Forecasting U.S. Recessions with Macro Factors

Sebastian Fossati*
University of Alberta

This version: May 19, 2015

Abstract

Dynamic factors estimated from panels of macroeconomic indicators are used to predict future recessions using probit models. Three factors are considered: a bond and exchange rates factor; a stock market factor; a real activity factor. Three results emerge. First, models that use only financial indicators exhibit a large deterioration in fit after 2005. Second, models that use factors yield better fit than models that use indicators directly. Out-of-sample forecasting exercises confirm these results for 3-, 6-, and 12-month horizons using both ex-post revised data and real-time data. Third, results show evidence that data revisions affect factors less than individual indicators.

Keywords: Recession, Forecasting, Factors, Probit Model.
JEL Codes: E32, C22, C25.

---

*Contact: Department of Economics, University of Alberta, Edmonton, AB T6G 2H4, Canada. Email: sfossati@ualberta.ca. Web: http://www.ualberta.ca/~sfossati/. I would like to thank Marcelle Chauvet and Jeremy Piger for providing part of the data used in this paper. I also thank Yu-chin Chen, Tim Cogley, Peter Fuleky, Chang-Jin Kim, and James Morley for helpful comments.

1 Introduction

Forecasting recessions (i.e., periods of decline in economic activity) is considered to be of special interest in macroeconomics as well as for policy makers and private economic agents. However, the prediction of business cycle phases in real-time (or shortly after) is particularly difficult since business conditions are never directly observable and the Business Cycle Dating Committee of the NBER makes its announcements long after the fact (often more than a year). For example, the NBER determined that a peak in economic activity (beginning of a recession) occurred in the U.S. economy in December 2007. This announcement, however, was not made until December 2008. In fact, over the past 30 years, the NBER has made its announcements between 6 to 20 months after the corresponding peak or trough.

In this context, a common strategy among those interested in modeling business conditions in real-time consists in generating recession probabilities using binary class models (e.g., probit, logit) for current or future NBER recession dates. The existing literature has focused mainly on probit models that use macroeconomic indicators directly. For example, Estrella and Mishkin (1998) find that the 3-month less 10-year term spread and stock price indexes are the most useful predictors of future U.S. recessions. Similarly, Wright (2006) finds that using the level of the federal funds rate together with the term spread improves the performance of the predictive probit models. Recently, Katayama (2010) analyzed the forecasting performance of several binary class models for NBER recessions using combinations of 33 macroeconomic indicators and a 6-month horizon. He concludes that the combination of the term spread, month-to-month changes in the S&P 500 index, and the growth rate of non-farm employment generates the sequence of out-of-sample recession probabilities that better fits NBER recession dates. Other relevant contributions to this literature include Dueker (1997), Chauvet and Potter (2005), Kauppi and Saikkonen (2008), Hamilton (2011), Owyang et al. (2012), Berge (2014), and Ng (2014), among others.

In this paper, I use dynamic factors estimated from small panels of macroeconomic indicators (macro factors) to forecast NBER recession dates using probit models. The goal is to compare forecasts from the factor-augmented probit models with forecasts from models that use only macroeconomic indicators. Three monthly macro factors are considered: (1) a bond and exchange rates factor extracted from 22 financial indicators; (2) a stock market factor extracted from 4 stock market indicators; (3) a real activity factor extracted from 4 "coincident" macroeconomic indicators. Recently, Chen et al. (2011) used static factors estimated by principal components from a large number of time series also to forecast future NBER recessions.[1] This approach based on large panels of macroeconomic indicators has been found useful in many forecasting exercises (see, e.g., Stock and Watson, 2002a,b, 2006). However, using dynamic factors estimated from small panels has some advantages. First, the small panel approach allows us to account for two important issues when evaluating the (pseudo) real-time out-of-sample performance of the forecasting models: (1) data availability at the time the forecast would have been made; (2) the effect of data revisions on the predicted probabilities. The first issue is addressed by properly taking into account the fact that real activity indicators are available with some lag. The second issue is addressed by comparing the out-of-sample forecasting performance of the models using both ex-post revised data and real-time data. In addition, factors estimated from small panels can be easier to interpret than factors estimated from large panels. Finally, recent work by Kauppi and Saikkonen (2008), Nyberg (2010), and Ng (2012), among others, has found

[1] Bellégo and Ferrara (2012) use a similar approach to forecast recessions in the euro area.

evidence that including dynamic elements (e.g., lags of the binary response variable) in the probit models can yield more accurate forecasts of U.S. recessions than standard probit models. But due to the long delay in NBER announcements, these dynamic models require important assumptions about what is known at the time of forecasting (specifically, what is the state of the economy in recent months). Since the real activity factor is a good predictor of the state of the economy (Chauvet and Piger, 2008), the models considered in this paper offer an alternative to the dynamic probit models that does not require knowledge of recent NBER turning points.[2]

The main results of this paper can be summarized as follows. First, probit models that use only financial indicators as predictors of future NBER recessions exhibit a large deterioration in fit after 2005. On the other hand, probit models that use both financial and real activity indicators directly or through a macro factor maintain their fit throughout the sample and exhibit a better forecasting performance during the 2008-09 recession. Second, probit models that use macro factors as predictors yield better in-sample fit than models that use indicators directly. Relative to the models proposed in Estrella and Mishkin (1998), Wright (2006), and Katayama (2010), the improvement can be substantial. Third, (pseudo) out-of-sample forecasting exercises designed to mimic real-time conditions confirm that forecasts from the probit models based on macro factors dominate forecasts from models previously considered in the literature. These results hold for 3-, 6-, and 12-month forecasting horizons using both ex-post revised data and real-time data. Finally, the results in this paper provide some

[2]Additional evidence supporting the use of small panels is provided in Camacho et al. (2013) and Fossati (2014). For example, Fossati (2014) compares the performance of a "small data" dynamic factor and a "big data" principal components factor as predictors of current NBER recessions using both binary class models and Markov-switching models. The results show that models based on the "small data" dynamic factor generate the sequence of out-of-sample class predictions that better approximates NBER recession dates. In addition, Camacho et al. (2013) show decreasing returns to adding more indicators with similar signal-to-noise ratios.

evidence on the issue of data revisions and factor models. In particular, data revisions appear to affect the real activity factor less than the individual real activity indicator (employment), a result conjectured in Berge and Jorda (2011) and Chen et al. (2011). As a result, probit models based on macro factors provide the best and most robust predictive performance for NBER recessions at all horizons considered in this paper.

This paper is organized as follows. Section 2 discusses the estimation of a dynamic macro factor from each of the three panels of macroeconomic indicators using Bayesian methods. Section 3 presents the predictive probit regressions and forecast evaluation statistics. Section 4 presents the empirical results. The in-sample results are presented in section 4.1. Out-of-sample results using both ex-post revised data and real-time data are presented in section 4.2. Section 5 concludes.

2 Estimation of Macro Factors

In this paper, instead of estimating latent common factors from a large panel of monthly macroeconomic indicators using principal components as in Stock and Watson (2002a,b, 2006), among others, I consider three small panels of indicators. These are: (1) a bond and exchange rates data set of 22 financial indicators including interest rates, interest rate spreads, and exchange rates; (2) a data set of 4 stock market indicators including stock price indexes, dividend yield, and price-earnings ratio; (3) a data set of 4 real activity indicators including industrial production, personal income less transfer payments, real manufacturing trade and sales, and employment. Dynamic factors estimated from each of these panels have been found useful in many forecasting exercises. For example, Ludvigson and Ng (2009) show that an important amount of variation in the two-year excess bond returns can be predicted by factors estimated from panels

(1) and (2).[3] Likewise, panel (3) has been used in Stock and Watson (1991), Diebold and Rudebusch (1996), Kim and Nelson (1998), Chauvet (1998), Chauvet and Piger (2008), Camacho et al. (2013), and Fossati (2014), among others, to model real-time business conditions.

For each of these three panels, I estimate a dynamic factor model using Bayesian methods and the following framework.[4] Let x be a $T \times N$ panel of macroeconomic indicators where x_{it}, $i = 1, \ldots, N$, $t = 1, \ldots, T$, has a factor structure of the form

$$x_{it} = \lambda_i(L) g_t + e_{it}, \tag{1}$$

where g_t is an unobserved dynamic factor, $\lambda_i(L) = \lambda_{i0} + \lambda_{i1} L + \ldots + \lambda_{is} L^s$, λ_{ij} are the dynamic factor loadings, and e_{it} is the idiosyncratic error. The dynamics of the latent factor are driven by an autoregressive process such that

$$\phi(L) g_t = \eta_t, \tag{2}$$

where $\phi(L)$ is a polynomial in L of order p_g and $\eta_t \sim i.i.d.\, N(0, \sigma_g^2)$. In addition, the dynamics of the idiosyncratic errors are also driven by autoregressive processes such that

$$\psi_i(L) e_{it} = \nu_{it}, \tag{3}$$

where $\psi_i(L)$ is a polynomial in L of order p_e and $\nu_{it} \sim i.i.d.\, N(0, \sigma_i^2)$ for $i = 1, \ldots, N$. For example, with $N = 4$, this is the dynamic factor model considered in Stock and Watson (1991).

For each of the three panels, the dynamic factor model is estimated recursively, starting with the sample period 1967:1-1988:1 and ending with the sample period

[3] See Ludvigson and Ng (2009) for a more detailed motivation for organizing the data into blocks.
[4] While the dynamic factors can also be estimated by maximum likelihood, Gibbs sampling provides a more robust alternative for the out-of-sample recursive exercises implemented below.

1967:1-2010:12 (i.e., the full sample). Prior to estimation, the data is transformed to ensure stationarity and standardized.[5] The factor model specification is completed by assuming $s = 2$ and $p_g = p_e = 1$ for every panel so that $\lambda_i(L) = \lambda_{i0} + \lambda_{i1}L + \lambda_{i2}L^2$, $\phi(L) = 1 - \phi L$, and $\psi_i(L) = 1 - \psi_i L$ for $i = 1, \ldots, N$. For estimation, the dynamic factor model is written in state-space form and estimated via Gibbs sampling following Kim and Nelson (1999) and Ludvigson and Ng (2009). Identification is achieved by setting the factor loading on the first time series in each panel to 1, i.e. $\lambda_{10} = 1$. Finally, the parameters λ_{ij} and ψ_i are initialized to zero, ϕ, σ_g^2, and σ_i^2 are initialized to 0.5, and principal components is used to initialize the dynamic factor. The Gibbs sampler runs 6,000 times. After discarding the first 1,000 draws (burn-in period), posterior means are computed using a thinning factor of 10, i.e. computed from every 10th draw. As a result, the subsequent analysis is based on the means of these 500 draws. The full sample estimated macro factors ($\hat{g}_{it}$ for $i = 1, 2, 3$) are presented in Figure 1.

[FIGURE 1 ABOUT HERE]

3 Predictive Regressions and Forecast Evaluation

The definition of the recession indicator follows Wright (2006) and is similar to the hitting probabilities considered in Chauvet and Potter (2005). Let $y_{t\,t+h}$ be a binary variable which equals 1 if the NBER's Business Cycle Dating Committee subsequently declared *any* of the months $t + 1$ through $t + h$ as a recession and 0 otherwise. A forecast of the probability of a recession in the next h months ($p_{t\,t+h}$) from a probit

[5] A complete description of the series and transformations is given in the appendix.

regression is then given by

$$p_{t\,t+h} = \text{Prob}(y_{t\,t+h} = 1 \,|\, z_t) = \Phi(\beta' z_t), \tag{4}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, β is a vector of coefficients, and z_t is a $k \times 1$ vector of predictors including an intercept.

Among the many potential predictors considered in the literature, the slope of the yield curve (or term spread) has been found to be a robust predictor of U.S. recessions. Estrella and Mishkin (1998), for example, conclude that the 3-month less 10-year term spread is the single best predictor of future recessions when looking at a horizon of two to four quarters. In addition, they find that stock price indexes can improve predictions and conclude that a model that uses these two financial indicators together gives a better out-of-sample predictive performance than one that uses the term spread alone. Similarly, Wright (2006) finds that a model using both the term spread and the level of the federal funds rate yields a better performance than a model using the term spread alone. Recently, Katayama (2010) analyzed the forecasting performance of 33 macroeconomic indicators using a 6-month horizon and concludes that the combination of the term spread, month-to-month changes in the S&P 500 index, and the growth rate of non-farm employment generates the sequence of out-of-sample recession probabilities that better fits subsequently declared NBER recession dates. Based on these studies, four indicators are selected as candidate regressors: (1) the 3-month less 10-year term spread (310TS); (2) the level of the federal funds rate (FFR); (3) the growth rate of the S&P 500 stock market index (SP500); (4) the growth rate of non-farm employment (EMP). In addition to these four indicators, I consider the three macro factors discussed in the previous section. As a result, in this paper I consider a regressor set of seven indicators ($310TS_t$, FFR_t, $SP500_t$, EMP_t, $\hat{g}_{1t}$, $\hat{g}_{2t}$,

$\hat{g}_{3t}$), with probit models restricted to a maximum number of three predictors (plus an intercept). In total, sixty-three alternative models are evaluated.[6]

Predicted recession probabilities for months $t+1$ through $t+h$ are generated based the information available at month t. In the case of the financial indicators, the factor $\hat{g}_{1t}$ (bond and exchange rates), and the factor $\hat{g}_{2t}$ (stock market), the information set includes data up to time t. In the case of the real activity indicator and the factor $\hat{g}_{3t}$ (real activity), however, the information set includes data only up to time $t-1$. As a result, real activity indicators enter the predictive regressions lagged one month.

I evaluate the in-sample fit of each candidate model using McFadden's pseudo-R^2 (R^2_{mf}) and the Bayesian Information Criterion (BIC). The R^2_{mf} is defined as

$$R^2_{mf} = 1 - \frac{\ln \hat{L}}{\ln L_0}, \tag{5}$$

where $\ln \hat{L}$ is the value of the log likelihood function evaluated at the estimated parameters and $\ln L_0$ is the log likelihood computed only with a constant term. The BIC for a model with k predictors is defined by

$$\text{BIC} = \ln \hat{\sigma}^2 + \frac{k \ln T}{T}, \tag{6}$$

where $\hat{\sigma}$ is the regression's standard error and T is the sample size.

Out-of-sample predicted probabilities of recession are evaluated using two statistics. The first statistic is the quadratic probability score (QPS), equivalent to the mean squared error, which is defined by

$$\text{QPS} = \frac{2}{T^*} \sum_{t=1}^{T^*} (y_{t\,t+h} - \hat{p}_{t\,t+h})^2, \tag{7}$$

[6] The 63 models include 35 three-predictor models, 21 two-predictor models, and 7 one-predictor models. In addition, 14 of these 63 models have only individual indicators and 7 models have only factors as predictors. The remaining 42 models have a combination of individual indicators and factors.

where T^* is the effective number of out-of-sample forecasts and $\hat{p}_{t\,t+h}$ is the predicted probability of recession for months $t+1$ through $t+h$ for a given model. The QPS can take values from 0 to 2 and smaller values indicate more accurate predictions. Finally, recession probabilities are evaluated using the log probability score (LPS), which is given by

$$\text{LPS} = -\frac{1}{T^*}\sum_{t=1}^{T^*}\left[y_{t\,t+h}\log(\hat{p}_{t\,t+h}) + (1-y_{t\,t+h})\log(1-\hat{p}_{t\,t+h})\right]. \tag{8}$$

The LPS can take values from 0 to $+\infty$ and smaller values indicate more accurate predictions. Compared to the QPS, the LPS score penalizes large errors more heavily. See, e.g., Katayama (2010) and Owyang et al. (2012).

4 Results

4.1 In-Sample Results

Each of the predictive probit regressions is first estimated using data starting in 1967:1 and ending in 2005:12, as in Wright (2006) and Katayama (2010). Next, the end of the sample is set at 2010:12 in order to include the 2008-09 recession. In both cases, data corresponds to the February 2011 vintage and the macro factors for the in-sample analysis are estimated using the full sample of time series information. The results are shown for three alternative forecast horizons: $h = 3$, 6, and 12 months.

While the analysis considers a total of sixty-three alternative predictive models, results for only ten models are discussed here.[7] Table 1 summarizes, in no particular order, the ten models discussed in this paper. Model 1, the baseline model, uses the 3-month less 10-year term spread as predictor. Model 2 uses both the term spread and

[7] Results for the remaining models are available upon request.

the level of the federal funds rate. This model is found to give the best performance in Wright (2006). Models 3 uses both the term spread and month-to-month percentage changes in the S&P 500 stock market index. This model is found to give the best performance in Estrella and Mishkin (1998). Model 4 adds the growth rate of non-farm employment to model 3. This is the best performing model in Katayama (2010). The remaining six models are the ones that exhibit the best performance in this paper: i.e., models that were ranked at least as top 3 by at least one of the forecast evaluation statistics discussed in the previous section. Note that all the top performing models include at least one macro factor as predictor. For example, models 5 and 6 include the real activity factor, model 7 includes the stock market factor, and models 8 and 9 include both the real activity and stock market factors. Finally, model 10 is a factor-only probit model.

[TABLE 1 ABOUT HERE]

Table 2 (panel A) reports the in-sample R^2_{mf} and BIC for $h = 3$ months. Several results stand out. First, probit models based only on financial indicators (models 1, 2, and 3) exhibit a large deterioration in fit after 2005. For example, in the case of model 2 (Wright, 2006), the R^2_{mf} falls 62% when the sample is extended to include the 2008-09 recession. This result is consistent with results reported in Ng and Wright (2013) who find that the predictive power of interest rate spreads substantially deteriorates at the end of the sample.[8] On the other hand, probit models that use the real activity

[8]Ng and Wright (2013) attribute this deterioration to changes in the causes of the last recession compared to previous recessions. In particular, they write:

> "The recessions of the early 1980s were caused by the Fed tightening monetary policy so as to lower inflation, with the effect of generating both an inverted yield curve and two recessions. The origins of the Great Recession were instead in excess leverage and a housing/credit bubble."

See also Stock and Watson (2012) for a detailed analysis of the 2008-09 recession.

indicator directly (models 4 and 7) or the real activity factor (models 5, 6, 8, 9, and 10) maintain their fit throughout the sample and exhibit a better forecasting performance during the 2008-09 recession. As noted in Estrella and Mishkin (1998), Katayama (2010), and Owyang et al. (2012), among others, real economic activity indicators can improve recession forecasts, particularly at short horizons. Second, probit models that use macro factors as predictors yield better in-sample fit than models that use macroeconomic indicators directly. For example, replacing employment with the real activity factor (i.e., comparing models 4 and 6) improves R^2_{mf} by 18% and the overall model's ranking, based on the BIC, from 26th to 4th. In fact, based on both the R^2_{mf} and the BIC, all the top ranked models include at least one macro factor as predictor (see Table 1). Overall, model 8 (which uses $310TS$, $\hat{g}_{2t}$, and $\hat{g}_{3t}$) is the best fitting model when looking at recessions over the next 3 months.

[TABLE 2 ABOUT HERE]

Tables 3 and 4 (panel A) report the in-sample R^2_{mf} and BIC for $h = 6$ and $h = 12$ months, respectively. The main results are similar to those found for $h = 3$. First, probit models that use both financial and real activity indicators yield better in-sample fit than models with financial indicators alone. Second, probit models that use macro factors as predictors of NBER recessions give better fit than models that use indicators directly. Relative to the models proposed in Estrella and Mishkin (1998), Wright (2006), and Katayama (2010), the improvement can be substantial. For example, model 8 improves the R^2_{mf} of these models by 14% to 285%. Finally, based on the BIC, model 8 is the best fitting model at all horizons.

[TABLES 3, AND 4 ABOUT HERE]

4.2 Out-of-Sample Results

To provide a more accurate assessment of the predictive regressions, in this section I evaluate the out-of-sample performance of the models in two (pseudo) real-time forecasting exercises. The first exercise uses ex-post revised data, corresponding to the February 2011 vintage, to generate out-of-sample predicted recession probabilities for each of the sixty-three models and the three forecast horizons (h = 3, 6, and 12 months). The first forecast is made for 1988:2 and the last for 2010:12 $- h$. As a result, the hold-out sample includes 272 out-of-sample predictions when $h = 3$, 269 predictions when $h = 6$, and 263 predictions when $h = 12$. In these three cases, the hold-out sample includes the last three recessions. The dynamic factors are estimated recursively, each period using revised data up to time t, and expanding the estimation window by one observation each month. The probit models are also estimated recursively and used to generate a recession probability for months $t+1$ through $t+h$ based the information available at month t. Again, in the case of the financial indicators and the macro factors $\hat{g}_{1t}$ and $\hat{g}_{2t}$, the information set includes data up to t. In the case of the real activity indicators and the real activity factor $\hat{g}_{3t}$, the information set includes data only up to $t - 1$ (i.e., lagged one month).

Table 2 (panel B) reports the out-of-sample QPS and LPS for $h = 3$ months. These forecast evaluation statistics suggest that the out-of-sample performance of the models that use macro factors is better than the models that use macroeconomic indicators directly. For example, replacing employment with the real activity factor (i.e., comparing models 4 and 6) improves the model's ranking from 8th to 2nd based on the QPS and from 12th to 2nd based on the LPS. Furthermore, all the top ranked models include at least one macro factor as predictor, a result consistent with what was found in-sample. Overall, model 8 (which uses $310TS$, $\hat{g}_{2t}$, and $\hat{g}_{3t}$) is the best fitting model when looking at recessions over the next 3 months.[9] Relative to model 4, the model found to give the best performance in Katayama (2010), model 8 reduces the QPS by 13% and the LPS by 12%. Again, the general result is that probit models that use real activity indicators directly or via the real factor exhibit a substantially better forecasting performance than models based only on financial indicators. Relative to models proposed in Estrella and Mishkin (1998) and Wright (2006), model 8 reduces the QPS by 44% to 48% and the LPS by 44% to 51% when looking at a 3-month horizon.

Tables 3 and 4 (panel B) report the out-of-sample QPS and LPS for $h = 6$ and $h = 12$ months, respectively. The main results are similar to those found for $h = 3$. First, probit models that use macro factors give better out-of-sample fit than models that use indicators directly. Additionally, models that use both financial and real activity indicators give better out-of-sample fit than models with financial indicators alone. Overall, model 8 is the best fitting model at all horizons. For $h = 6$ and $h = 12$, the improvement relative to the models proposed in Estrella and Mishkin (1998) and Wright (2006) can be substantial. For example, model 8 reduces the QPS by 41% to 45% and the LPS by 37% to 49%. The improvement in out-of-sample forecasting performance relative to the model proposed in Katayama (2010) is smaller. Specifically, model 8 reduces the QPS by 11% to 17% and the LPS by 10% to 11%.

The second exercise examines the robustness of the results obtained above using real-time vintage data (i.e., data as it was available at the time the prediction would have been generated) instead of using ex-post revised data. This, of course, is only relevant for the real activity indicators. Again, the first forecast is made for 1988:2

[9] Based on the out-of sample performance, model 8 exhibits only a slight edge over model 6. As a result, the main source of improvement appears to be the use of the real activity factor instead of employment in the forecasting models.

and the last for 2010:12 $-$ h. Macro factors are estimated recursively, each period using real-time data available at time t, and expanding the estimation window by one observation each month. The probit models are also estimated recursively and used to generate a recession probability for months $t+1$ through $t+h$ based the information available at month t with the real activity indicators and the real activity factor lagged one month.

Tables 2, 3, and 4 (panel C) report the out-of-sample QPS and LPS using real-time data for $h = 3$, 6, and 12 months, respectively. The results from this exercise confirm the overall conclusions from the previous exercise using revised data. In particular, models that use macro factors give better out-of-sample fit than models that use macroeconomic indicators directly. Based on real-time predicted recession probabilities, model 8 is again the best fitting model at all horizons. For example, relative to the models proposed in Estrella and Mishkin (1998) and Wright (2006), model 8 reduces the QPS by 32% to 41% and the LPS by 29% to 45%. Relative to the model proposed in Katayama (2010), model 8 reduces the QPS by 14% to 16% and the LPS by 12% to 14%. As a result, the improvement in out-of-sample forecasting performance relative to the model proposed in Katayama (2010) is more important when using real-time data.

Figures 2, 3, and 4 show the out-of-sample predicted probabilities of recession based on revised and real-time data for $h = 3$, 6, and 12 months, respectively. Results for models 2, 3, 4, and 8 are presented. NBER recession months are shown as shaded areas. Vertical lines before each recession indicate the date we would like to see the probabilities rise (i.e., h months before the beginning of the recession). As noted above, the out-of-sample predictive performance of model 2 (Estrella and Mishkin, 1998) and model 3 (Wright, 2006) is poor. For $h = 3$ and $h = 6$, recession probabilities from these models are low for most of the hold-out sample. For $h = 12$, predicted probabilities are high before actual recession periods, consistent with the improved R^2_{mf} found in-sample, but drop too soon. On the other hand, including real activity indicators contributes to make stronger predictions with probabilities that are closer to 1 before and during NBER recessions. As can be seen, model 4 (Katayama, 2010) and model 8 (which uses $310TS$, $\hat{g}_{2t}$, and $\hat{g}_{3t}$) exhibit a better performance, with high predicted probabilities preceding actual recession periods. Overall, model 8 is the best performing model as it generates recession probabilities that are smooth and closer to 0 during expansions and to 1 during recessions, a result consistent with the out-of-sample QPS and LPS scores reported above.

[FIGURES 2, 3, AND 4 ABOUT HERE]

Another result that emerges from these figures is that recession probabilities generated by model 8 are smooth when computed using both ex-post revised data as well as with real-time data. In fact, in the case of model 8, recessions probabilities generated with real-time data generally overlap with probabilities generated using revised data. On the other hand, model 4 generates recession probabilities that are smooth when computed using ex-post revised data while much more volatile when using real-time data. This result is consistent with the larger deterioration in out-of-sample forecasting performance observed in model 4 (relative to model 8 and other models using macro factors) when probabilities are generated using real-time data. Therefore, this result provides some evidence on the issue of data revisions and factor models. As conjectured in Berge and Jorda (2011) and Chen et al. (2011), data revisions appear to affect the real activity factor less than the individual real activity indicators (in this case employment).

5 Conclusion

This paper uses dynamic latent factors estimated from small panels of macroeconomic indicators to predict future NBER recession dates. The results show that probit models based on macro factors exhibit a better predictive performance than models that use macroeconomic indicators directly. These results hold in-sample and out-of-sample, for a forecasting horizon of 3, 6, and 12 months, and using both ex-post revised data and real-time data. Additionally, this paper shows that data revisions appear to affect the macro factors less than the individual indicators. Overall, probit models based on macro factors provide the best and most robust predictive performance for NBER recessions at all horizons considered in this paper.

6 Data Appendix

The following table lists the short name, transformation applied, and a data description of each series in the three groups considered. All bond, exchange rates, and stock market series are from FRED (St. Louis Fed), unless the source is listed as GFD (Global Financial Data), or AC (author's calculation). Vintage data for the real factor are from Camacho et al. (2013). The transformation codes are: 1 = no transformation; 2 = first difference; 3 = first difference of logarithms.

	Short Name	Trans.	Description
Bond and Exchange Rates Factor			
1	Fed Funds	2	Interest Rate: Federal Funds (Effective) (% per annum)
2	Comm paper	2	Commercial Paper Rate
3	3-m T-bill	2	Interest Rate: U.S.Treasury Bills, Sec Mkt, 3-Mo. (% per annum)
4	6-m T-bill	2	Interest Rate: U.S.Treasury Bills, Sec Mkt, 6-Mo. (% per annum)
5	1-y T-bond	2	Interest Rate: U.S.Treasury Const Maturities, 1-Yr. (% per annum)
6	5-y T-bond	2	Interest Rate: U.S.Treasury Const Maturities, 5-Yr. (% per annum)
7	10-y T-bond	2	Interest Rate: U.S.Treasury Const Maturities, 10-Yr. (% per annum)
8	AAA bond	2	Bond Yield: Moody's AAA Corporate (% per annum) (GFD)
9	BAA bond	2	Bond Yield: Moody's BAA Corporate (% per annum) (GFD)
10	CP spread	1	Comm paper – Fed Funds (AC)
11	3-m spread	1	3-m T-bill – Fed Funds (AC)
12	6-m spread	1	6-m T-bill – Fed Funds (AC)
13	1-y spread	1	1-y T-bond – Fed Funds (AC)
14	5-y spread	1	5-y T-bond – Fed Funds (AC)
15	10-y spread	1	10-y T-bond – Fed Funds (AC)
16	AAA spread	1	AAA bond – Fed Funds (AC)
17	BAA spread	1	BAA bond – Fed Funds (AC)
18	Ex rate: index	3	Exchange Rate Index (Index No.) (GFD)
19	Ex rate: Swit	3	Foreign Exchange Rate: Switzerland (Swiss Franc per U.S.$)
20	Ex rate: Jap	3	Foreign Exchange Rate: Japan (Yen per U.S.$)
21	Ex rate: U.K.	3	Foreign Exchange Rate: United Kingdom (Cents per Pound)
22	Ex rate: Can	3	Foreign Exchange Rate: Canada (Canadian$ per U.S.$)
Stock Market Factor			
1	S&P 500	3	S&P's Common Stock Price Index: Composite (1941-43=10) (GFD)
2	S&P indst	3	S&P's Common Stock Price Index: Industrials (1941-43=10) (GFD)
3	S&P div yield	3	S&P's Composite Common Stock: Dividend Yield (% per annum) (GFD)
4	S&P PE ratio	3	S&P's Composite Common Stock: Price-Earnings Ratio (%) (GFD)
Real Factor			
1	IP	3	Industrial Production Index - Total Index
2	PILT	3	Personal Income Less Transfer Payments
3	MTS	3	Manufacturing and Trade Sales
4	Emp: total	3	Employees On Nonfarm Payrolls: Total Private

References

Bellégo, C., and Ferrara, L. (2012): “Macro-financial linkages and business cycles: A factor-augmented probit approach”, *Economic Modelling*, 29, 1793–1797.

Berge, T.J. (2014): “Predicting recessions with leading indicators: Model averaging and selection over the business cycle”, *Federal Reserve Bank of Kansas City Working Paper 13-05*.

Berge, T.J., and Jorda, O. (2011): “Evaluating the Classification of Economic Activity in Recessions and Expansions”, *American Economic Journal: Macroeconomics*, 3, 246–277.

Camacho, M., Perez-Quiros, G., and Poncela, P. (2013): “Extracting Nonlinear Signals from Several Economic Indicators”, Banco de España.

Chauvet, M. (1998): “An Econometric Characterization of Business Cycle Dynamics with Factor Structure and Regime Switches”, *International Economic Review*, 39(4), 969–996.

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. (2005): “Forecasting Recessions Using the Yield Curve”, *Journal of Forecasting*, 24(2), 77–103.

Chen, Z., Iqbal, A., and Lai, H. (2011): “Forecasting the Probability of Recessions: a Probit and Dynamic Factor Modelling Approach”, *Canadian Journal of Economics*, 44 (2), 651–672.

Diebold, F., and Rudebusch, G. (1996): "Measuring Business Cycles: A Modern Perspective", *Review of Economics and Statistics*, 78, 66–77.

Dueker, M.J. (1997): "Strengthening the Case for the Yield Curve as a Predictor of U.S. Recessions", *Federal Reserve Bank of St. Louis Economic Review*, 79, 41–51.

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

Fossati, S. (2014): "Dating U.S. Business Cycles with Macro Factors", unpublished.

Hamilton, J. (2011): "Calling Recessions in Real Time", *International Journal of Forecasting*, 27, 1006–1026.

Katayama, M. (2010): "Improving Recession Probability Forecasts in the U.S. Economy", unpublished.

Kauppi, H., and Saikkonen, P. (2008): "Predicting U.S. Recessions with Dynamic Binary Response Models", *Review of Economics and Statistics*, 90(4), 777–791.

Kim, C.J., and Nelson C.R. (1998): "Business Cycle Turning Points, a New Coincident Index, and Tests of Duration Dependence Based on a Dynamic Factor Model with Regime Switching", *Review of Economics and Statistics*, 80, 188–201.

Kim, C.J., and Nelson C.R. (1999): *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, The MIT Press.

Ludvigson, S.C., and Ng, S. (2009): "A Factor Analysis of Bond Risk Premia", forthcoming in *Handbook of Applied Econometrics*.

Ng, E.C.Y. (2012): "Forecasting U.S. Recessions with Various Risk Factors and Dynamic Probit Models", *Journal of Macroeconomics*, 34(1), 112–125.

Ng, S. (2014): "Boosting Recessions", *Canadian Journal of Economics*, Vol. 47 (1).

Ng, S., and Wright, J.H. (2013): "Facts and Challenges from the Great Recession for Forecasting and Macroeconomic Modeling", NBER Working Paper No. 19469.

Nyberg, H. (2010): "Dynamic Probit Models and Financial Variables in Recession Forecasting", *Journal of Forecasting*, 29, 215–230.

Owyang, M.T., Piger, J.M., and Wall, H.J. (2012): "Forecasting National Recessions Using State Level Data", working paper 2012-013A, Federal Reserve Bank of St. Louis.

Stock, J.H., and Watson, M.W. (1991): "A Probability Model of the Coincident Economic Indicators", in *Leading Economic Indicators: New Approaches and Forecasting Records*, edited by K. Lahiri and G, Moore, Cambridge University Press.

Stock, J.H., and Watson, M.W. (2002a): "Forecasting Using Principal Components From a Large Number of Predictors", *Journal of the American Statistical Association*, 97, 1167–1179.

Stock, J.H., and Watson, M.W. (2002b): "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business and Economic Statistics*, 20, 147–162.

Stock, J.H., and Watson, M.W. (2006): "Forecasting with Many Predictors", in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. Granger, and A. Timmermann, 1, 515–554. Elsevier.

Stock, J.H., and Watson, M.W. (2012): "Disentangling the Channels of the 2007-2009 Recession", NBER Working Paper No. 18094.

Wright, J.H. (2006): "The Yield Curve and Predicting Recessions", Finance and Economics Discussion Series, Federal Reserve Board.

Table 1: Selected Forecasting Models for NBER Recessions

Predictors			Model									
			1	2	3	4	5	6	7	8	9	10
1	$310TS_t$	3-Month less 10-Year Spread	✓	✓	✓	✓	✓	✓	✓	✓		
2	FFR_t	Federal Funds Rate		✓			✓				✓	
3	$SP500_t$	S&P 500 (% change)			✓	✓		✓				
4	EMP_{t-1}	Employment (% change)				✓			✓			
5	$\hat{g}_{1t}$	Bond and Exchange Rates Factor										✓
6	$\hat{g}_{2t}$	Stock Market Factor							✓	✓	✓	✓
7	$\hat{g}_{3t-1}$	Real Factor					✓	✓		✓	✓	✓

Table 2: Forecasting NBER Recessions Over the Next 3 Months

Sample	Statistic	Model									
		1	2	3	4	5	6	7	8	9	10
(A) In-Sample Fit											
1967:01 – 2005:12	R^2_{mf}	0.10	0.21	0.11	0.39	0.49	0.47	0.41	**0.49**	0.49	0.49
1967:01 – 2010:12	R^2_{mf}	0.05	0.08	0.07	0.38	0.45	0.45	0.40	**0.48**	0.47	0.47
	BIC	-1.92	-1.98	-1.95	-2.30	-2.39	-2.41	-2.34	**-2.47**	-2.46	-2.46
	Ranking	63	55	60	26	7	4	17	**1**	2	3
(B) Out-of-Sample Fit: Revised Data											
1988:02 – 2010:09	QPS	0.27	0.27	0.25	0.16	0.17	0.15	0.16	**0.14**	0.16	0.15
	Ranking	61	59	46	8	12	2	7	**1**	5	3
	LPS	0.44	0.47	0.41	0.26	0.25	0.23	0.26	**0.23**	0.25	0.25
	Ranking	47	61	45	12	8	2	11	**1**	4	5
(C) Out-of-Sample Fit: Real-Time Data											
1988:02 – 2010:09	QPS	—	—	—	0.19	0.19	0.16	0.19	**0.16**	0.17	0.17
	Ranking				10	15	2	13	**1**	4	6
	LPS	—	—	—	0.30	0.29	0.26	0.30	**0.26**	0.27	0.28
	Ranking				14	7	2	13	**1**	4	5

Note: R^2_{mf} is McFadden's pseudo-R^2 and BIC is the Bayesian Information Criterion from the maximum likelihood estimation of the probit models at a horizon of h months. QPS is the quadratic probability score and LPS is the log probability score. Models are ranked from 1 to 63.

Table 3: Forecasting NBER Recessions Over the Next 6 Months

Sample	Statistic	Model									
		1	2	3	4	5	6	7	8	9	10
(A) In-Sample Fit											
1967:01 – 2005:12	R^2_{mf}	0.18	0.29	0.19	0.43	0.52	0.51	0.45	**0.53**	0.50	0.52
1967:01 – 2010:12	R^2_{mf}	0.10	0.14	0.13	0.41	0.47	0.47	0.43	**0.50**	0.46	0.49
	BIC	-1.84	-1.92	-1.89	-2.26	-2.34	-2.36	-2.31	**-2.42**	-2.33	-2.39
	Ranking	61	53	58	19	4	3	9	**1**	6	2
(B) Out-of-Sample Fit: Revised Data											
1988:02 – 2010:06	QPS	0.31	0.31	0.29	0.19	0.20	**0.17**	0.19	0.17	0.20	0.20
	Ranking	57	56	46	3	11	**1**	5	2	10	8
	LPS	0.49	0.53	0.46	0.30	0.30	0.27	0.30	**0.27**	0.31	0.30
	Ranking	47	59	45	8	5	2	9	**1**	12	10
(C) Out-of-Sample Fit: Real-Time Data											
1988:02 – 2010:06	QPS	—	—	—	0.22	0.23	0.19	0.23	**0.19**	0.22	0.22
	Ranking				7	8	2	12	**1**	5	6
	LPS	—	—	—	0.35	0.34	0.30	0.35	**0.30**	0.33	0.33
	Ranking				11	9	2	9	**1**	5	6

Note: R^2_{mf} is McFadden's pseudo-R^2 and BIC is the Bayesian Information Criterion from the maximum likelihood estimation of the probit models at a horizon of h months. QPS is the quadratic probability score and LPS is the log probability score. Models are ranked from 1 to 63.

Table 4: Forecasting NBER Recessions Over the Next 12 Months

Sample	Statistic	Model									
		1	2	3	4	5	6	7	8	9	10
(A) In-Sample Fit											
1967:01 – 2005:12	R^2_{mf}	0.31	0.45	0.31	0.47	**0.59**	0.52	0.48	0.53	0.45	0.58
1967:01 – 2010:12	R^2_{mf}	0.21	0.25	0.23	0.42	0.50	0.47	0.44	0.49	0.39	**0.51**
	BIC	-1.82	-1.92	-1.87	-2.14	-2.31	-2.27	-2.18	**-2.32**	-2.08	-2.30
	Ranking	46	33	43	17	2	5	13	**1**	21	3
(B) Out-of-Sample Fit: Revised Data											
1988:02 – 2009:12	QPS	0.36	0.36	0.34	0.24	0.26	0.21	0.24	**0.21**	0.29	0.28
	Ranking	48	45	38	6	7	2	3	**1**	14	10
	LPS	0.54	0.63	0.51	0.37	0.39	0.33	0.36	**0.33**	0.42	0.41
	Ranking	43	56	36	6	8	2	5	**1**	13	10
(C) Out-of-Sample Fit: Real-Time Data											
1988:02 – 2009:12	QPS	—	—	—	0.27	0.28	0.24	0.27	**0.23**	0.30	0.30
	Ranking				4	7	2	3	**1**	10	13
	LPS	—	—	—	0.41	0.43	0.36	0.41	**0.36**	0.44	0.45
	Ranking				6	7	2	5	**1**	11	12

Note: R^2_{mf} is McFadden's pseudo-R^2 and BIC is the Bayesian Information Criterion from the maximum likelihood estimation of the probit models at a horizon of h months. QPS is the quadratic probability score and LPS is the log probability score. Models are ranked from 1 to 63.

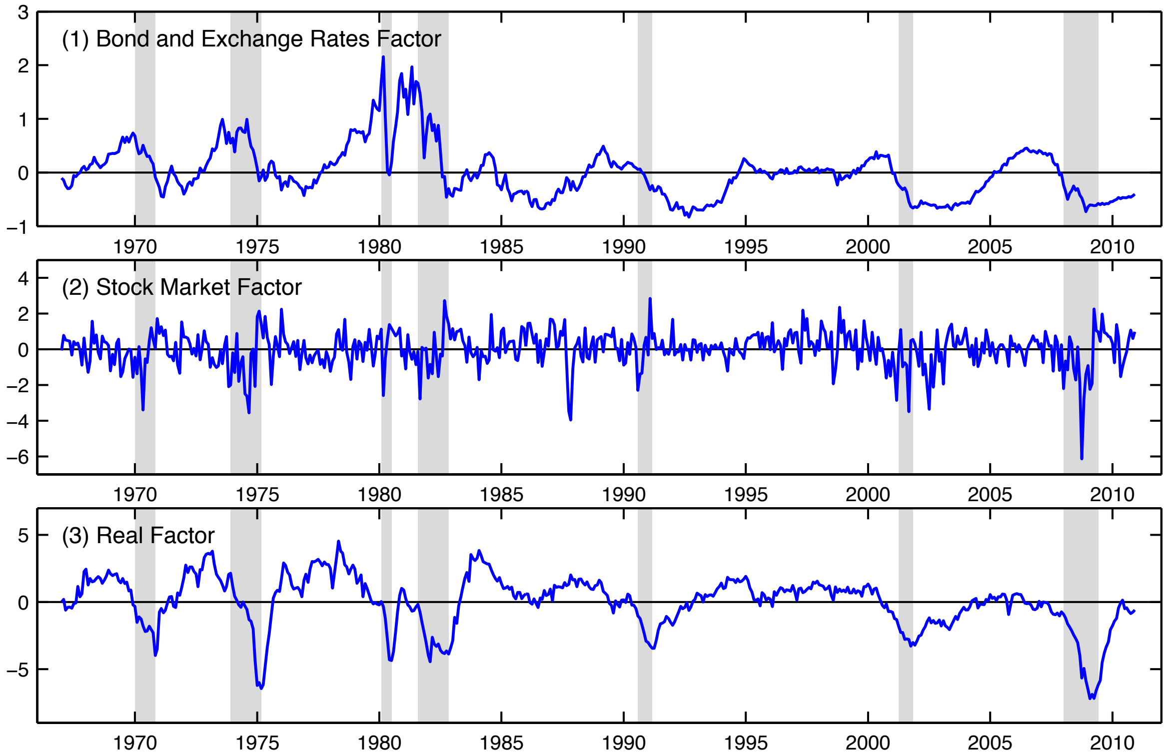

Figure 1: Estimated dynamic macro factors (posterior means) for the full sample. Shaded areas denote NBER recession months.

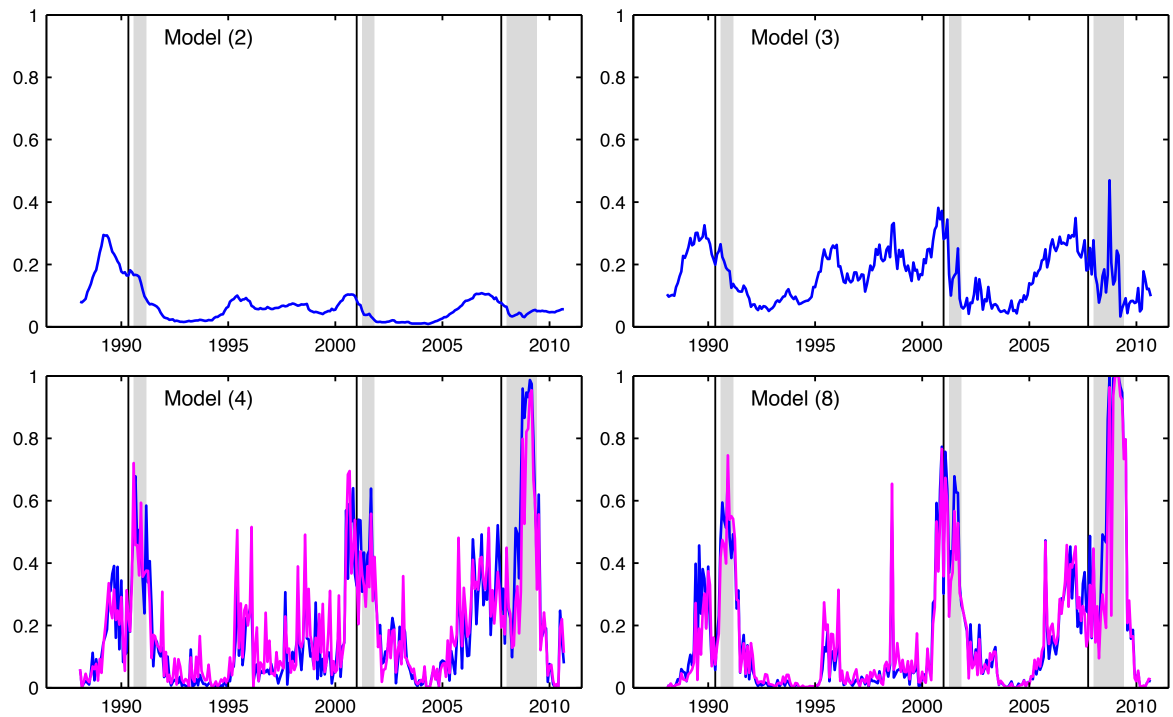

Figure 2: Out-of-sample predicted probabilities of a recession within the next 3 months for models 2, 3, 4, and 8. Revised (blue / dark) and real-time data (magenta / light). Shaded areas denote NBER recession months. Vertical lines denote the date we would like to see the probabilities rise (i.e., 3 months before the beginning of the recession).

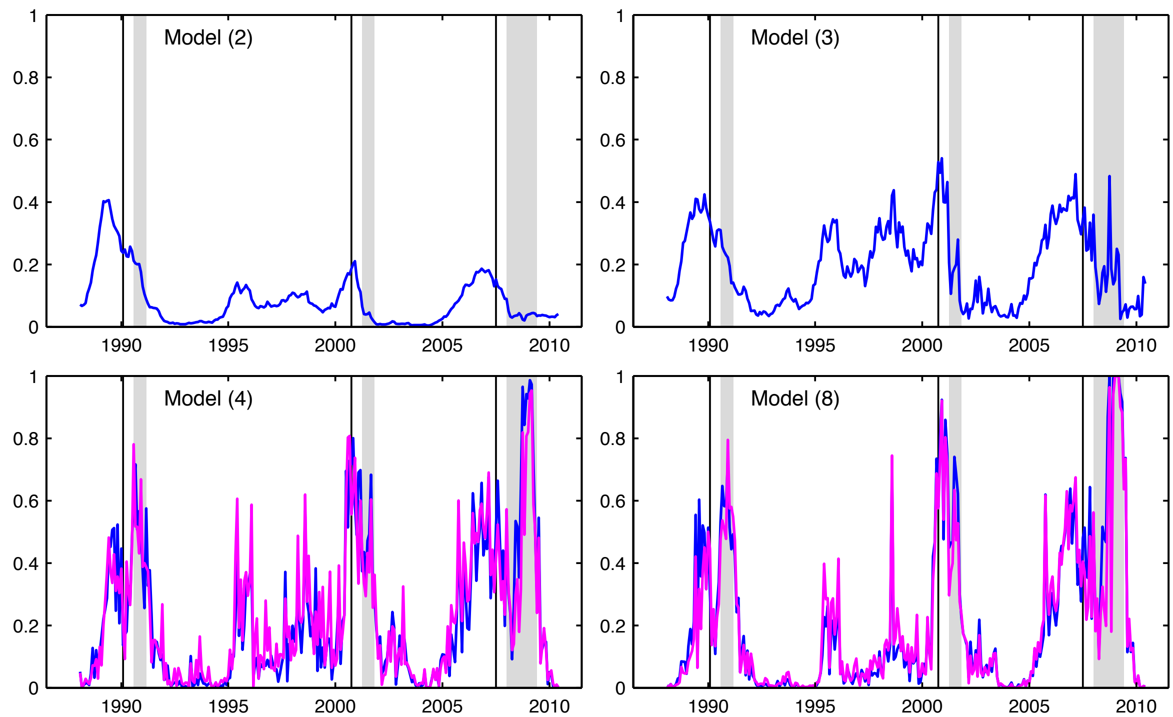

Figure 3: Out-of-sample predicted probabilities of a recession within the next 6 months for models 2, 3, 4, and 8. Revised (blue / dark) and real-time data (magenta / light). Shaded areas denote NBER recession months. Vertical lines denote the date we would like to see the probabilities rise (i.e., 6 months before the beginning of the recession).

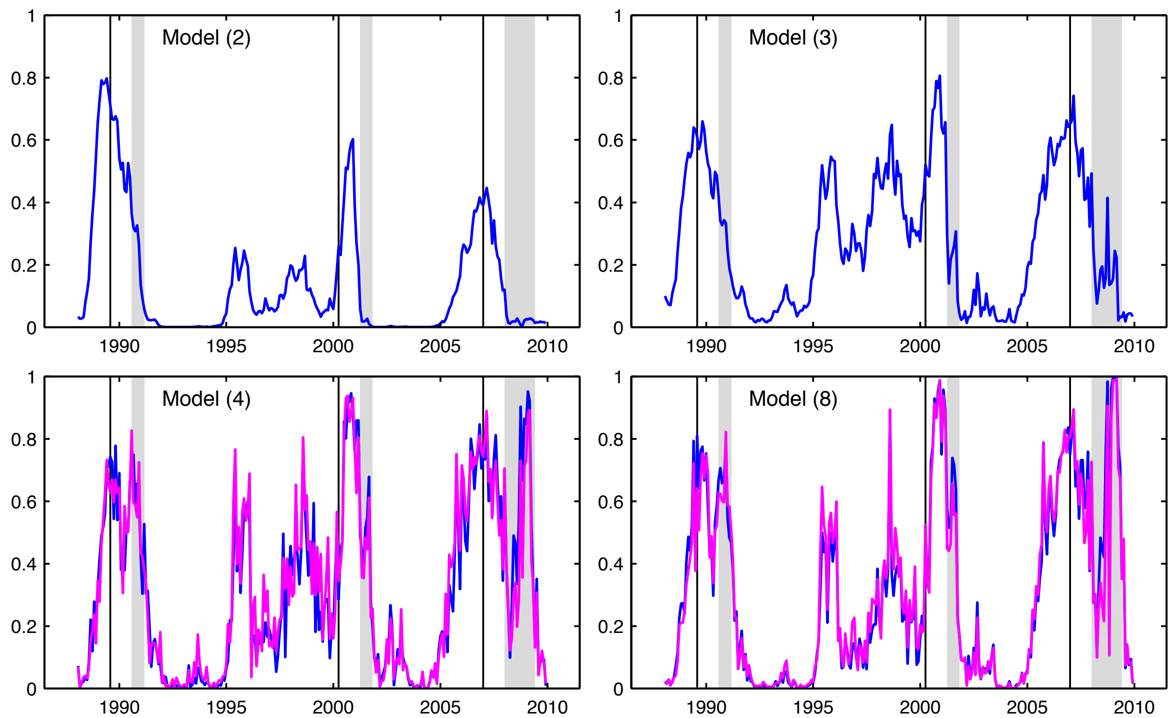

Figure 4: Out-of-sample predicted probabilities of a recession within the next 12 months for models 2, 3, 4, and 8. Revised (blue / dark) and real-time data (magenta / light). Shaded areas denote NBER recession months. Vertical lines denote the date we would like to see the probabilities rise (i.e., 12 months before the beginning of the recession).