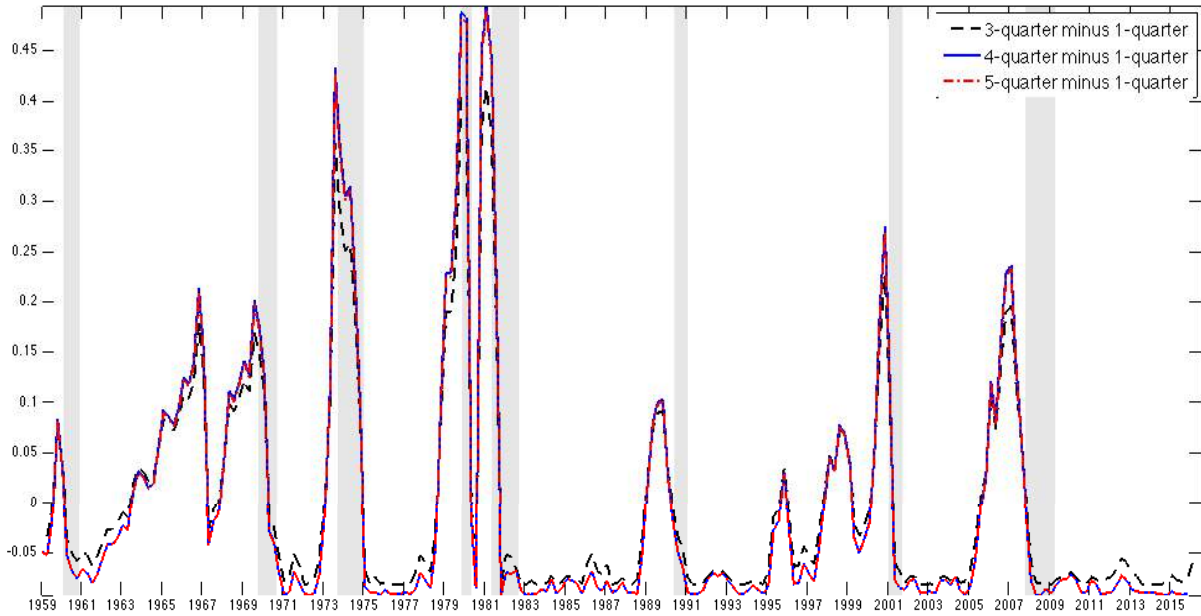
We now compare the performance of the AAR model to that of the benchmark models (namely the AR, ADL and MS) at predicting GDP growth. Figure 5 shows the adjusted R-squared of the AR, ADL and AAR models on the left vertical axis and the Student t-stat associated with  $\hat{\delta}_h$  on the right vertical axis. Recall that the AAR models reduces to an AR model when  $\delta_h = 0$ . We see that the adjusted R-squared is larger for the AAR model than for the AR at horizons  $h = 1$  to  $h = 7$ . Accordingly, the parameter  $\delta_h$  is estimated to be significant for these lags (i.e., t-stat larger than

Figure 3: Full-sample Term Structure of recession probabilities



This figure plots the full-sample Term Structure of recession probabilities, 1 to 8 quarters ahead, from static Probit model having the term spread as the only predictor. The periods correspond to information set when the forecasts have been constructed.

Figure 4: Full-sample Term Spread of recession probabilities

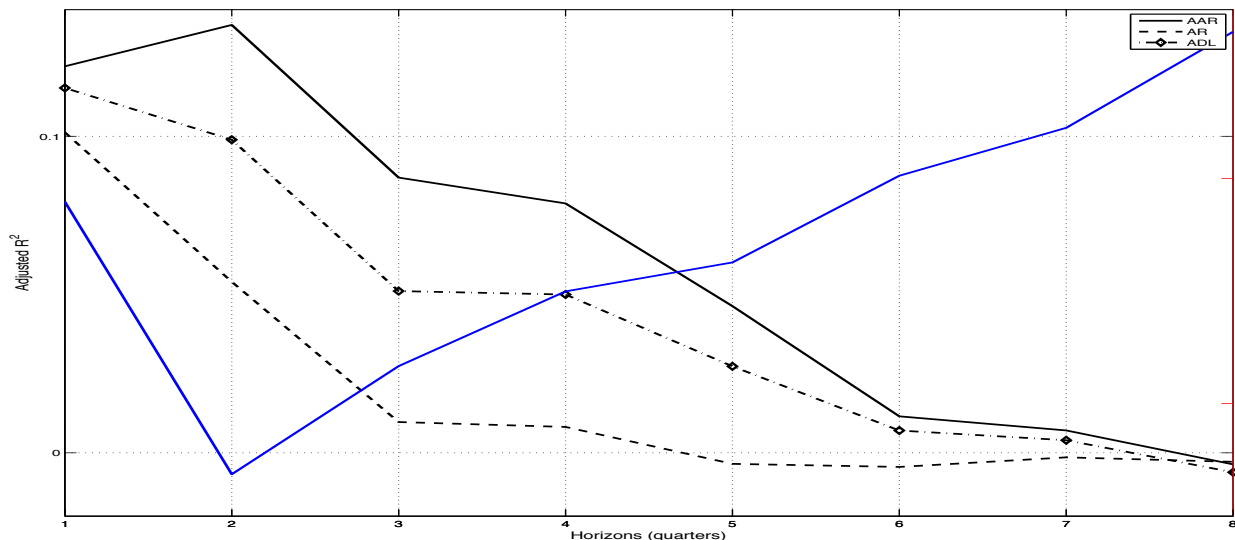


This figure shows the full-sample Term Spread of recession probabilities from static Probit model having the term spread as the only predictor. 3-quarter minus 1-quarter stands for 3-quarter minus 1-quarter ahead forecasted recession probabilities.

2 in absolute value). The gap between the adjusted R-squares of the two models decreases with  $h$  and the AR model underperforms the historical average at lags beyond  $h = 4$ .

The ADL model fits that data better than the AR model but underperforms the AAR model. This suggests that the nonlinear transformation applied to  $X_t$  prior to its inclusion in the right hand side of Equation (4) matters. Putting it differently, the probability of recession at a given horizon is a relevant predictor of GDP growth at that horizon.

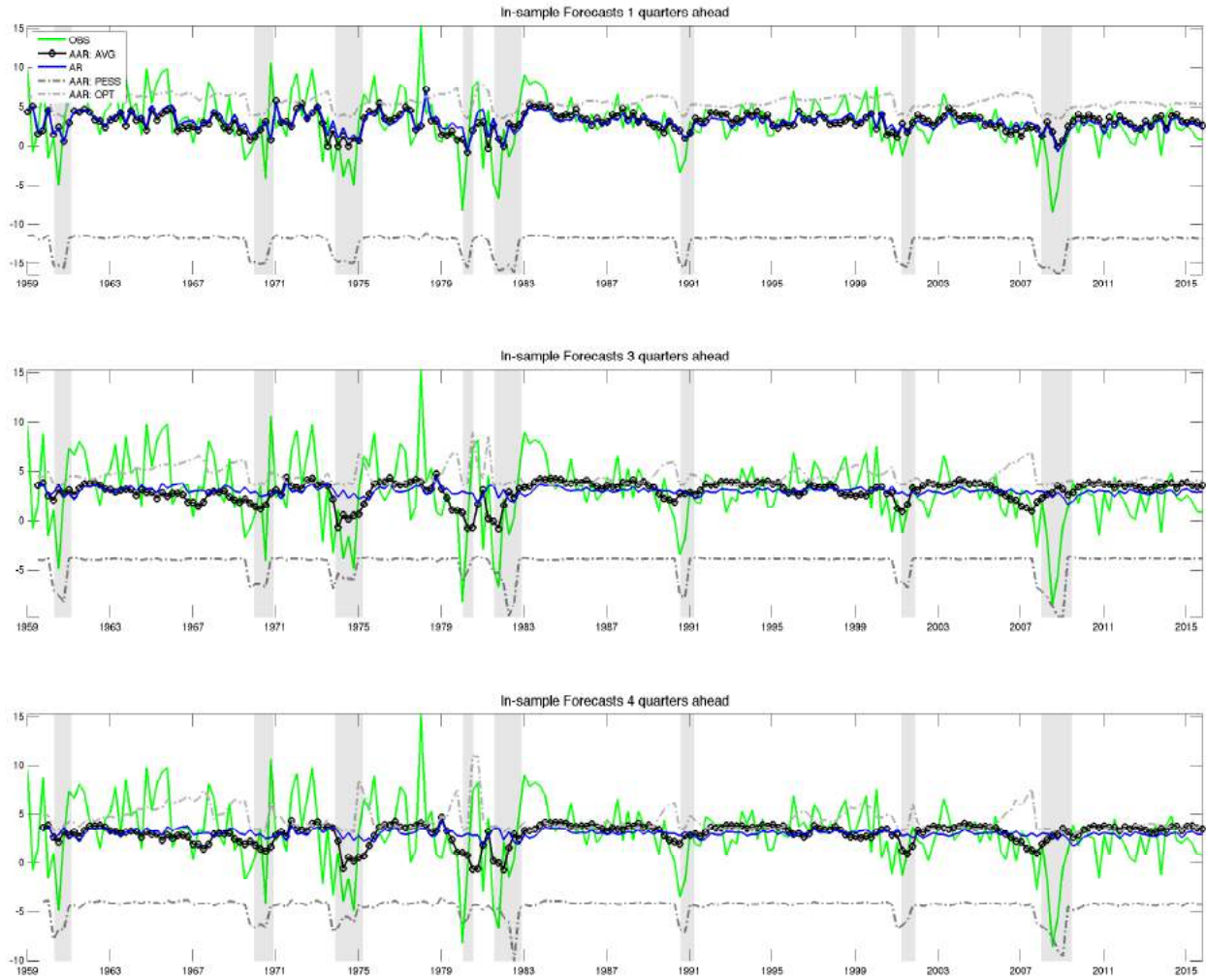
Figure 5: Predicting GDP growth: In-sample goodness-of-fit



This figure shows the adjusted  $R^2$  of the AAR, AR and ADL models on left vertical axis (full black line and dotted lines respectively), and the Student  $t$ -stat associated with  $\hat{\delta}_{i,h}$  on the right vertical axis. The probabilities of recession have been estimated from the static Probit model conditioned on TS only.

Figure 6 shows the in-sample fit of the AR and AAR models as well as the optimistic and pessimistic forecasts. We note that the trajectory of fitted values produced by the AR model becomes roughly flat beyond horizon  $h = 3$ . At a glance, the AAR model clearly dominates the AR model at horizons  $h \geq 3$ . This is consistent with the term structure of adjusted R-squares seen on Figure 5. The optimistic forecast is not optimistic enough as it often falls below the actual realizations of GDP growth. The pessimistic scenario is too pessimistic at horizon 1 and more realistic at longer horizons. One plausible explanation of this asymmetric pattern is that the state  $R_t = 0$  is actually a “no recession” state, which merges both stagnation and expansion.

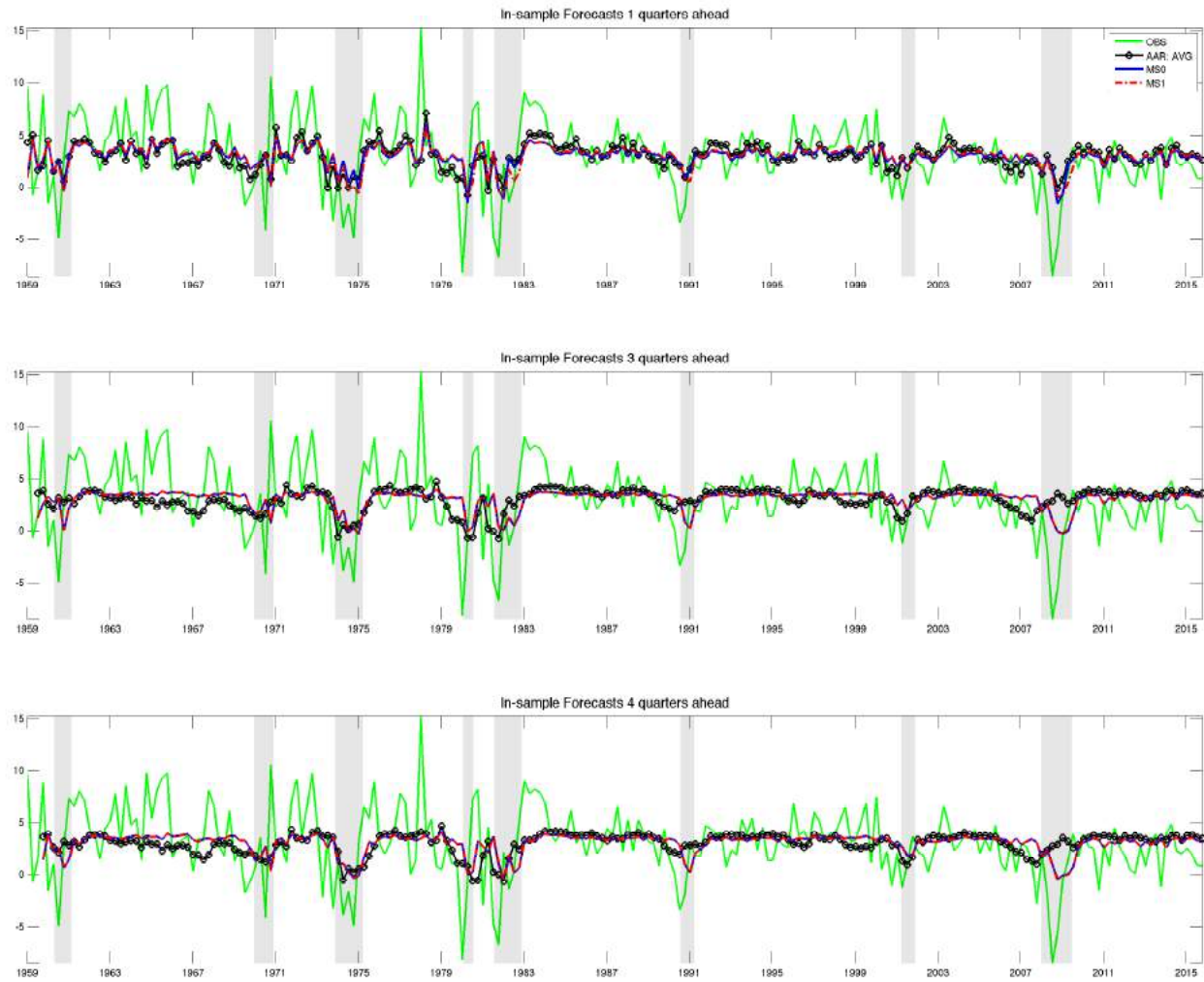
Figure 6: In-sample predictions of GDP growth: AAR vs. AR



This figure shows 1, 3 and 4 quarters ahead in-sample forecasts of GDP growth obtained from the AR and AAR models as well as the optimistic and pessimistic forecast scenarios. Grey areas indicate NBER recessions.

Figure 7 compares the trajectory of fitted values for the AAR and MS models. First, we note that the MS model that allows both the intercept and the autoregressive root to be state-dependent does not outperform its more parsimonious version that constrains  $\rho$  to be constant across states. Second, the in-sample fit of the MS models is better than that of the AAR during certain recessions (e.g., the great recession of 2007-2009) and worse during others recessions. Third, trajectories of fitted values obtained from the MS models are roughly flat during expansion periods as soon as the forecast horizon exceeds  $h = 3$ . Indeed, the MS models inherit the limitations of the AR model identified on Figure 5 so that the autoregressive coefficients are not significantly different from zero at long horizons.

Figure 7: In-sample predictions of GDP: AAR vs. MS



*This figure compares the trajectories of 1, 3 and 4 quarters ahead in-sample forecasts of GDP growth obtained from the AAR and two Markov Switching models models: MS0 and MS1 (11) and (12) respectively.*

## 5.2 Real-Time Out-of-Sample Analysis

We now explore the performance of our forecasting strategies in a real-time out-of-sample forecast exercise. The OOS evaluation period spans 1981:Q4-2015:Q4. We stop the sample at the end of 2015 in order to have all data releases available. First, we re-examine the performance of our method to predict turning points based on term structures of recession probabilities that are computed out-of-sample. Next, we examine the performance of the AAR model at predicting GDP growth. Finally, we perform an in-depth analysis of the last two recessions of the US economy.

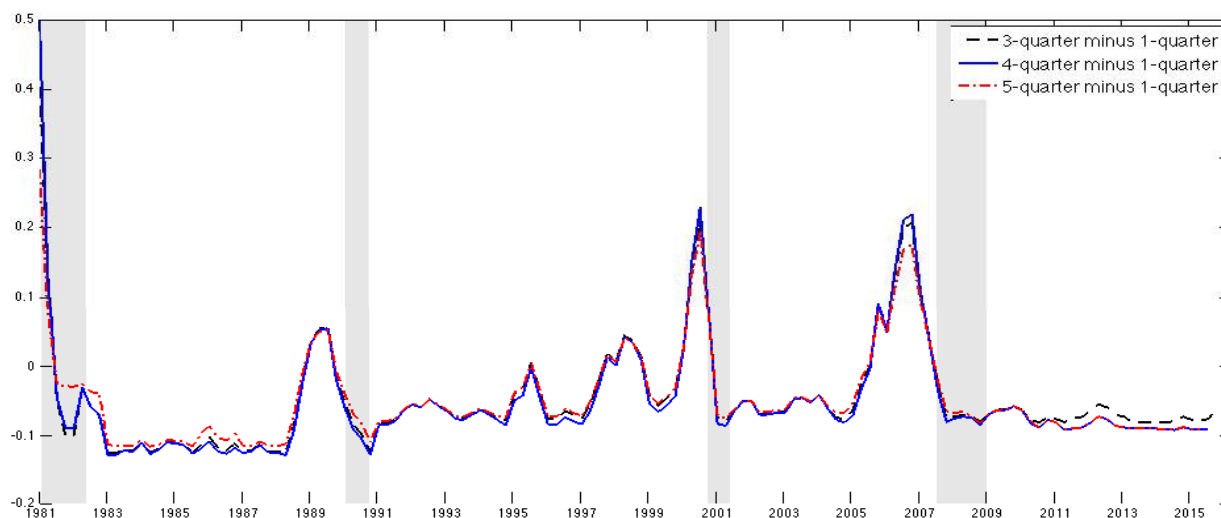


### 5.2.1 Predicting Turning Points in Real Time

Figure 3 and Figure 4 were computed using in-sample predictions of models that are estimated using final releases. We redo the same exercise in real time. That is, we estimate probit models for the probability of recession using a sample that stops at the latest official NBER turning point. The parameter estimates obtained from these models are then used to make out-of-sample predictions. Figure 8 shows the out-of-sample predictions of the Term Spread of recession probabilities over the whole out-of-sample period. We see that the results are qualitatively similar to what we have found in the in-sample analysis. Recessions are always announced by a large peak in the term spread of recession probabilities. Quantitatively, the out-of-sample predictions of the term spreads tends to be uniformly lower than their in-sample counterparts.

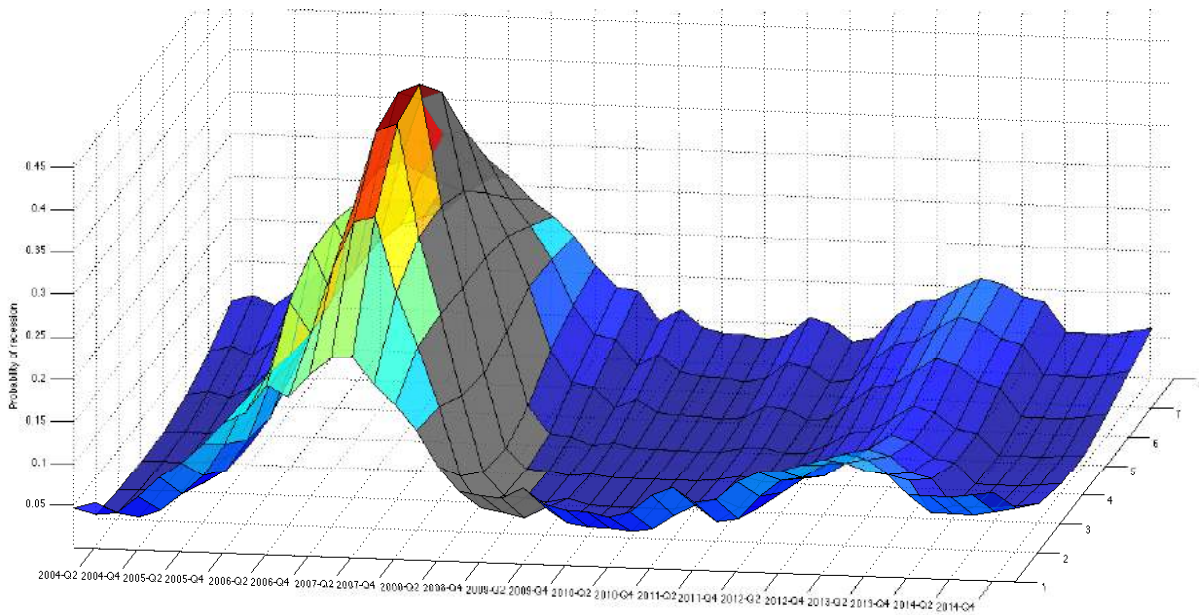
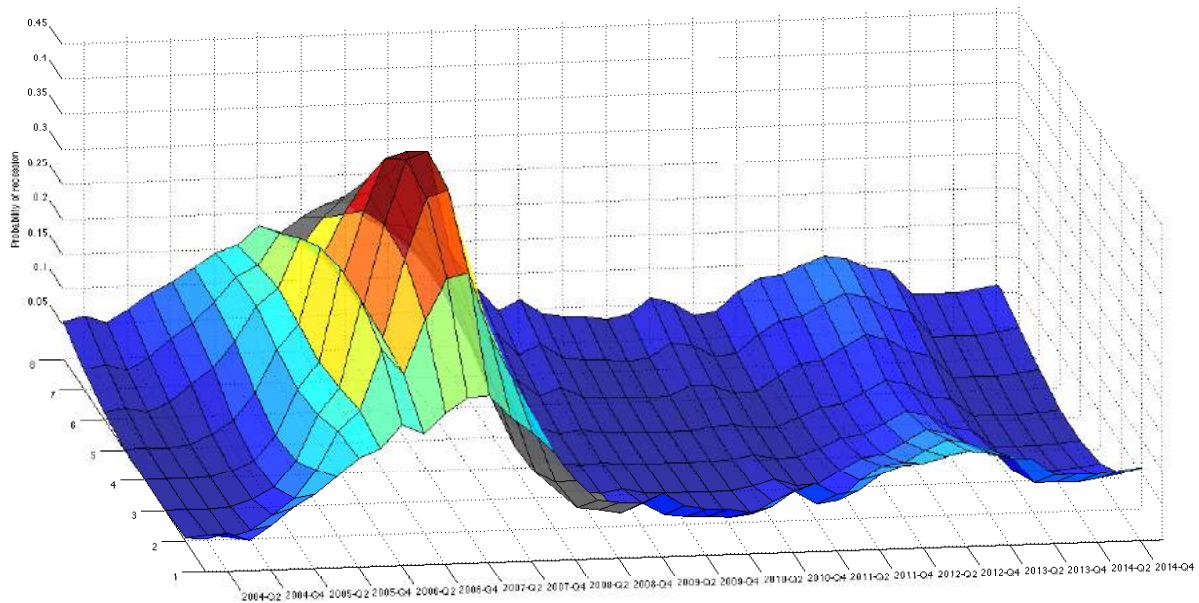
Figure 9 shows a 3-D plot of the term structure of recession probabilities that focuses on the most recent recession. Clearly, the switching of the term structure curve from convex to concave has been a warning sign of the recession. Likewise, as the recession approached its end, the term structure curve moved slowly from concave to convex. The level of the term structure curves have been changing since 2010, but its shape remains convex. The fact that the out-of-sample predictions of the term structure of recession probabilities behaves qualitatively as the in-sample predictions is quite reassuring but not surprising. This is attributable to the parsimonious parameterization of the probit models on which the predictions are based.

Figure 8: Out-of-sample Term Spread of recession probabilities



*This figure shows the out-of-sample Term Spread of recession probabilities from static Probit model having the term spread as the only predictor. 3-quarter minus 1-quarter stands for 3-quarter minus 1-quarter ahead forecasted recession probabilities.*

Figure 9: Out-of-sample Term Structure of recession probabilities



*This figure plots the out-of-sample Term Structure of recession probabilities, 1 to 8 quarters ahead, from static Probit model having the term spread as the only predictor. The periods correspond to information set when the forecasts have been constructed.*

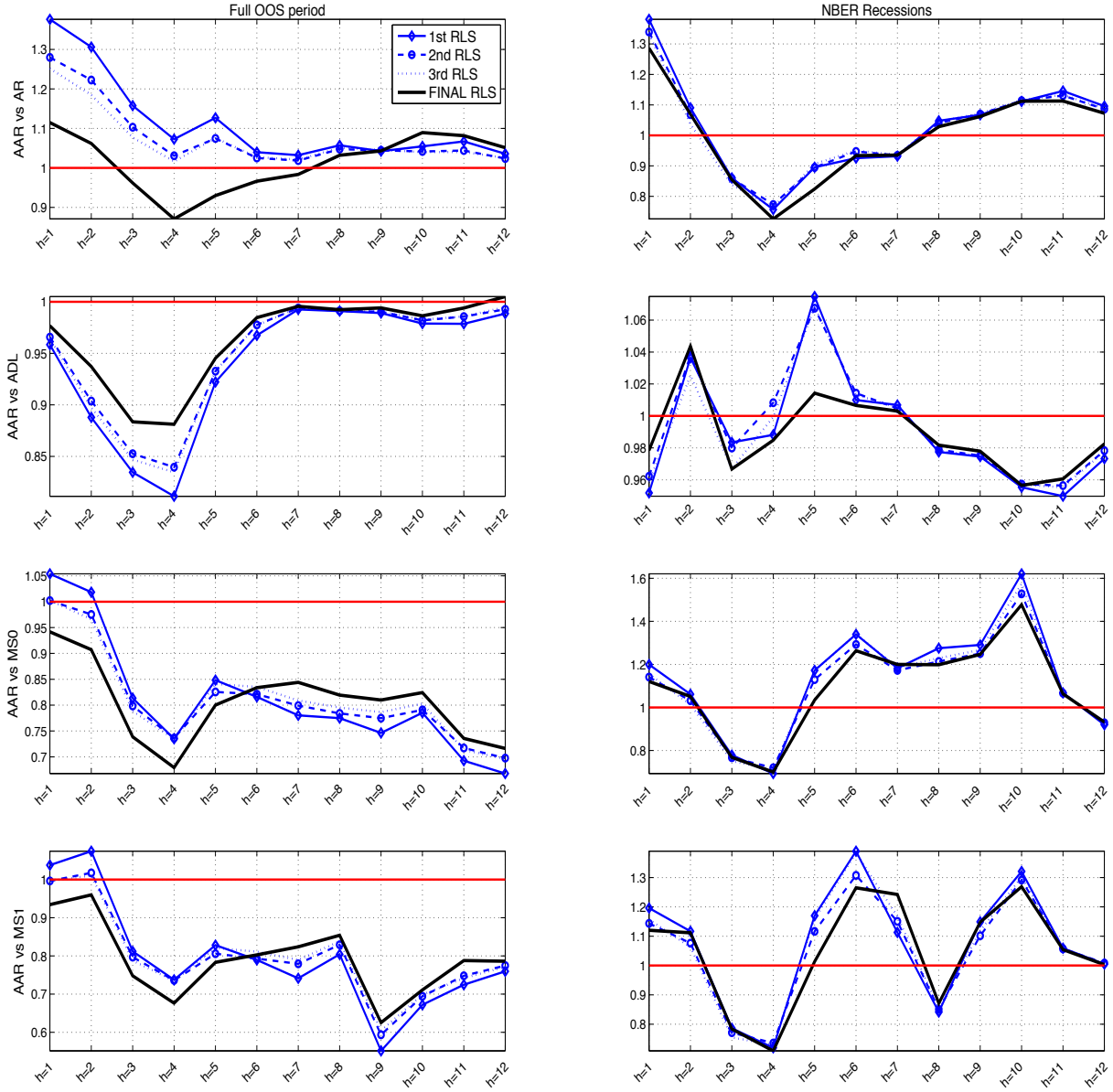
### 5.2.2 Forecasting GDP Growth

In this section we compare the forecasting performance of our AAR model versus the linear and nonlinear benchmarks. The analysis is done in real time and on an out-of-sample (OOS) basis. The OOS evaluation period spans 1981:Q4-2015:Q4. The models are estimated recursively from 1959:Q1 and are updated only at the NBER announcements. We consider forecasting economic activity variables 1 to 12 quarters ahead. The performance metrics advocated is the mean squared error (MSE). In all tables, we always show the ratio of MSEs between the AAR and the competing model. If this ratio is lower than 1, this means that the AAR produces more accurate forecasts than the competing model. In the main text, we show the results that are based on the Probit model that uses the TS only as predictor.

Figure 10 shows the performance of the models on the full out-of-sample period and during recession periods only. As the analysis is done in real time, we are able to construct forecast errors with respect to the first, second, third and the final release of the GDP. The final releases are the best approximations of the actual data of interest which, in the realm of real-time analysis, are considered latent. Therefore, it is important to assess the ex-post accuracy of real-time forecasts for subsequent releases.

Comparing the AAR to the AR model on the period 1981-2015 (Figure 10, upper left panel), we see that adding the probability of recession to the AR improves its forecasting accuracy for final releases by more than 10% at one-year horizon. The AAR model delivers a smaller relative MSE for horizons 3 to 7 quarters for final releases, but it is outperformed at all horizons when predicting the initial releases. As expected, the relative out-of-sample performance of the AAR improves significantly during NBER recessions (Figure 10, upper right panel). The largest performance improvements are observed at horizons between 3 and 7 quarters. At one-year horizon, the relative efficiency of the AAR is as large as 30%. The results are qualitatively similar across releases, but the relative efficiency of the AAR is maximized for the final release. This is an important result given that revisions tends to be important during recession (see the section on the Great Recession in the sequel). Having an approach in hand to predict accurately the final releases in real-time is crucial for policy makers, investors and statistical agencies that are in charge of nowcasting and revision of the national accounts forecasts.

Figure 10: Out-of-sample performance of AAR model



This figure shows the ratio of the AAR mean squared errors over the competing models for the full out-of-sample period, the left column, and during the NBER recessions, the right column. ADL correspond to model in (5). MS0 and MS1 are the Markov Switching models (11) and (12) respectively. 1st RLS, 2nd RLS, 3th RLS and Final RLS stand for data releases.

The second row of Figure 10 compares the AAR and ADL model. Recall that the difference between the AAR and the ADL rests on the treatment of the Term Spread (i.e., nonlinear for the AAR and linear for ADL). Considering the full out-of-sample period, the AAR outperforms the ADL uniformly over the forecast horizons with the smallest relative MSE occurring around horizons 3 and 4 quarters (12% for the final release and nearly 20% for the 1st release). This

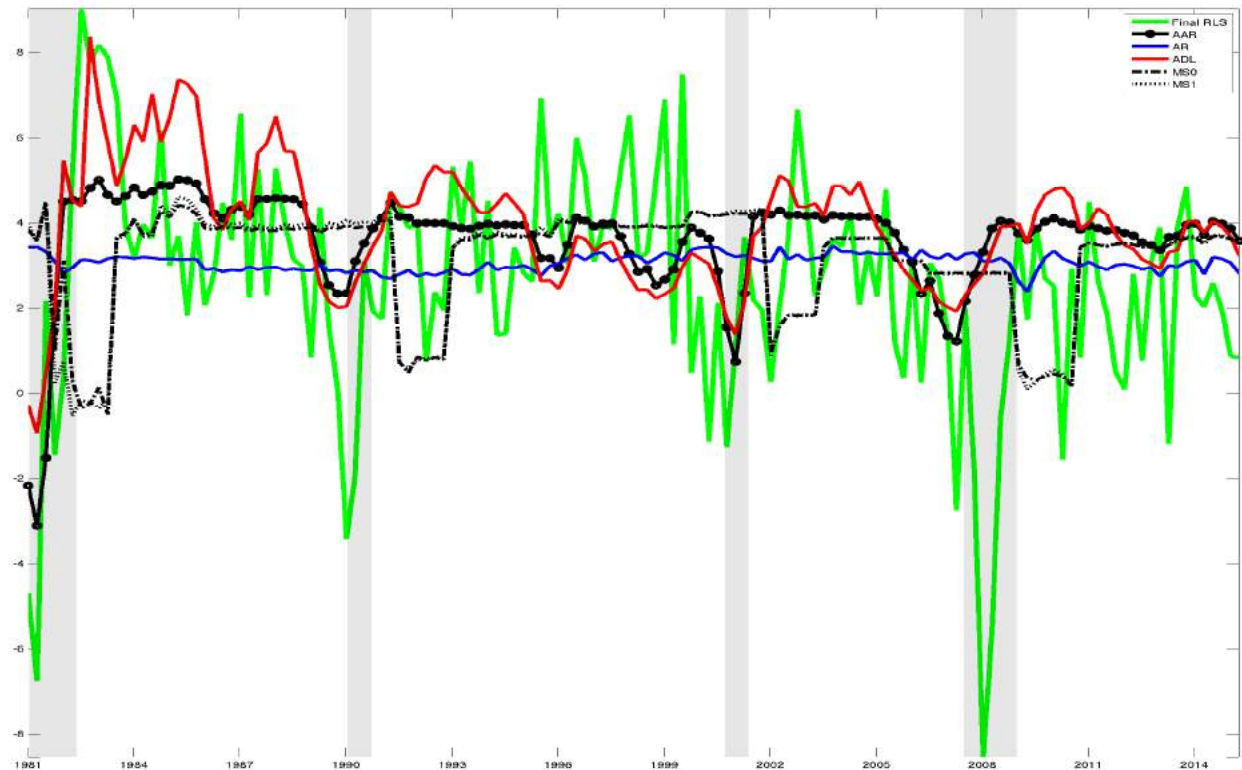
good performance of the AAR is attributable to the non-linear treatment of the TS done via the probability of recessions. In fact, the contribution of Term Spread to the predictions of the AAR is negligible during no-recession periods where the probability of a recession is close to zero while the ADL always imposes the same marginal effect for TS. Note however that the AAR does not uniformly dominate the ADL across horizons during NBER recession periods. The AAR still displays a smaller relative MSE at several horizons for the final release.

Finally, the last two rows of Figure 10 compares the AAR model to two popular nonlinear alternatives. MS0 is the two-state Markov switching model (11) where only the intercept is state dependent. The AAR model outperforms the Markov switching models on the full OOS period and at all forecast horizons for the final release. The improvement attains more than 30% at one-year horizon compared to MS0 and nearly 40% at  $h = 9$  with respect to MS1. During recessions, the Markov switching models perform better except at horizons 3 to 4 quarters where the predictive power of the Probit model is maximized (hence the good performance of AAR at these horizons). Indeed, the AAR model delivers predictions that are up to 20% more accurate than the forecasts made by MS models at horizons 3 and 4. The reason is rather simple. The AAR model relies on a first step probabilistic model that exploits the forward looking information contained in the TS. This forward looking information is brought into the forecasting equation much earlier than in the MS model, where a rather large change in GDP must occur before we observe a change in the value of the state variable. Moreover, GDP data are available with one quarter lag and subject to large revision during crises. The observability of the NBER recessions therefore gives some mileage to the AAR approach at horizons where the signal that leads recessions (i.e., the probability of recessions) is maximized.

Figure 11 plots the real-time out-of-sample forecasts of the GDP growth 4-quarters ahead as well as the final release. Table (1) presents the usual statistics: MSPE, bias and variance ratio of the AAR model over the alternative models and the  $p$ -values of Diebold-Mariano's accuracy test. The AR forecast is weakly oscillating around the unconditional average of the GDP growth, hence the bias is low but the variance of the forecast error is high. The ADL is closer to the AAR during recessions but is more biased otherwise. The AAR performs better than Markov Switching for a reason that was mentioned previously. Namely, the signal that leads recessions is taken into account early by the AAR via the first step probit that conditions the probability of recessions

on forward looking financial variables. The MS models relies more on a delayed signal. Table (1) shows that over the full out-of-sample, the AAR model outperforms significantly the MS models in terms of MSPE. This good performance of the AAR model compared to the AR and Markov Switching models is mainly attributable to the variance. With respect to ADL, the AAR produce more than 20% less biased forecasts. During NBER recessions, the over-performance of the AAR is attributable to bias reduction.<sup>11</sup> This is again relevant for policy makers since the bias is very important during economic downturns.

Figure 11: Out-of-sample 4-quarter ahead forecasts



*This figure shows the real-time 4 quarters ahead forecasts of the GDP growth. ADL correspond to model in (5). MS0 and MS1 are the Markov Switching models (11) and (12) respectively. Final RLS stand for the final data release.*

### 5.3 Real-Time Analysis of the Latest Recessions of the U.S. Economy

Policy makers and investors attach a high value to forecast accuracy during recession episodes. Typically, recession episodes are shorter than expansion episodes and characterized by rapid changes in the values of many economic indicators. Moreover, economic activity data that are released

<sup>11</sup>Diebold-Mariano  $p$ -values should be taken with a grain of salt since there are only 16 periods of recessions.

Table 1: Performance statistics for 4-quarter ahead forecasting

<i>AAR w/r</i>	Full out-of-sample				NBER Recessions			
	MSPE	DM	Bias	Variance	MSPE	DM	Bias	Variance
AR	0,8709	0,0459	2,3169	0,7588	0,7260	0,0811	0,7165	1,2968
ADL	0,8812	0,0370	0,7873	0,9520	0,9847	0,4256	0,9052	1,2537
MS0	0,6793	0,0468	1,9161	0,5940	0,7002	0,0946	0,6946	1,3463
MS1	0,6763	0,0450	1,9205	0,5911	0,7090	0,1067	0,7194	1,1686

*Note: This table shows the real-time out-of-sample forecasting statistics of the AAR model against the alternatives. Forecasting errors are calculated with final release data. MSPE stands for the ratio of AAR mean squared predictive error over the alternative. DM stands for p-value of Diebold-Mariano test (Clark-West adjustment has been used for comparison between AAR and AR because the models are nested). Bias (Variance) stand for ratio of bias (variance) of AAR over the alternative models.*

during recession episodes tend to undergo more or less important revisions. It is therefore of interest to assess how well our models perform during those periods where economic data are harder to track than usual.

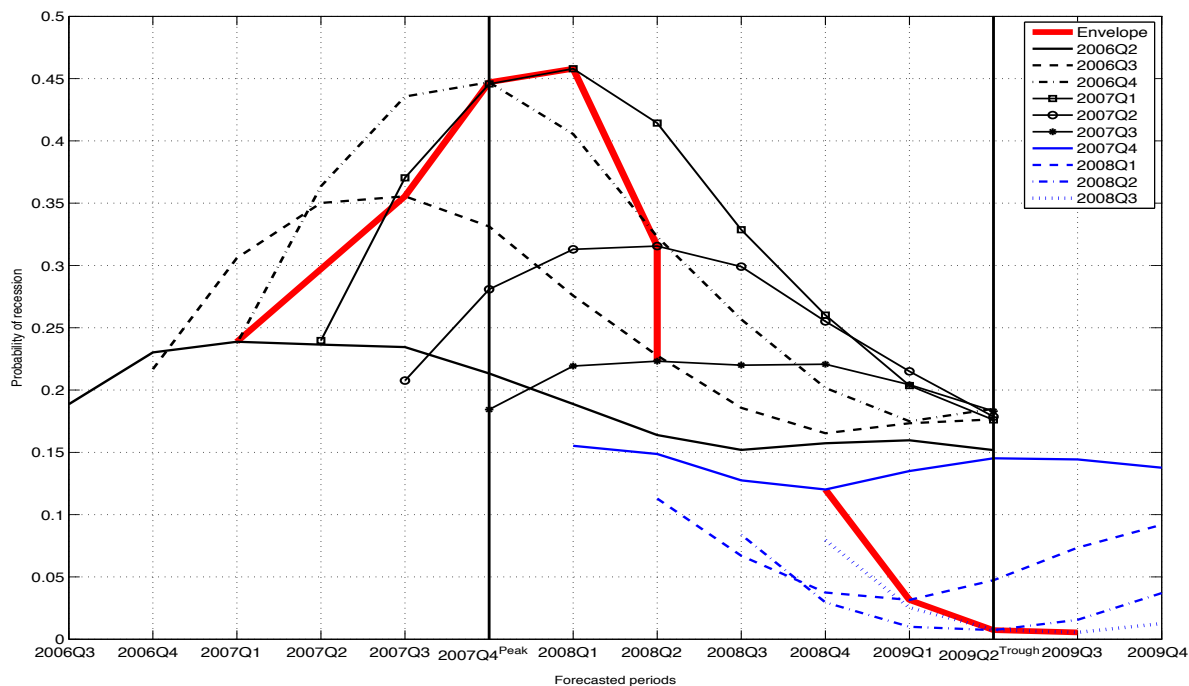
In this section, we conduct a real-time analysis of the last two recessions experienced by the U.S. economy. For each recession, we define a time window that starts around four periods before the recession and ends around two periods after the recession. Considering each recession in turn, we generate the term structure of probabilities of recession for each period on the selected window. Second, we compare the term structure of forecasts of the AAR versus AR and ADL models on two separate figures. Third, we compare the term structure of forecasts of the AAR and MS models on another graph. The MS model considered here is the one with changing intercept and constant autoregressive coefficient. Finally, we generate trajectories of recursive forecasts for the AAR model at fixed horizons. The analysis is done in real time, meaning that the forecasts are generated using the most recent GDP release and parameters that are estimated from a sample that stops at the previous official NBER turning point.

### 5.3.1 The Great Recession (2007-2009)

Figure 12 shows ten curves, each of them representing the term structure of the probability of recession at a given date. Two bold red lines are superimposed in order to ease the visualization of the upper envelope of the six concave curves that are at the beginning of the recession and the lower envelope of the four convex curves that are at the end of the recession. The maximum of the

upper envelope curve is located at 2008Q1, one quarter only after the date officially designated by the NBER as the peak of the business cycle. Likewise, the minimum of the lower envelope curve is located one quarter after the official end of the recession. This suggests that the upper and lower envelope curves are informative about the business cycle turning points as hypothesized previously.

Figure 12: Forecasting the Great Recession turning points in real time



*This figure shows out-of-sample predictions of the term structure of recession probabilities obtained from the static Probit model that uses TS only as predictor around the great recession. The black line 2006Q2 corresponds to the forecasts conditional on 2006Q2 information, the black dotted line forecasts is made conditional on 2006Q3, and so on. Thick red lines stand for the upper and lower envelopes of the predicted term structures of recession probabilities.*

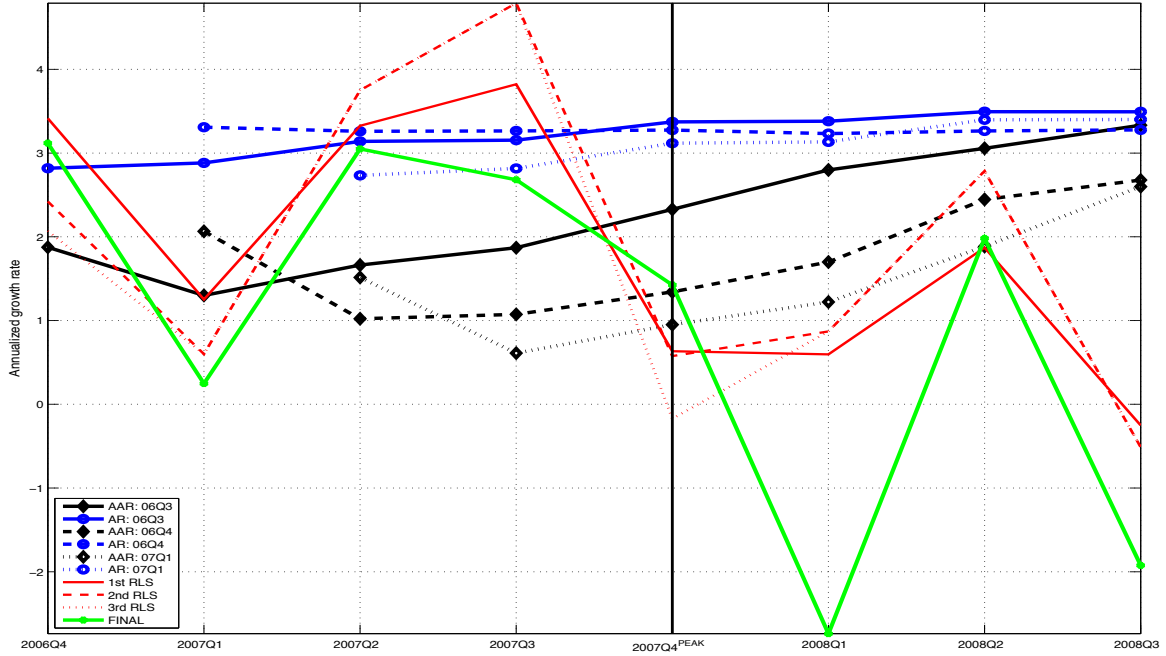
Figure 13 compares the term structures of out-of-sample predictions for the AAR and AR models. The term structures of out-of-sample forecasts produced by the AR model are roughly flat while those produced by the AAR model exhibit more correlation with the actual data. Clearly, the out-of-sample forecasts of the AR model are more disconnected from reality than those of the AAR model.

Figure 14 compares the term structure of forecasts of the AAR and ADL models. The forecasts of the ADL model are more responsive to the actual data than those of the AR model, but less responsive than those of the AAR model. The gap between the term structures of forecasts of the AAR and ADL models increases with the horizon.

Figure 15 compares the term structures of out-of-sample predictions for the AAR and MS models. All term structure of forecasts are flat for the MS model. Their levels a few quarters



Figure 13: Direct OOS forecasts of GDP growth during the Great Recession: AR vs. AAR

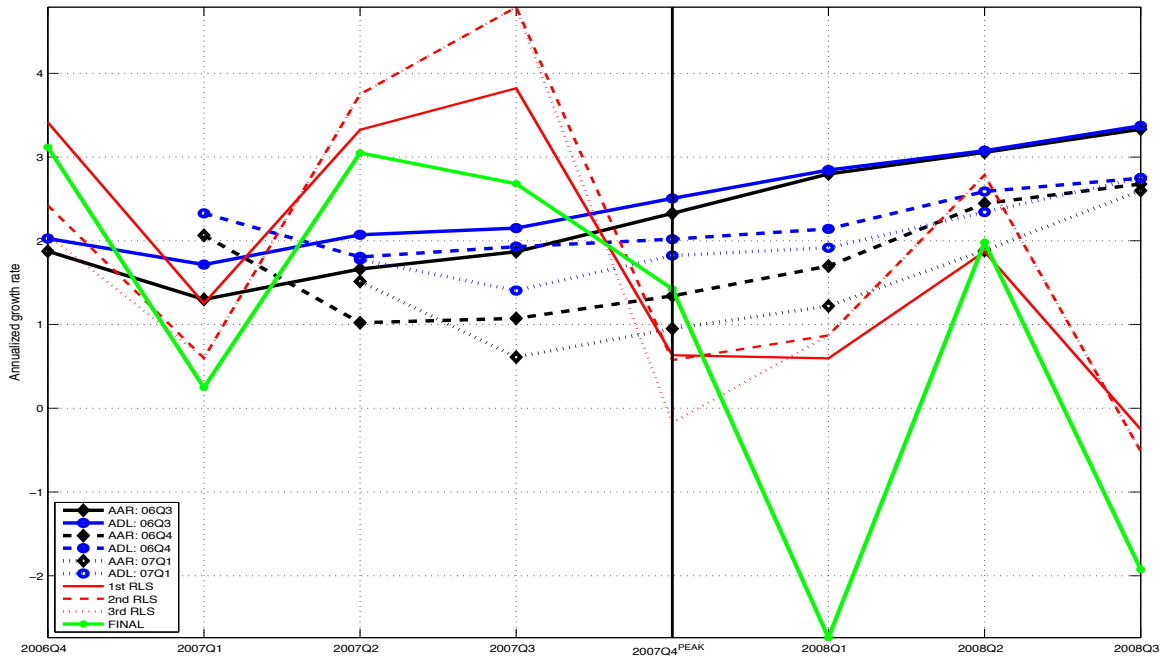


This figure shows out-of-sample forecasts of GDP growth at different horizons obtained from the AR and AAR models around the great recession. For instance, AAR: 06Q3 stands for the forecasts of the AAR model based on the first release of 2006Q3 GDP, as it was available in 2006Q4, and so on. The other lines represent the 1st, 2nd and 3rd releases in real time as obtained from the corresponding vintages. The final data are the most recent values of the 2015Q1 vintage.

before and through the middle of the recession are quite deceptive about reality. Indeed, one has to wait until 2009Q1 (one quarter before the official end of the recession) before seeing a sudden drop in the level of the term structure of forecasts of the MS model. Overall, the term structure of out-of-sample forecasts of the MS model has an uninformative shape while its level signals the recession quite late.

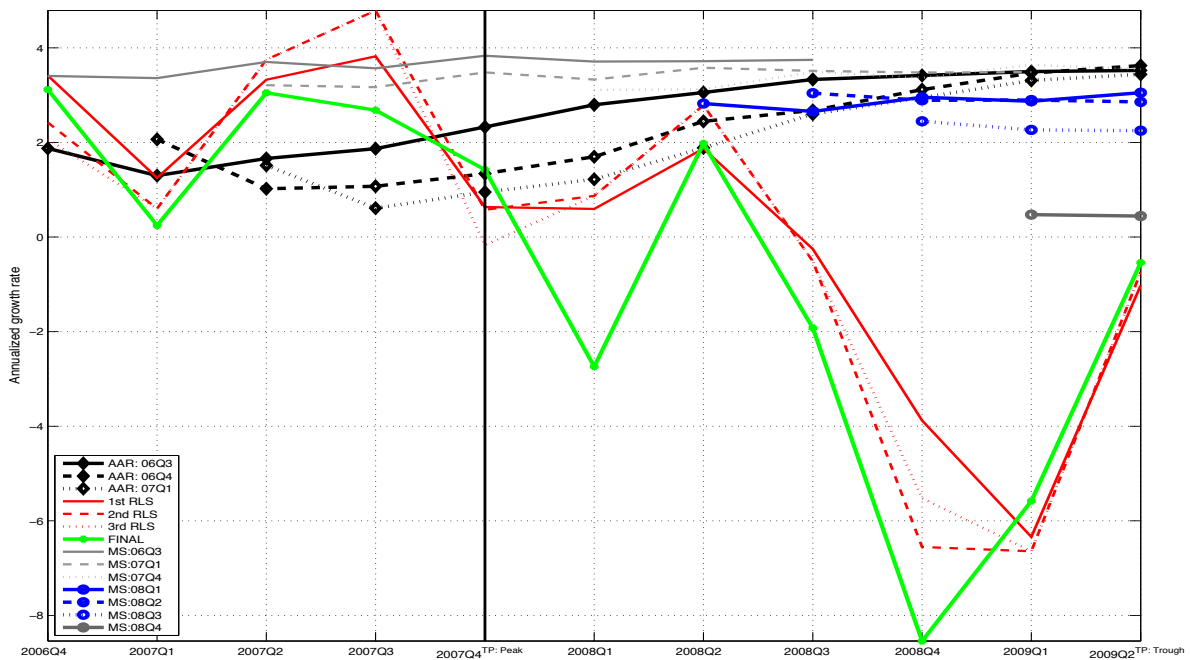
Figure 16 shows the trajectories of recursive forecasts of the AAR model for  $h = 3$  and  $h = 4$  along with the corresponding optimistic and pessimistic scenarios. The results are quite similar for both horizons. The average forecasts are upward trending during the recession periods while the optimistic and pessimistic forecasts are decreasing. The actual realizations of GDP growth are much lower than the average forecasts. Indeed, actual realizations are closer to the pessimistic forecasts between 2008Q4 and 2009Q1. Having a pessimistic scenario available in advance can help mitigate the impact of bad surprise during recessions.

Figure 14: Direct OOS forecasts of GDP growth during the Great Recession: AAR vs. ADL



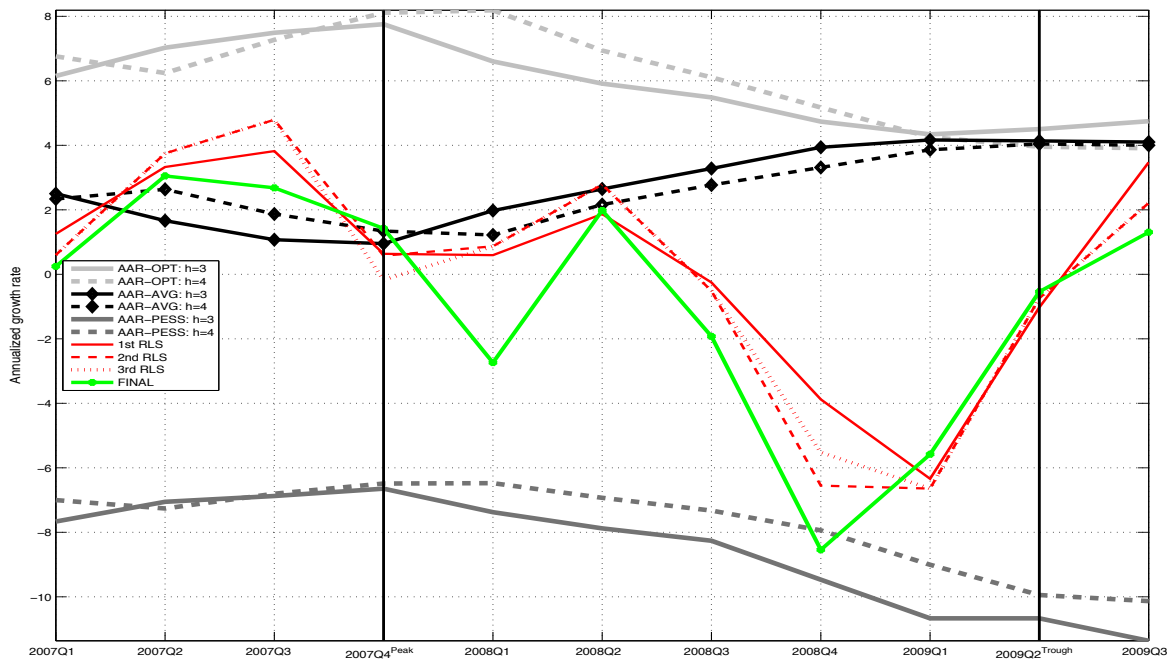
This figure shows out-of-sample forecasts of GDP growth at different horizons obtained from the AAR and ADL models around the great recession. For instance, AAR: 06Q3 stands for the forecasts of the AAR model based on the first release of 2006Q3 GDP, as it was available in 2006Q4, and so on. The other lines represent the 1st, 2nd and 3rd releases in real time as obtained from the corresponding vintages. The final data are the most recent values of the 2015Q1 vintage.

Figure 15: Direct OOS forecasts of GDP growth during the Great Recession: AAR vs. MS



This figure shows direct out-of-sample forecasts of GDP growth at different horizons obtained from the AAR model and the MS model with state-dependent intercept around the great recession.

Figure 16: Recursive OOS forecasts of GDP growth during the Great Recession: average, optimistic and pessimistic scenarios



This figure shows the trajectories of recursive forecasts of the AAR model for  $h = 3$  and  $h = 4$  along with the corresponding optimistic (AAR-OPT) and pessimistic (AAR-PESS) scenarios around the great recession.

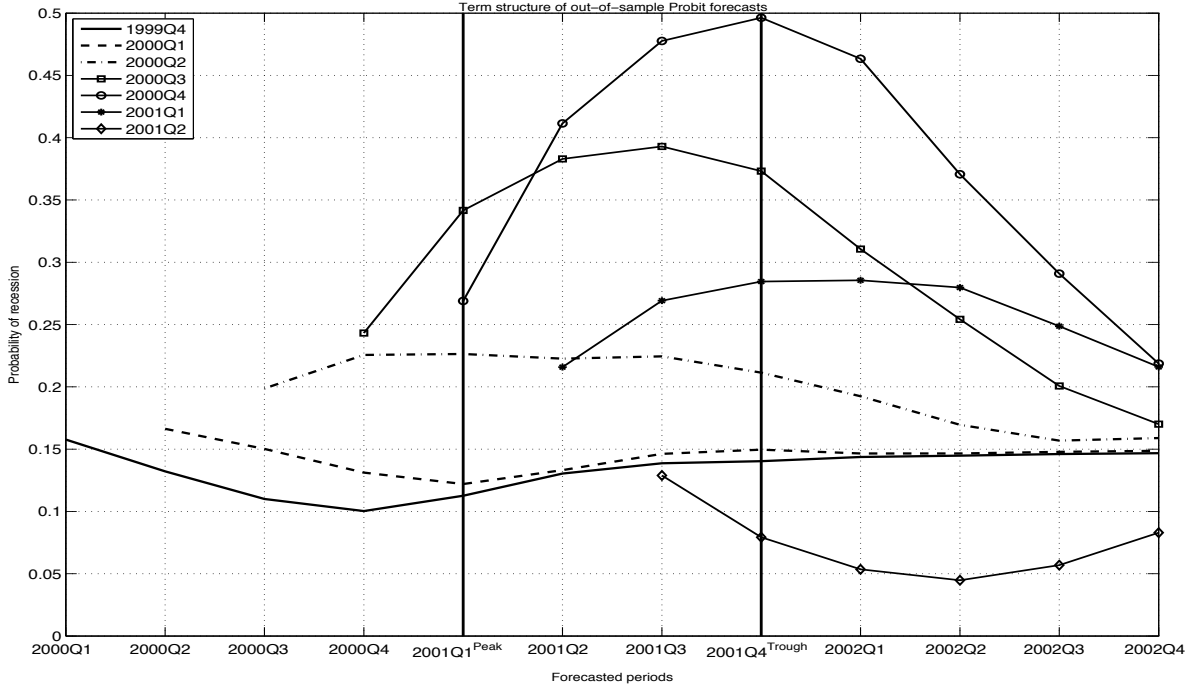
### 5.3.2 The 2001 Recession

The strategy which consists of using the maximum of the upper envelope of the term structure of recession probabilities and the minimum of its lower envelope as turning points does not work well for the 2001 recession (See Figure 17). This might be due to the fact that this recession has been much shorter than the Great recession, or that it has different causes from the other recessions so that it is not predictable by spreads. See discussions in Ng and Wright (2013), Kim and Murray (2002) and French (2005). The beginning and end of this recession are signaled by sudden changes in the shape of the term structure of the probability of recession. Indeed, this curve switched from convex to concave at the beginning of the recession and from concave to convex at the end of the recession. This suggests that changes in the shape of the term structure of recession probabilities signal turning points.<sup>12</sup>

As previously, the term structure of out-of-sample forecasts of GDP growth generated by the AR model remain quite flat over time (see Figure 18). The term structure of out-of-sample forecasts

<sup>12</sup>This further suggest that the great recession of 2007 begun at least a year before the official date announced by the NBER. See Figure 12.

Figure 17: Forecasting the 2001 recession turning points in real time



This figure shows out-of-sample predictions of the term structure of recession probabilities obtained from the static Probit model that uses TS only as predictor around the 2001 recession. The black line 1999Q4 corresponds to forecasts conditional on 1999Q4 predictors value, etc..

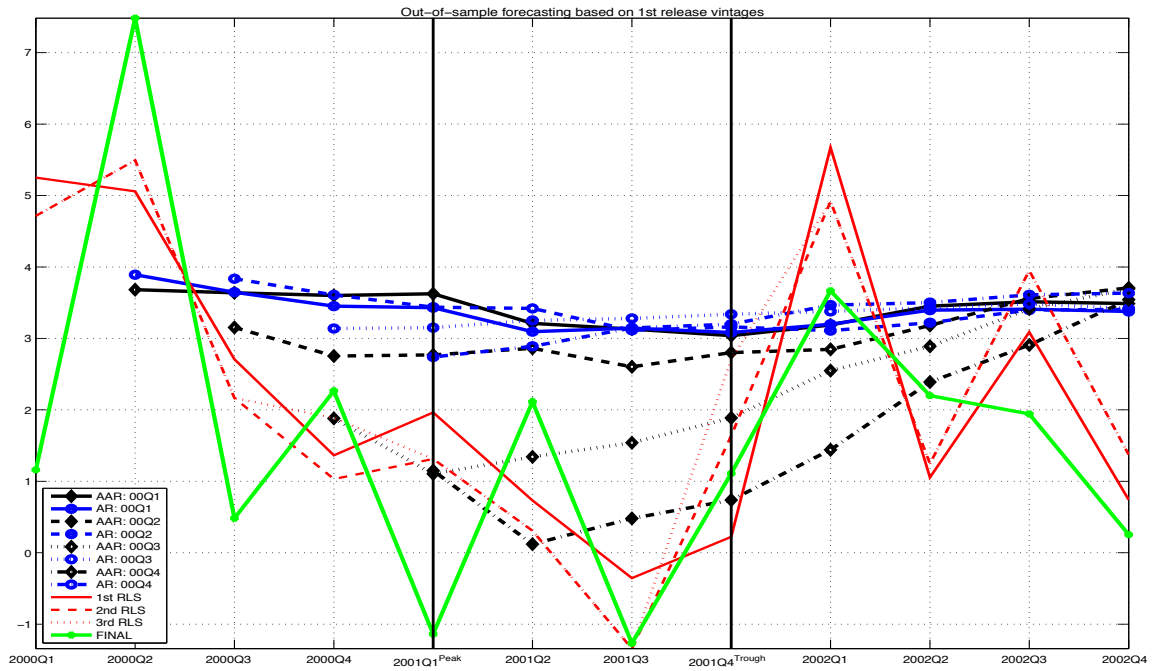
generated by the AAR model are more informative about the actual economic condition. At glance, we see that the fit of the AAR model is better here than for the Great recession.

Figure 19 shows the term structure of forecasts for the AAR and ADL models. The conclusions are consistent with what was seen previously. That is, the AAR model performs better than the ADL model at glance. As we update the term structure of forecasts, both models fit the data better than during the great recession.

Figure 20 compares the term structure of out-of-sample forecasts for the AAR and MS models. Again, the results are similar to those obtained for the great recession. The shape of the term structure of out-of-sample forecasts of the MS model is uninformative and flat while its level dropped significantly only after the official end of the recession.

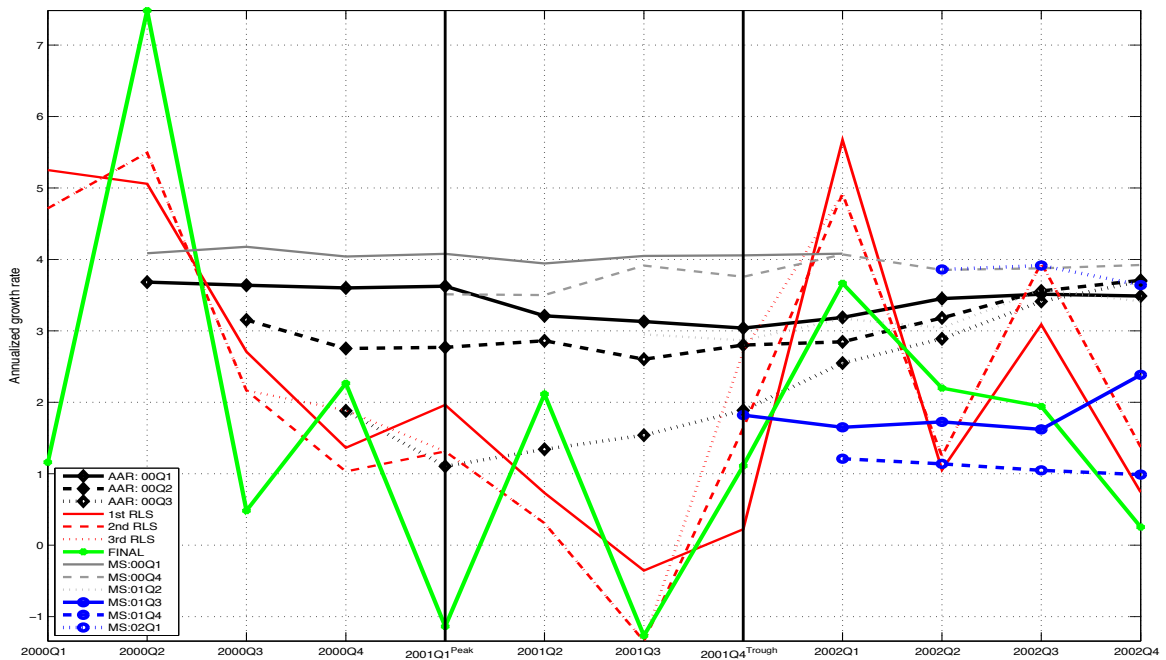
Figure 21 shows the trajectory of recursive forecasts at horizons  $h = 3$  and  $h = 4$ . The actual data are closely tracked by the average forecast. The latter forecast is closer to the optimistic forecast than to the pessimistic forecast, a situation that is probably attributable to the brevity of the 2001 recession.

Figure 18: Direct OOS forecasts of GDP growth during the 2001 Recession: AAR vs. AR



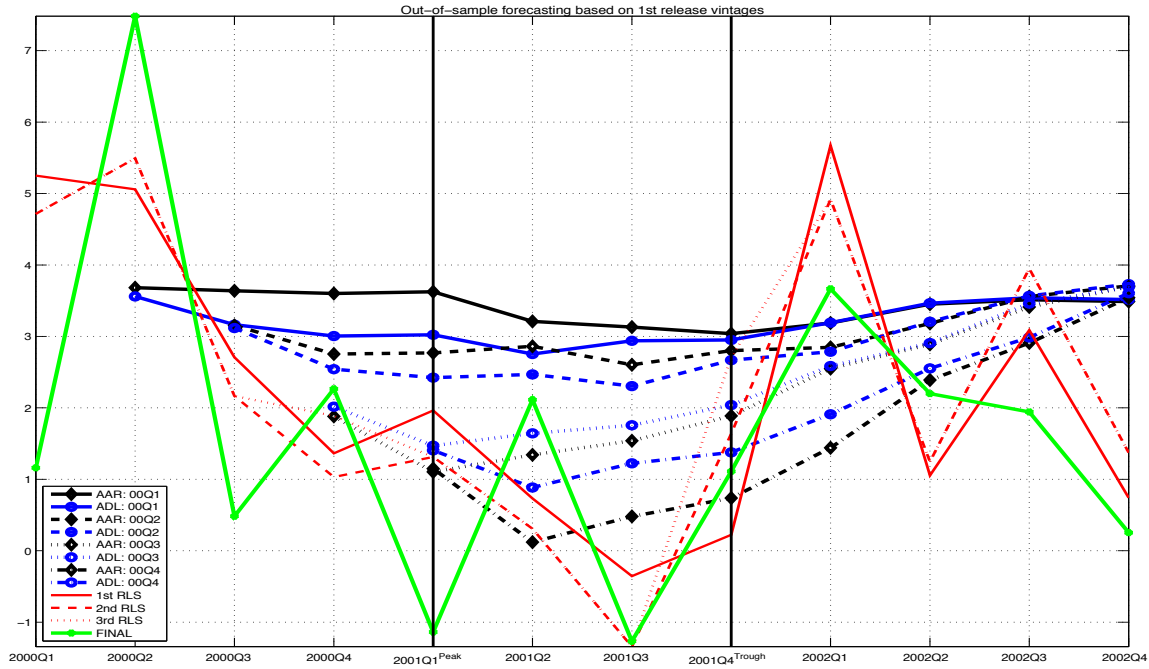
This figure shows out-of-sample forecasts of GDP growth at different horizons obtained from the AAR and AR models around the 2001 recession. For instance, AAR: 00Q1 stands for the AAR model forecasts using the first release of 2000Q1 GDP, as it was available in 2000Q2, etc.

Figure 20: Direct OOS forecasts of GDP growth during the 2001 recession: AAR vs. MS



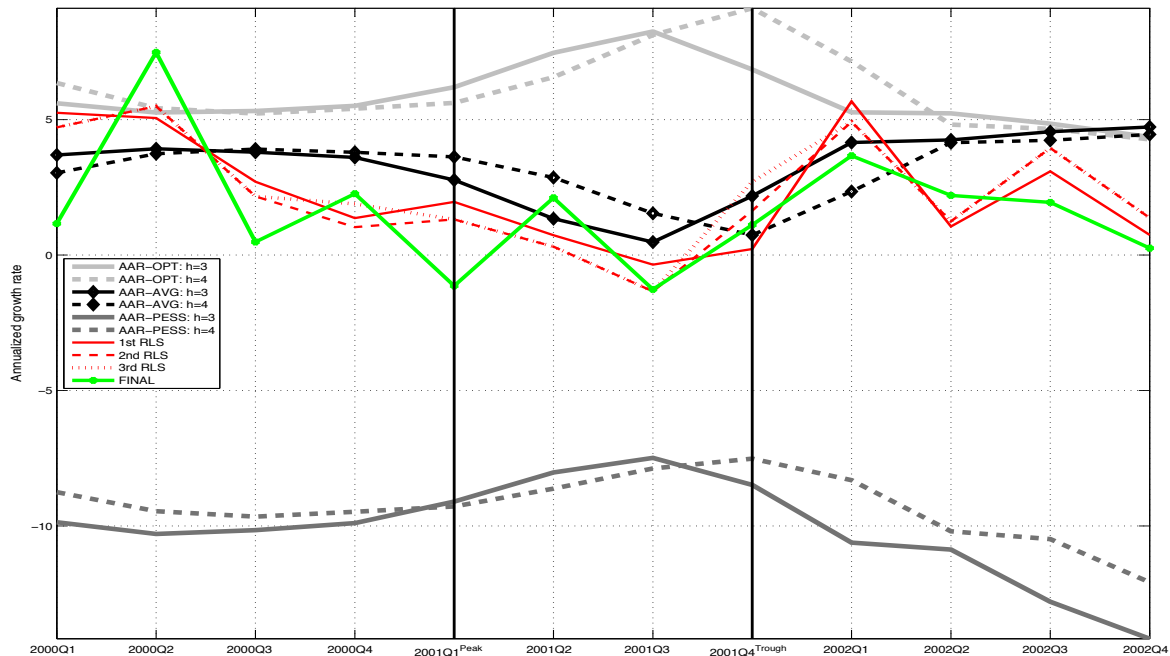
This figure shows direct out-of-sample forecasts of GDP growth at different horizons obtained from the AAR model and the MS model with state-dependent intercept around the 2001 recession.

Figure 19: Direct OOS forecasts of GDP growth during the 2001 recession: AAR vs. ADL



This figure shows out-of-sample forecasts of GDP growth at different horizons obtained from the AAR and ADL models around the 2001 recession. For instance, AAR: 00Q1 stands for the AAR model forecasts using the first release of 2000Q1 GDP, as it was available in 2000Q2, and so on. The other lines represent the 1st, 2nd and 3rd releases in real time as obtained from the corresponding vintages. The final data are the most recent values of the 2015Q1 vintage.

Figure 21: Recursive OOS forecasts of GDP growth during the 2001 recession: average, optimistic and pessimistic scenarios



This figure shows the trajectories of recursive forecasts of the AAR model for  $h = 3$  and  $h = 4$  along with the corresponding optimistic (AAR-OPT) and pessimistic (AAR-PESS) scenarios around the 2001 recession.

## 6 Conclusion

This paper explores an approach based on augmented autoregressive models (AAR) to forecast future economic activity. Average forecasts are obtained from an AR(1) model that is augmented with a variable that measures the probability of a recession conditional on forward-looking financial variables. AR(1) models augmented with IMRs are used to produce forecasts of economic activity that are conditional on whether the horizon of interest is a recession period (pessimistic forecast) or not (optimistic forecast). The implementation of these models require a prior estimation of a Probit model for the probability of a recession. Overall, our methodology is simple, parsimonious and easy to replicate. It can be easily adapted to other contexts by replacing the economic activity variable by another variable of interest (unemployment rate, credit volume, etc.) and adding more predictors to the first step Probit model. In particular, our pessimistic forecast can be used as input for stress testing exercises in the banking and real estate sectors.

Our empirical results suggest that a static Probit model that uses only the TS as regressor provides comparable in-sample and out-of-sample fit to the data as more sophisticated Probit models. We also find that the dynamic patterns of the term structure of recession probabilities are informative about the business cycle turning points. Indeed, these term structure curves switch from convex to concave near to the beginning of recessions and from concave to convex near to the end of recessions. Our most parsimonious AAR model delivers better out-of-sample forecasts of GDP growth than the AR, ADL and Markov Switching models.

## References

- Aastveit, K. A., Carriero, A., Clark, T. E., and Marcellino, M. (2014). Have standard vars remained stable since the crisis? Technical report, Manuscript.
- Anderson, H. M. and Vahid, F. (2001). Predicting the probability of a recession with nonlinear autoregressive leading-indicator models. *Macroeconomic Dynamics*, 5:482–505.
- Berge, T. and Jorda, O. (2011). Evaluating the classification of economic activity into recessions and expansions. *American Economic Journal: Macroeconomics*, 3:246–277.
- Chauvet, M. (1998). An econometric characterization of business cycles with factor structure and markov switching. *International Economic Review*, 39:969–996.
- Chauvet, M. and Hamilton, J. (2006). Dating business cycle turning points. In C” Milas, P. R.

- and van Dikj, D., editors, *Nonlinear Time Series Analysis of Business Cycles*, volume 77, pages 1–54. North Holland, Amsterdam.
- Chauvet, M. and Piger, J. (2008). A comparison of the real time performance of business cycle dating methods. *Journal of Business and Economic Statistics*, 26:42–49.
- Chauvet, M. and Potter, S. (2002). Predicting a recession: Evidence from the yield curve in presence of structural breaks. *Economic Letters*, 77:245–253.
- Chauvet, M. and Potter, S. (2005). Forecasting recessions using the yield curve. *Journal of Forecasting*, 24:77–103.
- Chauvet, M. and Potter, S. (2013). Forecasting output. In Elliott, G. and Timmermann, A., editors, *Handbook of Forecasting*, volume 2, pages 1–56. North Holland, Amsterdam.
- Christiansen, C., Eriksen, J. N., and Møller, S. V. (2013). Forecasting us recessions: The role of sentiments. Technical report, CREATES Research Paper 2013-14.
- Clark, T. E. K. and McCracken, W. (2013). Evaluating the accuracy of forecasts from vector autoregressions. In Fomby, T., Kilian, L., and Murphy, A., editors, *Advances in Econometrics: VAR models in Macroeconomics New Developments and Applications*.
- Doan, T., Litterman, R., and Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions (with discussion). *Econometric Reviews*, 3:1–144.
- Dueker, M. (2005). Dynamic forecasts of qualitative variables: A qual var model of u.s. recessions. *Journal of Business & Economic Statistics*, 23:96–104.
- Dueker, M. and Wesche, K. (2005). Forecasting macro variables with a qual var business cycle turning point index. *Applied Economics*, 42(23):2909–2920.
- Estrella, A. and Mishkin, F. (1998). Predicting us recessions: Financial variables as leading indicators. *Review of Economics and Statistics*, 80:45–61.
- French, M. W. (2005). A nonlinear look at trend mfp growth and the business cycle: Results from a hybrid kalman/markov switching model. Technical Report 2005-12, Finance and Economics Discussion Series (FEDS), Divisions of Research and Statistics and Monetary Affairs, Federal Reserve Board. Washington, D.C., Staff Working Paper, Cambridge, Massachusetts.
- Gargano, A. and Timmermann, A. (2014). Forecasting commodity price indexes using macroeconomic and financial predictors. *International Journal of Forecasting*, 30:825–843.
- Giannone, D., Lenza, M., Momferatou, D., and Onorante, L. (2010). Short-term inflation projections: A bayesian vector autoregressive approach. Technical report, European Central Bank.
- Gilchrist, S., Yankov, V., and Zakrajsek, E. (2009). Credit market shocks and economic fluctuations: Evidence from corporate bond and stock markets. *Journal of Monetary Economics*, 56:471–493.
- Grover, S. and McCracken, M. (2014). Factor-based prediction of industry-wide bank stress. *St. Louis Federal Reserve Bank Review*, 96:173–194.
- 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. (2011). Calling recessions in real time. *International Journal of Forecasting*, 27:1006–1026.
- Hao, L. and Ng, E. C. (2011). Predicting canadian recessions using dynamic probit modelling approaches. *Canadian Journal of Economics*, 44(4):1297–1330.
- Issler, J. and Vahid, F. (2006). The missing link: using the nber recession indicator to construct coincident and leading indices of economic activity. *Journal of Econometrics*, 132:281–303.
- Kauppi, H. and Saikkonen, P. (2008). Predicting u.s. recessions with dynamic binary response models. *Review of Economics and Statistics*, 90:777–7791.
- Kim, C.-J. and Murray, C. J. (2002). Permanent and transitory components of recessions. *Empirical Economics*, 27(2):163–183.
- Leroux, M., Kotchoni, R., and Stevanovic, D. (2017). Macroeconomic forecast accuracy in a data-



- rich environment. Technical report, CIRANO, 2017s-05.
- Meyer, B. and Zaman, S. (2013). Its not just for ination: The usefulness of the median cpi in bvar forecasting. Working Paper 13-03, Federal Reserve Bank of Cleveland.
- Ng, S. (2014). Boosting recessions. *Canadian Journal of Economics*, 47(1):1–34.
- Ng, S. and Wright, J. (2013). Facts and challenges from the great recession for forecasting and macroeconomic modeling. *Journal of Economic Literature*, 51(4):1120–1154.
- Nyberg, H. (2010). Dynamic probit models and financial variables in recession forecasting. *Journal of Forecasting, Special Issue: Advances in Business Cycle Analysis and Forecasting*, 29:215–230.
- Rapach, D. E. Strauss, J. K. and Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies*, 23(2):821–862.
- Rudebusch, G. D. and Williams, J. (2009). Forecasting recessions: The puzzle of the enduring power of the yield curve. *Journal of Business and Economic Statistics*, 27(4):492–503.
- Schorfheide, F. and Song, D. (2013). Real time forecasting with a mixed-frequency var. Technical report, NBER.
- Sims, C. (1982). Policy analysis with econometric models. *Brookings Papers on Economic Activity*, pages 107–152.
- Stock, J. H. and Watson, M. W. (1989). New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*, pages 351–393.
- Stock, J. H. and Watson, M. W. (2010). Indicators for dating business cycles: Cross-history selection and comparisons. *American Economics Review: Papers and Proceedings*, 100.
- Stock, J. H. and Watson, M. W. (2012). Estimating turning points using large data sets. *Journal of Econometrics*, page forthcoming.
- Wright, J. (2006). The yield curve and predicting recessions. Technical report, Federal Reserve Board.