Dating U.S. Business Cycles with Macro Factors

Sebastian Fossati*

University of Washington
Job Market Paper

This version: November 12, 2009

Abstract

I propose a framework for the assessment of current business conditions using a factor-augmented autoregressive probit model where the dependent variable is the state of the economy as defined by the NBER. Results show that latent common factors estimated by principal components analysis from a large number of macroeconomic series have important predictive power for NBER recession dates. The models generate in-sample recession probabilities that almost perfectly reproduce NBER dates. A pseudo out-of-sample forecasting exercise, designed to approximate real time conditions, shows that predicted recession probabilities consistently rise during subsequently declared NBER recession dates. With the appropriate classification rule, the models exhibit good performance as real time dating algorithms. The latent variable in the probit model can be interpreted as an index of business conditions which is used to assess the strength of an expansion or the depth of a recession.

Keywords: Business Cycle, Forecasting, Factors, Probit Model, Bayesian Methods.
JEL Codes: E32, E37, C01, C22, C25.

---

*Department of Economics, University of Washington, Box 353330, Seattle, WA 98195-3330. Email: sfossati@u.washington.edu. http://students.washington.edu/sfossati/. I thank Eric Zivot and Drew Creal for helpful advice and support. I also thank Byron Tsang and seminar participants at the 2009 NBER–NSF Time Series Conference (University of California, Davis) for helpful comments. Support from the Grover and Creta Ensley Fellowship is gratefully acknowledged.

1 Introduction

Is the U.S. economy in recession? This was one of the central questions in the business and policy communities during the year 2008. While the consensus among analysts was that the economy was in fact in recession, most business cycle indicators failed to signal the downturn.[1] This question was answered in December 2008 when the Business Cycle Dating Committee of the NBER determined that a peak in economic activity (beginning of a recession) occurred in the U.S. economy in December 2007. The year 2009 brought forth several related questions: Is the U.S. economy still in recession? How deep is the current recession? Is it a depression? What is the shape of the recession? V, U, L-shaped? Answering these questions in real time is not an easy task since business conditions are not observable, and NBER announcements come out long after the fact.[2]

With these questions in the background, I propose a framework for the assessment of current business conditions. Instead of relying on a small number of observed variables (see, e.g., Chauvet and Potter, 2008; Aruoba et al., 2009), I consider the information contained in a large number of macroeconomic series using a factor-augmented autoregressive probit model where the dependent variable is the state of the economy (recession or expansion) as defined by the NBER. This paper contributes to the literature on business cycle modeling by showing that latent common factors estimated by principal components from a large number of macroeconomic time series have important predictive power for NBER recession dates. The main driving force of this result

[1] For example, Krugman (2008) writes: "Suddenly, the economic consensus seems to be that the implosion of the housing market will indeed push the U.S. economy into a recession, and that it's quite possible that we're already in one". Leamer (2008), on the other hand, concludes that: "[The recession-dating] algorithm indicates that the data through June 2008 do not yet exceed the recession threshold, and will do so only if things get much worse".

[2] The NBER takes between 6 to 20 months to announce peaks and troughs.

is a factor that loads heavily on measures of real output and employment, a 'real' factor. This result is in line with recent empirical research using factor models which has found that a few factors extracted from a large number of series can be useful in many forecasting exercises; see, e.g., Stock and Watson (2002a,b, 2006), Ludvigson and Ng (2008a,b), and Giannone et al. (2008).

While recession probabilities have traditionally been generated using Markov switching models as in Hamilton (1989), Chauvet (1998), Chauvet and Hamilton (2006), and Chauvet and Piger (2008), the use of binary class models to predict NBER recession dates is not new in this literature. For example, Estrella and Mishkin (1998), Dueker (1997), Chauvet and Potter (2002, 2005), Kauppi and Saikkonen (2008), and Katayama (2009) examine the usefulness of several economic and financial variables, e.g. the interest rate spread, as predictors of future U.S. recessions. The approach I take is closer to Chauvet and Potter (2008) who consider the performance of four monthly coincident macroeconomic variables as predictors of current (rather than future) business conditions.

The main results of this paper can be summarized as follows. First, the factor-augmented autoregressive probit models proposed here generate in-sample recession probabilities that reproduce NBER dates almost perfectly. Chauvet and Potter (2008) achieve similar in-sample results, but they do so by allowing for recurrent structural breaks. A consequence of using business cycle dependent parameters is that a real time application of their models is not feasible. As shown in Stock and Watson (2008), factor models can be robust to structural instability. The results presented here are obtained using a simple model specification where there is no need to allow for structural breaks or time-varying parameters.

Second, a pseudo out-of-sample forecasting exercise, designed to approximate real time conditions, shows that predicted recession probabilities consistently rise during subsequently declared NBER recession dates. With the appropriate classification rule, the models exhibit good performance as real time dating algorithms. However, since the models' predictive performance is different for expansion and recession periods, this paper emphasizes that model selection requires choosing a loss function that reflects the preferences of the forecaster.

Third, the latent variable in the probit model can be interpreted as an index of business conditions which is used to assess the strength of an expansion or the depth of a recession; see, e.g., Dueker (2005). Similarly, Aruoba et al. (2009) propose a framework that allows for the extraction of the latent state of business conditions at high frequency. Even though the information sets used in these papers differ substantially, the results show that the (standardized) latent variable from the factor-augmented probit models almost perfectly overlap with the index extracted by Aruoba et al. (2009).

The paper is organized as follows. Section 2 presents the factor-augmented autoregressive probit model and discusses its estimation using Bayesian methods. Section 3 presents preliminary results using single-regressor traditional probit models. Section 4 presents in-sample estimation results and out-of-sample forecast results in the form of posterior means. An evaluation of the out-of-sample forecasts and the calibration of an optimal classification rule are also presented in this section. Section 5 concludes.

2 The Econometric Model

This section presents the econometric framework. First, I present the factor-augmented autoregressive probit model and discuss the use of principal components to estimate latent common factors from a large number of macroeconomic series. Subsequently, I discuss how to estimate the autoregressive probit model using Gibbs sampling.

2.1 A Factor-Augmented Autoregressive Probit Model

Define a latent variable y_t^*, which represents the state of the economy as measured by the Business Cycle Dating Committee of the NBER, such that

$$y_t^* = \alpha + \beta' x_t + \epsilon_t, \tag{1}$$

where x_t is a vector of exogenous predictors, (α, β') are regression coefficients, and $\epsilon_t | x_t \sim i.i.d.\ N(0,1)$.[3] We do not observe y_t^*, but rather y_t, which represents the observable recession indicator according to the following rule

$$y_t = \begin{cases} 1 & \text{if} \quad y_t^* \geqslant 0 \\ 0 & \text{if} \quad y_t^* < 0 \end{cases}, \tag{2}$$

where y_t is 1 if the observation corresponds to a recession and 0 otherwise. In the case of the traditional probit model, the conditional probability of recession is

$$p_t = P(y_t = 1 | x_t) = P(y_t^* \geqslant 0 | x_t) = \Phi(\alpha + \beta' x_t), \tag{3}$$

where $\Phi(\cdot)$ is the distribution function of the standard normal.

Since business cycle phases should have a certain duration, an autoregressive term can be included in (1) in order to capture dependence in the latent variable such that

$$y_t^* = \alpha + \beta' x_t + \theta y_{t-1}^* + \epsilon_t, \tag{4}$$

where $|\theta| < 1$. This model is similar to the models considered in Dueker (1999) and Chauvet and Potter (2005, 2008). As in the case of the traditional probit model, the conditional probability of recession is given by

$$p_t = P(y_t = 1 | x_t, y_{t-1}^*) = P(y_t^* \geqslant 0 | x_t, y_{t-1}^*) = \Phi\big(\alpha + \beta' x_t + \theta y_{t-1}^*\big). \tag{5}$$

[3] Note that since y_t^* is not observable, if $\epsilon_t | x_t \sim i.i.d.\ N(0, \sigma^2)$ is assumed, the regression coefficients (α, β') and σ are not separately identified. As a consequence, it is standard to normalize σ to 1.

Chauvet and Potter (2008) analyze the performance of four coincident macroeconomic variables (industrial production, sales, personal income, and employment) as predictors of y_t. Instead of relying on a small number of observable variables, I consider the information contained in a large number of macroeconomic time series. As in Stock and Watson (2002a,b, 2006) and Ludvigson and Ng (2008a,b), among others, consider the case where we observe a $T \times N$ panel of macroeconomic data, where N is large, and possibly larger than T. I want to estimate (1), where x_t denotes the $N \times 1$ vector of panel observations at time t. One way of dealing with the possible degrees of freedom problem is by summarizing the information in the panel using a small number of common factors. Assume x_{it}, $i = 1, ..., N$, $t = 1, ..., T$, has a factor structure of the form

$$x_{it} = \lambda_i' f_t + e_{it}, \tag{6}$$

where f_t is a $r \times 1$ vector of latent factors, λ_i is a $r \times 1$ vector of latent factor loadings, and e_{it} is the idiosyncratic error. Since $r \ll N$, an important dimension reduction can be obtained by considering the factor-augmented regression

$$y_t^* = \alpha + \delta' F_t + \theta y_{t-1}^* + \epsilon_t, \tag{7}$$

where $F_t \subseteq f_t$. Note that F_t does not have to include all elements of f_t, only those that are relevant for predicting y_t^*.

Since the common factors are not observed, we must replace f_t with an estimate $\hat{f}_t$. Stock and Watson (2002a) show that, when $N, T \to \infty$, f_t can be consistently estimated by principal components analysis. Results in Bai and Ng (2006) provide the framework for inference in the linear factor-augmented regression model and show that estimated factors can be used instead of the true factors in this model; see, also, Stock and Watson (2002b, 2006) and Ludvigson and Ng (2008a,b). Similar results for non-linear models, including the probit model, are provided in Bai and Ng (2008). Finally, the number of latent common factors, r, to be estimated by principal components analysis can be determined using model selection criteria as in Bai and Ng (2002). The final regression for the factor-augmented autoregressive probit model is then

$$y_t^* = \gamma' z_t + \theta y_{t-1}^* + \epsilon_t, \tag{8}$$

where $\gamma = (\alpha, \delta')'$ and $z_t = (1, \hat{F}_t')'$.

2.2 Model Estimation

I estimate the model given by (2) and (8) in two steps. First, I estimate the latent common factors by principal components analysis, as explained above, and then I estimate the autoregressive probit model using the estimated factors as predictors. The likelihood function for the model is

$$L\left(y|z, \gamma, \theta, y_0\right) = \prod_{t=1}^{T} \left[\Phi(\gamma' z_t + \theta y_{t-1}^*)\right]^{y_t} \left[1 - \Phi(\gamma' z_t + \theta y_{t-1}^*)\right]^{1-y_t}. \tag{9}$$

Maximum likelihood estimation of dynamic probit models can be quite difficult. The problem is the evaluation of the likelihood function which requires numerical evaluation of a T-variate normal distribution (see Eichengreen et al., 1985). Bayesian methods, however, can greatly simplify the problem. The approach I take consists of using data augmentation via Gibbs sampling, allowing me to treat y_t^* as observed data. This strategy turns the probit model into a standard linear regression model. The implementation of the Gibbs sampler in this context is similar to that of Dueker (1999) and Chauvet and Potter (2005, 2008). After generating initial values of the latent variable y_t^*, the sampler proceeds as follows: (i) generate draws of the latent variable y_t^* conditional on (γ', θ) and the observed data; (ii) generate draws of γ' conditional on (y_t^*, θ) and the observed data; (iii) generate draws of θ conditional on (y_t^*, γ') and the observed data. Prior and posterior distributions are discussed below.

2.2.1 Generating Draws of the Latent Variable

Initial values of the latent variable, $y_t^{*(0)}$ for $t = 1, ..., T$, are drawn from $f(y_t^{*(0)}|y_{t-1}^{*(0)}, y_t)$ with $y_0^{*(0)} = 0$. Conditional on y_{t-1}^* and y_t, y_t^* has a truncated normal distribution where $y_t^* \geq 0$ if $y_t = 1$ and $y_t^* < 0$ if $y_t = 0$. The truncation imposes a sign condition on y_t^* based on the observed value y_t. Then, potential values of $y_t^{*(0)}$ are drawn from $y_t^{*(0)} \sim N(\gamma' z_t + \theta y_{t-1}^{*(0)}, 1)$. Draws are discarded if the sign condition is not satisfied.

Obtaining subsequent draws of the latent variable y_t^* conditional on the parameters and the observed data requires the derivation of the the conditional distribution $y_t^*|y_{t-1}^*, y_{t+1}^*$. Since the vector $(y_{t+1}^*, y_t^*, y_{t-1}^*)$ has a joint normal distribution, the conditional distribution $y_t^*|y_{t-1}^*, y_{t+1}^*$ is also normal. Starting with (8) and substituting backwards for lagged y^*'s on the right side, the following results can be derived

$$y_t^* = \sum_{s=0}^{t-1} \theta^s \gamma' z_{t-s} + \sum_{s=0}^{t-1} \theta^s \epsilon_{t-s}$$

$$E(y_t^*) = A_t = \sum_{s=0}^{t-1} \theta^s \gamma' z_{t-s} = \gamma' z_t + \theta A_{t-1}$$

$$Var(y_t^*) = B_t = \sum_{s=0}^{t-1} \theta^{2s} = 1 + \theta^2 B_{t-1}$$

$$Cov(y_t^*, y_{t-1}^*) = \theta B_{t-1}$$

The joint distribution of the vector $(y_{t+1}^*, y_t^*, y_{t-1}^*)$ is then

$$\begin{bmatrix} y_{t+1}^* \\ y_t^* \\ y_{t-1}^* \end{bmatrix} \sim N\left(\begin{bmatrix} A_{t+1} \\ A_t \\ A_{t-1} \end{bmatrix}, \begin{bmatrix} B_{t+1} & \theta B_t & \theta^2 B_{t-1} \\ & B_t & \theta B_{t-1} \\ & & B_{t-1} \end{bmatrix} \right).$$

Using standard results for the multivariate normal distribution, $y_t^*|y_{t+1}^*, y_{t-1}^* \sim N\big(\tilde{\mu}_t, \tilde{\Sigma}_t\big)$ for $t = 2, ..., T-1$, with truncation such that $y_t^* \geq 0$ if $y_t = 1$ and $y_t^* < 0$ if $y_t = 0$ and

$$\tilde{\mu}_t = A_t + \theta \begin{pmatrix} B_t \\ B_{t-1} \end{pmatrix}' \begin{pmatrix} B_{t+1} & \theta^2 B_{t-1} \\ & B_{t-1} \end{pmatrix}^{-1} \begin{pmatrix} y_{t+1}^* - A_{t+1} \\ y_{t-1}^* - A_{t-1} \end{pmatrix},$$

$$\tilde{\Sigma}_t = B_t - \theta^2 \begin{pmatrix} B_t \\ B_{t-1} \end{pmatrix}' \begin{pmatrix} B_{t+1} & \theta^2 B_{t-1} \\ & B_{t-1} \end{pmatrix}^{-1} \begin{pmatrix} B_t \\ B_{t-1} \end{pmatrix}.$$

Finally, assuming $y_0^* = 0$, $y_1^*|y_2^* \sim N\big(\tilde{\mu}_1, \tilde{\Sigma}_1\big)$, with truncation such that $y_1^* \geq 0$ if $y_1 = 1$ and $y_1^* < 0$ if $y_1 = 0$ and

$$\tilde{\mu}_1 = A_1 + \theta B_1 B_2^{-1}\big(y_2^* - A_2\big) = A_1 + \frac{\theta}{1+\theta^2}\big(y_2^* - A_2\big),$$

$$\tilde{\Sigma}_1 = B_1 - \theta^2 B_1 B_2^{-1} B_1 = 1 - \frac{\theta^2}{1+\theta^2}.$$

Based on these results, subsequent draws of the latent variable, $y_t^{*(i)}$ for $t = 1, ..., T$, are taken from $f(y_t^{*(i)}|y_{t-1}^{*(i-1)}, y_{t+1}^{*(i)}, y_t)$ for $t = 1, ..., T-1$ and $f(y_t^{*(i)}|y_{t-1}^{*(i-1)}, y_t)$ for $t = T$ where i denotes the ith cycle of the Gibbs sampler. As in Chauvet and Potter (2005, 2008), I start drawing a value of y_T^* conditional on a value of y_{T-1}^* and y_T from $y_T^{*(i)} \sim N(\gamma' z_T + \theta y_{T-1}^{*(i-1)}, 1)$, with truncation such that $y_T^{*(i)} \geq 0$ if $y_T = 1$ and $y_T^{*(i)} < 0$ if $y_T = 0$. With this value of y_T^*, I generate draws of y_t^* for $t = 1, ..., T-1$ backwards using the results described above. Potential draws of y_t^* are discarded if the sign condition is not satisfied.

2.2.2 Prior and Posterior for γ

Following Albert and Chib (1993) and Dueker (1999), I use a flat non-informative prior for γ. Initial values for γ in the first cycle of the Gibbs sampler are the least squares estimates from a regression on the observed variable y_t without autoregressive terms. Let $W_t^\gamma = y_t^* - \theta y_{t-1}^*$, then draws of γ are generated from the multivariate normal distribution $\gamma | y^*, \theta, y \sim N\left(\hat{\gamma}, (z'z)^{-1}\right)$ where $\hat{\gamma} = (z'z)^{-1} z' W^\gamma$.

2.2.3 Prior and Posterior for θ

Similarly, I use a flat non-informative prior for the autoregressive parameter θ. The initial value of θ to start the Gibbs sampler is set at 0.5. Let $W_t^\theta = y_t^* - \gamma' z_t$ and $W_t^y = y_{t-1}^*$, with $W_1^y = 0$. Then, potential draws of θ are generated from $\theta | y^*, \gamma, y \sim N\left(\hat{\theta}, (W^{y\prime} W^y)^{-1}\right)$ where $\hat{\theta} = (W^{y\prime} W^y)^{-1} W^{y\prime} W^\theta$. Draws are discarded if the stationarity condition $|\theta| < 1$ is not satisfied.

2.2.4 Recession Probabilities

Conditional recession probabilities are generated at each draw of the Gibbs sampler using (5) such that

$$p_t^{(i)} = \Phi\left(\gamma^{(i)\prime} z_t + \theta^{(i)} y_{t-1}^{*(i)}\right), \tag{10}$$

where i denotes the ith cycle of the Gibbs sampler. The posterior mean probability of recession is given by

$$\hat{p}_t = \frac{1}{I} \sum_{i=1}^{I} p_t^{(i)}, \tag{11}$$

where I denotes the total number of draws.

3 Data and Preliminary Results

The sample period is 1960:4 – 2007:12, and the recession indicator, y_t, is coded according to the business cycle turning points of the NBER: y_t is 1 if the observation corresponds to a recession and 0 otherwise. Common factors are estimated from a balanced panel of 131 monthly U.S. macroeconomic time series, spanning the period 1960:1 – 2007:12. The data set, generously provided by Serena Ng and Sydney Ludvigson, is used in Ludvigson and Ng (2008b) and is an extended version of the one used in Stock and Watson (2002b, 2005, 2006) and Ludvigson and Ng (2008a). The series include a wide range of macroeconomic variables in the broadly defined categories: output and income; employment, hours, and unemployment; inventories, sales, and orders; housing and consumption; international trade; prices and wages; money and credit; interest rates and interest rates spreads; stock market indicators and exchange rates. The data in x_t were transformed in order to ensure stationarity and standardized prior to estimation.[4]

As in Ludvigson and Ng (2008a,b), model selection criteria indicate eight static common factors that are estimated by principal components analysis. Note that the normalization imposed for identification purposes implies that estimated factors are mutually orthogonal. The first factor accounts for the largest amount of total variation in the panel, the second factor accounts for the largest variation in the panel that was not accounted for by the first factor, and so on.[5] Since factors that are important for explaining the total variation in the panel data x_{it} need not be relevant for modeling y_t, the first question is then which estimated factors have predictive power for y_t. To address this question, I estimate eight single-regressor traditional probit models with $y_t^* = \alpha + \delta \hat{f}_{it} + \epsilon_t$ for $i = 1, ..., 8$ and $t = 1, ..., T$ by maximum likelihood. Table 1 reports parameter estimates, McFadden's pseudo-R^2, the value of the log likelihood, and the likelihood ratio (LR) test statistic for the hypothesis that $\delta = 0$ with its associated probability value. Only the estimated first, second, and fourth factors are significant

[4] A complete description of the series and transformations is given in Ludvigson and Ng (2008a,b).
[5] "Total variation" is the sum of the variances of the variables in the panel x.

(p-value < 0.1). The estimated first factor not only explains most of the variation in the panel x, but also has the largest (in-sample) predictive power for y_t with pseudo-$R^2 = 0.555$. The second and fourth estimated factors, although significant, exhibit very low values of pseudo-R^2.

[TABLE 1 ABOUT HERE]

While economic interpretation of the individual factors is difficult because of identification issues, it is sometimes possible to interpret the factors by measuring on which series in the panel they load heavily. Results in Ludvigson and Ng (2008a, Figure 1) show that the first factor loads heavily on real variables such as employment, production, capacity utilization, and manufacturing orders. Figure 1 presents the estimated first factor along with the (standardized) index of capacity utilization. The series are similar, with major troughs corresponding closely to NBER recession dates (shaded areas). As concluded in Stock and Watson (2002b) and Ludvigson and Ng (2008a,b), the first factor appears to be an index of real economic activity. Figure 2 presents the probability of recession estimated from the traditional probit model using the first factor as predictor. Recession probabilities consistently rise during NBER recession dates and the model signals recessions with high probability values. The model, however, shows probabilities that are relatively volatile during recessions and exhibits several false positives during expansions.

[FIGURE 1 ABOUT HERE]

[FIGURE 2 ABOUT HERE]

4 Empirical Results

Before we can estimate the factor-augmented autoregressive probit model presented above, we need to determine which common factors should be included in the regression. The results in the previous section suggest that a probit model with the first estimated factor as predictor is a good starting point. I consider three model specifications: one where only the first factor is included as predictor; one where the first two factors are included; one where the first four factors are included. The equation to be estimated takes the form $y_t^* = \alpha + \sum_{i=1}^{k} \delta_i \hat{f}_{it} + \theta y_{t-1}^* + \epsilon_t$. I will refer to these models as AFP(k) with $k = 1, 2, 4$.

The next section presents in-sample results where the common factors $\hat{f}_t$ at each date t are estimated using the full sample of time series information, and where it is assumed that the entire series of NBER dates is known. To provide a more accurate evaluation of the models, section 4.2 presents out-of-sample results for a pseudo real time recursive exercise. In this case, the factors are estimated recursively, each period using data only up to time t. Furthermore, since NBER dates are not known for some time, I assume that at time t the forecaster does not know whether the true state of the economy has changed over the last 12 months such that $y_{t-i} = y_{t-12}$ for $i = 0, 1, ..., 11$. Further details are given below. Finally, section 4.3 considers the calibration of an optimal classification rule that translates recession probabilities into class predictions using a sequence of observed forecasts.

4.1 In-Sample Results

To estimate (8) for k = 1, 2, and 4, the Gibbs sampler was run with 50,000 iterations each time. After discarding the first 10,000 draws (burn-in period), posterior means are computed using a thinning factor of 40, i.e. computed from every 40th draw. As a consequence, the subsequent analysis is based on the means of these 1000 draws.

The top part of Table 2 reports the posterior mean and standard deviation of the models' parameters. For the three autoregressive models, the posterior distributions of δ_1 and δ_2 are concentrated away from zero and the first two factors appear to be important. The third and fourth factors, however, do not appear to be important since the posterior distributions of δ_3 and δ_4 are concentrated near zero.[6] The results also show a posterior mean value of the autoregressive coefficient that is just above 0.7 for the three models. This result indicates that the data exhibits strong persistence, while standard deviations indicate that the coefficients have posterior distributions concentrated away from zero. Bayes factors are the main tool of Bayesian model selection. With improper priors, however, Bayes factors are not well defined. As a consequence, I compute standard frequentist goodness of fit statistics using the posterior means. These statistics can be directly compared with the maximum likelihood estimates in Table 1. The results for the AFP(1) indicate that the inclusion of an autoregressive term yields important improvements in pseudo-R^2. The inclusion of the second estimated factor also yields improvements in pseudo-R^2 and the AIC and BIC both confirm the contribution. The inclusion of the third and fourth factors does not yield an important improvement and is rejected by the information criteria. In sum, both the AFP(1) and AFP(2) models exhibit a good fit with high values of pseudo-R^2. The inclusion of the third and fourth factors does not appear to improve the fit and, as a consequence, the AFP(4) is discarded.

[6] δ_1 and δ_2 exhibit 95% highest posterior density intervals that do not include zero. δ_3 and δ_4, however, exhibit 95% highest posterior density intervals that include zero. As discussed in Koop (2003), this is evidence against the inclusion of the third and fourth factors. The intervals are not reported, but are available upon request.

[TABLE 2 ABOUT HERE]

The latent variable in the probit model can be interpreted as an index of business cycle conditions that can be used to assess the strength of an expansion or the depth of a recession. Figure 3 plots the negative posterior mean (at each t) latent variable from the AFP(1) and AFP(2) models for the full sample. By construction, the indices take negative values during recessions and perfectly match NBER dates. While the index from the AFP(2) exhibits more variation than the one from the AFP(1), both indices suggest that the 1973-75, 1980, and 1981-82 recessions were relatively deeper than the other recessions in the sample. However, since business conditions are not observable, it is not possible to formally evaluate these results. Aruoba et al. (2009) propose an index of business conditions –the Aruoba-Diebold-Scotti (ADS) busines conditions index– that is designed to track real business conditions at high frequency and is regularly updated by the Federal Reserve Bank of Philadelphia. Figure 4 plots the standardized negative posterior mean latent variable for the two models considered here together with the ADS index.[7] The figure shows that, for both models, the latent variable almost perfectly overlaps with the ADS index, and the latter appears to be fluctuating around the business conditions indices from the probit models. Consequently, the ADS index shows an important degree of correlation with the indices extracted from the models proposed here: 0.75 with the AFP(1) index and 0.78 with the AFP(2) index. Since the econometric methods and, more importantly, the information set used in this paper differ substantially from those used in Aruoba et al. (2009), this

[7] The ADS index series was taken from the left-most column in a spreadsheet of vintages of ADS business conditions indices available for download at the Federal Reserve Bank of Philadelphia's website. Since the index is constructed at a daily frequency, the observation corresponding to the first day of a given month was assigned to the previous month (e.g., the value of the ADS index on 1/1/2008 was assigned to December 2007). The spreadsheet is available at:
http://www.philadelphiafed.org/research-and-data/real-time-center/business-conditions-index/

result is quite remarkable.

[FIGURE 3 ABOUT HERE]

[FIGURE 4 ABOUT HERE]

Figure 5 plots the posterior mean probabilities of recession estimated from the factor-augmented autoregressive probit models using (11). In both cases, recession probabilities consistently rise during NBER recession dates and almost perfectly identify recession periods with very high probability values. Comparing the estimated recession probabilities from the AFP(1) (Figure 5, top) with the ones from the traditional probit model with the first factor as predictor (Figure 2) can be useful to understand the effect of including an autoregressive term in the regression. The autoregressive probit model generates recession probabilities that are smooth and eliminates, for the most part, false alarms. Consistent with the results in Table 2, including more factors in the regression improves the fit by generating recession probabilities that are marginally closer to 1 during recessions and closer to 0 during expansions.

[FIGURE 5 ABOUT HERE]

4.2 Out-of-Sample Results

To provide a more realistic assessment of the AFP(k) models, I evaluate their predictive performance in a pseudo out-of-sample forecasting exercise. This exercise requires that we make some assumptions about what was known at each time t. First, the factors are estimated recursively, each period using data only up to time t. This requires assuming that all series in the panel were available up to time t at time t.[8] Second, since recent

[8] This is not likely since some series are only available after a few weeks or months. Giannone et al. (2008), however, develop a formal framework for forecasting in real time using a large number of series released with different lags that could be used here.

NBER dates are not known, I assume that the forecaster does not know whether the true state of the economy has changed over the last 12 months. This implies that, at time t, each model is estimated assuming that $y_{t-i} = y_{t-12}$ for $i = 0, 1, ..., 11$. As a consequence, the sign condition on y_t^* is not imposed on these last 12 observations when generating draws of the latent variable in the Gibbs sampler. Since recession probabilities for time t at time t ($\hat{p}_{t,t}$) are generated without making use of y_t, these are in fact out-of-sample recession probabilities.

I use the hold-out sample period 1988:1 – 2007:12 to generate the out-of-sample forecasts $\hat{p}_{t,t}$. The models are estimated recursively, expanding the estimation window by one observation each month. At each time t, the Gibbs sampler was run 12,000 iterations and, after discarding the first 2,000 draws to allow the sampler to converge, posterior mean probabilities of recession are computed using a thinning factor of 10. Figure 6 presents out-of-sample recession probabilities from the AFP(1) and AFP(2) models. In both cases, recession probabilities consistently rise during subsequently declared NBER recession dates. The fact that recession probabilities do not always cross the commonly used cut-off value of 0.5 has to be evaluated with caution. Since the sample proportion of recession periods is only 0.1431, the sample considered here is unbalanced and, as a consequence, a predicted probability of recession as large as 0.5 may be unlikely (see, e.g., Greene, 2008). Cramer (1999) shows that, with unbalanced samples, it is normally found that estimated probabilities are very low for the outcome with the smaller share.

[Figure 6 ABOUT HERE]

A formal evaluation of the out-of-sample results requires the selection of a loss function that reflects the preferences of the forecaster. In the case of recession indicators, the loss is greater in the case of missed signals and, hence, an asymmetric loss function may be more appropriate. The cost-weighted misclassification loss function assumes that the two types of misclassifications (false positives and false negatives) involve differing costs while assuming that the sum of costs add to 1 (see, e.g., Buja et al., 2005). The loss function is given by

$$ML = \frac{1}{N} \sum_{t=1}^{N} \big((1-q) y_t (1-\hat{y}_{t,t}) + q(1-y_t)\hat{y}_{t,t}\big), \tag{12}$$

where N is the number of out-of-sample forecasts, $\hat{y}_{t,t}$ is the predicted class, q is the cost of a false positive, and $(1-q)$ is the cost of a false negative. The loss is 0 if the predicted classification is perfect and takes positive values otherwise. In order to compute the loss I need to select a classification rule that translates the out-of-sample recession probabilities into class predictions. A simple rule is given by

$$\hat{y}_{t,t} = \begin{cases} 1 & \text{if } \hat{p}_{t,t} \geq c \\ 0 & \text{otherwise} \end{cases}, \tag{13}$$

for some c to be chosen by the forecaster, with $0 < c < 1$. While the usual choice is $c = 0.5$ (see, e.g., Chauvet and Potter, 2008), an alternative cut-off considered in the literature consists on setting c equal to the sample proportion of recession periods $\bar{p}$, yielding the rule $c = \bar{p}$ (see, e.g., Birchenhall et al., 1999). Cramer (1999) analyzes the use of classification rules for class prediction and concludes that, for unbalanced samples, the sample proportion is a better choice for the cut-off rather than 0.5.

The horizontal dashed lines on Figure 6 represent the two possible decision rules, $c = 0.1431$ and $c = 0.5$. The results show that setting the cut-off at 0.5 delivers a conservative rule where the AFP(1) (partially) identifies the 1990 recession and fails to recognize the 2001 recession, while the AFP(2) fails to recognize both recessions. Setting the cut-off at $\bar{p}$ substantially reduces the number of false negatives, but implies too many false positives for the AFP(1). The AFP(2), however, matches subsequently declared NBER dates relatively well with very few false positives. To compute the misclassification loss (12) we need to specify the relative cost of false positives and false negatives. Since the cost is greater in the case of a missed signal, I specify $q = 1/3$ and $(1 - q) = 2/3$; i.e., the cost of a false negative is twice the cost of a false positive. The choice of q, although arbitrary, is not important for the results. Table 3 presents the misclassification loss for $c = 0.5$ and $c = \bar{p}$. The results for the hold-out sample confirm that recession probabilities from the AFP(2) and the decision rule $c = \bar{p}$ generate the sequence of class predictions that better approximate subsequently declared NBER recession dates. Since the predictive performance of the models is different for expansion and recession periods, Table 3 also provides the loss for these subperiods. The decision rule $c = 0.5$ misses most recession signals and, as a consequence, exhibits a very large loss during recessions. The rule $c = \bar{p}$ on the other hand, shows a much lower loss during recessions at the cost of a larger loss during expansions.[9]

[TABLE 3 ABOUT HERE]

4.3 Calibration of an Optimal Classification Rule

In the previous section, the models' performance was evaluated using two arbitrary rules, $c = 0.5$ and $c = \bar{p}$. Elliott and Lieli (2005), however, argue that the cut-off should not be arbitrary but rather chosen to reflect the preferences of the forecaster. In this section, I consider the problem of how to determine the cut-off from a sequence of observed forecasts. This is a calibration exercise that implies finding the value of c that minimizes a pre-defined loss function (see, e.g., Gneiting and Raftery, 2007). The

[9] Since the hold-out sample includes only two recessions, the results must be taken with caution.

optimal cut-off c^* can be estimated by minimizing the cost-weighted misclassification loss function (12):

$$c^* = \arg\min_c \frac{1}{N}\sum_{t=1}^{N}\big((1-q)y_t(1-\hat{y}_{t,t}(c)) + q(1-y_t)\hat{y}_{t,t}(c)\big), \tag{14}$$

with $\hat{y}_{t,t}(c)$ given by (13). Ideally, i.e. with a large sample that includes many recession and expansion periods, the hold-out sample would be divided in two. The first subsample would be used to calibrate the decision rule, i.e. find c^*, using the sequence of out-of-sample forecasts. The second subsample would then be used to formally evaluate the out-of-sample performance of the model and the decision rule jointly. Unfortunately, such an exercise is not feasible since the hold-out sample only includes two recessions and, as a consequence, I only perform the calibration exercise.

To minimize (12) we need to specify the relative cost of false positives and false negatives. Again, the choice of q is rather arbitrary but, as long as $q < 0.5$, it is not important for the results. I specify $q = 1/3$ and $(1-q) = 2/3$. For the hold-out sample period 1988:1 – 2007:12, $c^* = 0.26$ for the AFP(1) and $c^* = 0.21$ for the AFP(2). Table 3 presents the misclassification loss for $c = c^*$. Results are presented for the hold-out sample, for expansion periods, and for recession periods. The AFP(2) exhibits a smaller overall loss, at the cost of some missed signals and a larger loss for recession periods. The AFP(1) shows a smaller loss during recessions, at the cost of a larger loss during expansions. The horizontal solid lines on Figure 6 represent the decision rule $c = c^*$ for both models. In sum, with the classification rule $c = c^*$, the AFP(1) detects both peaks accurately while troughs are called with some lag. On the other hand, the AFP(2) is more conservative, calling troughs earlier than the AFP(1) and generating fewer false positives at the cost of some missed signals. As a consequence, the selection of the best performing model depends on the preferences of the forecaster.

5 Conclusion

In this paper, I propose a framework for the assessment of current business conditions using the information contained in a large number of macroeconomic time series. The main driving force of the results is a factor that loads heavily on measures of real output and employment. In-sample results show that the model fits the data as well as other more complicated models. The latent variable in the probit model can be interpreted as an index of business conditions. The indices from the factor-augmented autoregressive probit models are highly correlated with the index extracted by Aruoba et al. (2009). Out-of-sample predicted recession probabilities consistently rise during subsequently declared NBER recession dates and the models exhibit good performance as real time dating algorithms.

The model I consider can be extended in a number of ways. First, it can be extended to allow for non-linear dynamics. Expansions and recessions may be probabilistically different regimes, and a Markov switching dynamic probit model (as in Dueker, 1999) may be more adequate. Another extension involves using a two-regime self-exciting threshold autoregressive probit model which allows for asymmetric effects over the business cycle. Second, the model can be used to evaluate the predictive power of the macro factors for future (rather than current) business conditions. In particular, it is of interest to evaluate which factors are relevant at different horizons. These extensions are topics for future research.

References

Albert, J.H., and Chib, S. (1993): "Bayesian Analysis of Binary and Polychotomous Response Data", *Journal of the American Statistical Association*, 88, 669-679.

Aruoba, S.B., Diebold, F.X., and Scotti, C. (2009): "Real-Time Measurement of Business Conditions", forthcoming in *Journal of Business and Economic Statistics.*

Bai, J., and Ng, S. (2002): "Determining the Number of Factors in Approximate Factor Models", *Econometrica*, 70, 191-221.

Bai, J., and Ng, S. (2006): "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions", *Econometrica*, 74, 1133-1150.

Bai, J., and Ng, S. (2008): "Extremum Estimation when the Predictors are Estimated from Large Panels", unpublished.

Birchenhall, C.R., Jessen, H., Osborn, D.R., Simpson, P. (1999): "Predicting U.S. Business-Cycle Regimes", *Journal of Business and Economic Statistics*, 17(3), 313-323.

Buja, A., Stuetzle, W., and Shen, Y. (2005): "Loss Functions for Binary Class Probability Estimation and Classification: Structure and Applications", unpublished.

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 Hamilton, J. D. (2006): "Dating Business Cycle Turning Points", in *Non-linear Time Series Analysis of Business Cycles*, ed. by D. van Dijk, C. Milas, and P. Rothman, Contributions to Economic Analysis series, 276, 1-54. Elsevier.

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 Recessions: Evidence from the Yield Curve in the Presence of Structural Breaks", *Economic Letters*, 77(2), 245-253.

Chauvet, M., and Potter, S. (2005): "Forecasting Recessions Using the Yield Curve", *Journal of Forecasting*, 24(2), 77-103.

Chauvet, M., and Potter, S. (2008): "Business Cycle Monitoring with Structural Changes", forthcoming in *International Journal of Forecasting.*

Cramer, J.S. (1999): "Predictive Performance of the Binary Logit Model in Unbalanced Samples", *Journal of the Royal Statistical Society*, Series D (The Statistician), 48(1), 85-94.

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.

Dueker, M.J. (1999): "Conditional Heteroscedasticity in Qualitative Response Models of Time Series: A Gibbs Sampling Approach to the Bank Prime Rate", *Journal of Business and Economic Statistics*, 17(4), 466-472.

Dueker, M.J. (2005): "Dynamic Forecasts of Qualitative Variables: A Qual VAR Model of U.S. Recessions", *Journal of Business and Economic Statistics*, 23(1), 96-104.

Eichengreen, B., Watson, M.W., and Grossman, R.S. (1985): "Bank Rate Policy Under Interwar Gold Standard", *Economic Journal*, 95, 725-745.

Elliott, G., and Lieli, R.P. (2005): "Predicting Binary Outcomes", unpublished.

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.

Giannone, D., Reichlin, L., and Small, D. (2008): "Nowcasting: The Real-Time Informational Content of Macroeconomic Data", *Journal of Monetary Economics*, 55(4), 665-676.

Gneiting, T., and Raftery, A.E. (2007): "Strictly Proper Scoring Rules, Prediction, and Estimation", *Journal of the American Statistical Association*, 102, 359-378.

Greene, W. (2008): "Discrete Choice Modeling", *The Handbook of Econometrics: Vol. 2, Applied Econometrics*, Part 4.2., ed. T. Mills and K. Patterson, Palgrave, London.

Hamilton, J. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica*, 57, 357-384.

Katayama, M. (2009): "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.

Koop, G. (2003): *Bayesian Econometrics*, Wiley.

Krugman, P. (2008): "Responding to Recession", *The New York Times*.

Leamer, E. (2008): "What's a Recession, Anyway?", NBER Working Paper No. 14221.

Ludvigson, S.C., and Ng, S. (2008): "Macro Factors in Bond Risk Premia", forthcoming in *The Review of Financial Studies*.

Ludvigson, S.C., and Ng, S. (2008): “A Factor Analysis of Bond Risk Premia”, unpublished .

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. (2005): “Implications of Dynamic Factor Models for VAR Analysis”, unpublished.

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. (2008): “Forecasting in Dynamic Factor Models Subject to Structural Instability”, in *The Methodology and Practice of Econometrics - A Festschrift in Honour of Professor David F. Hendry*, ed. by J. Castle and N. Shephard. Oxford: Oxford University Press.

Table 1: Single-Factor Probit Models for y_t

Regressor	$\hat{f}_{1t}$	$\hat{f}_{2t}$	$\hat{f}_{3t}$	$\hat{f}_{4t}$	$\hat{f}_{5t}$	$\hat{f}_{6t}$	$\hat{f}_{7t}$	$\hat{f}_{8t}$
α	-1.831	-1.075	-1.064	-1.075	-1.063	-1.064	-1.059	-1.062
	(0.138)	(0.066)	(0.065)	(0.066)	(0.065)	(0.065)	(0.065)	(0.065)
δ	-1.655	0.187	-0.097	-0.177	0.095	0.091	0.022	-0.084
	(0.159)	(0.058)	(0.063)	(0.060)	(0.060)	(0.062)	(0.055)	(0.060)
R^2	0.555	0.022	0.005	0.019	0.005	0.005	0.000	0.004
$ln\hat{L}$	-105.425	-231.854	-235.849	-232.644	-235.769	-235.953	-236.954	-236.059
LR	263.221	10.362	2.372	8.781	2.532	2.164	0.161	1.953
p-value	0.000	0.001	0.123	0.003	0.112	0.141	0.688	0.162

Note: Probit models with $y_t^* = \alpha + \delta\hat{f}_{it} + \epsilon_t$ for $i = 1, ..., 8$ and $t = 1, ..., T$ are estimated by maximum likelihood. Top panel reports parameter estimates and standard errors (in parentheses). $R^2 = 1 - ln\hat{L}/lnL_0$ is McFadden's pseudo-R^2, where $ln\hat{L}$ is the value of the log likelihood function evaluated at the estimated parameter values and lnL_0 is the log likelihood computed only with a constant term. $LR = -2[ln\hat{L} - lnL_0]$ is the likelihood ratio test statistic and p-value is the associated probability value.

Table 2: Factor-Augmented Autoregressive Probit Models for y_t

	AFP(1)		AFP(2)		AFP(4)	
	Mean	SD	Mean	SD	Mean	SD
α	-0.672	0.150	-1.007	0.254	-1.083	0.261
δ_1	-0.539	0.145	-0.847	0.251	-0.987	0.269
δ_2			0.503	0.097	0.520	0.101
δ_3					-0.049	0.150
δ_4					0.213	0.129
θ	0.741	0.067	0.719	0.078	0.710	0.069
R^2	0.838		0.890		0.891	
$ln\hat{L}$	-38.439		-26.127		-25.949	
AIC	0.145		0.105		0.112	
BIC	0.167		0.136		0.157	

Note: Top panel reports posterior mean parameters and standard deviations for the three models. $R^2 = 1 - ln\hat{L}/lnL_0$ is McFadden's pseudo-R^2, where $ln\hat{L}$ is the value of the log likelihood function evaluated at the posterior means and lnL_0 is the log likelihood computed only with a constant term. AIC $= -2(ln\hat{L})/T + k2/T$ and BIC $= -2(ln\hat{L})/T + k(lnT)/T$ are the traditional information criteria.

Table 3: Out-of-Sample Misclassification Loss

	AFP(1)			AFP(2)		
	$c = 0.5$	$c = \bar{p}$	$c = c^*$	$c = 0.5$	$c = \bar{p}$	$c = c^*$
Hold-out sample	0.040	0.061	0.022	0.053	0.019	0.015
Expansions	0.002	0.066	0.021	0.000	0.015	0.005
Recessions	0.491	0.000	0.035	0.667	0.070	0.140

Note: Results for the hold-out sample period 1988:1 − 2007:12. $\bar{p} = 0.1431$ is the full sample proportion of recession periods. c^* is the optimal cut-off. $c^* = 0.26$ for the AFP(1) and $c^* = 0.21$ for the AFP(2).

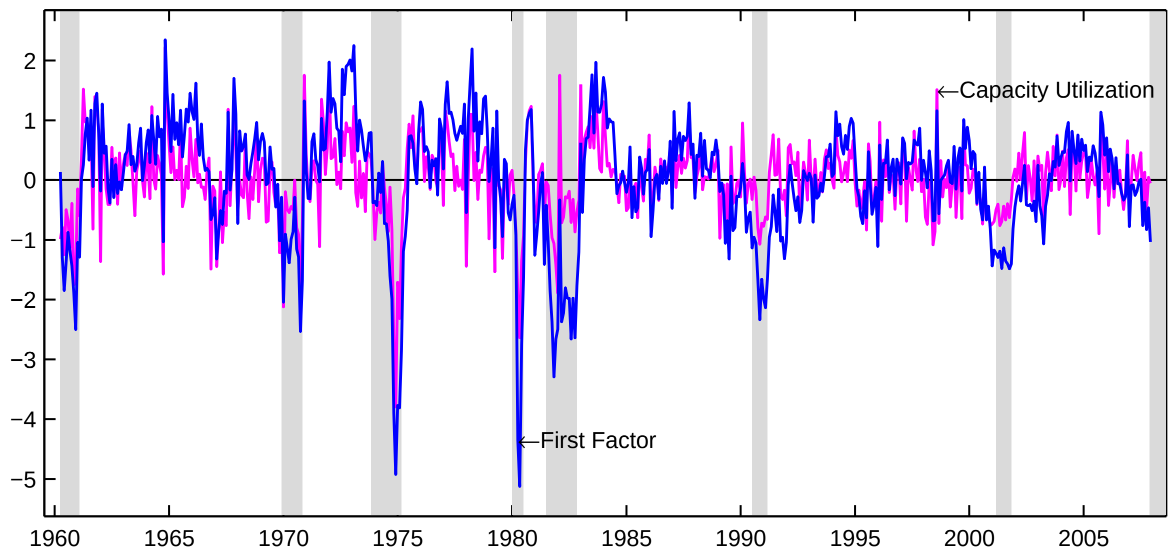

Figure 1: First Factor and Capacity Utilization. "First Factor" denotes the first estimated factor ($\hat{f}_{1t}$). Standardized units are reported. Shaded areas denote NBER recession months.

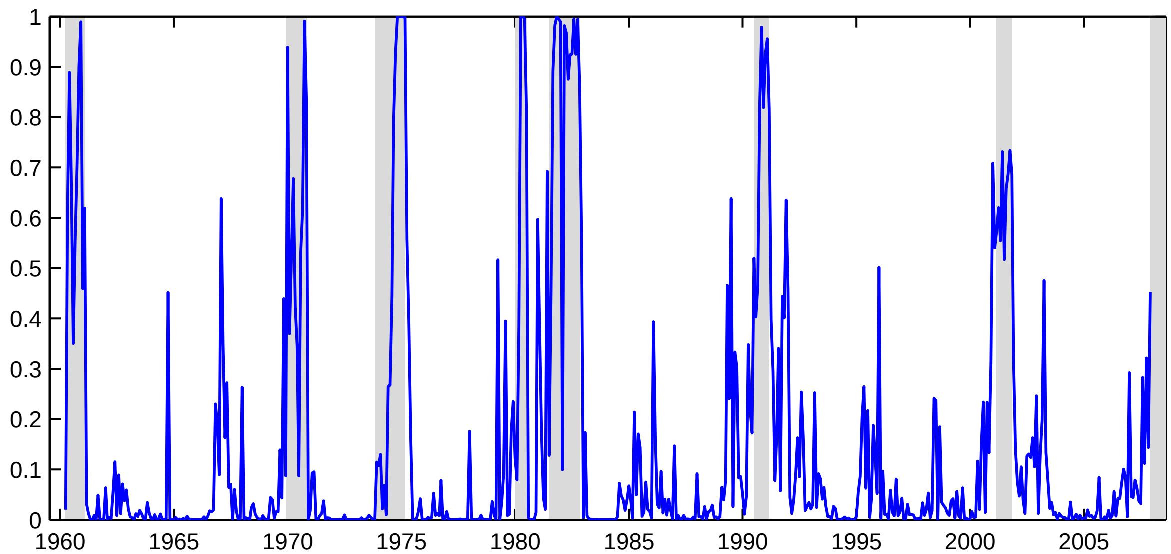

Figure 2: In-sample probabilities of recession from the single-factor probit model using the first estimated factor ($\hat{f}_{1t}$). Shaded areas denote NBER recession months.

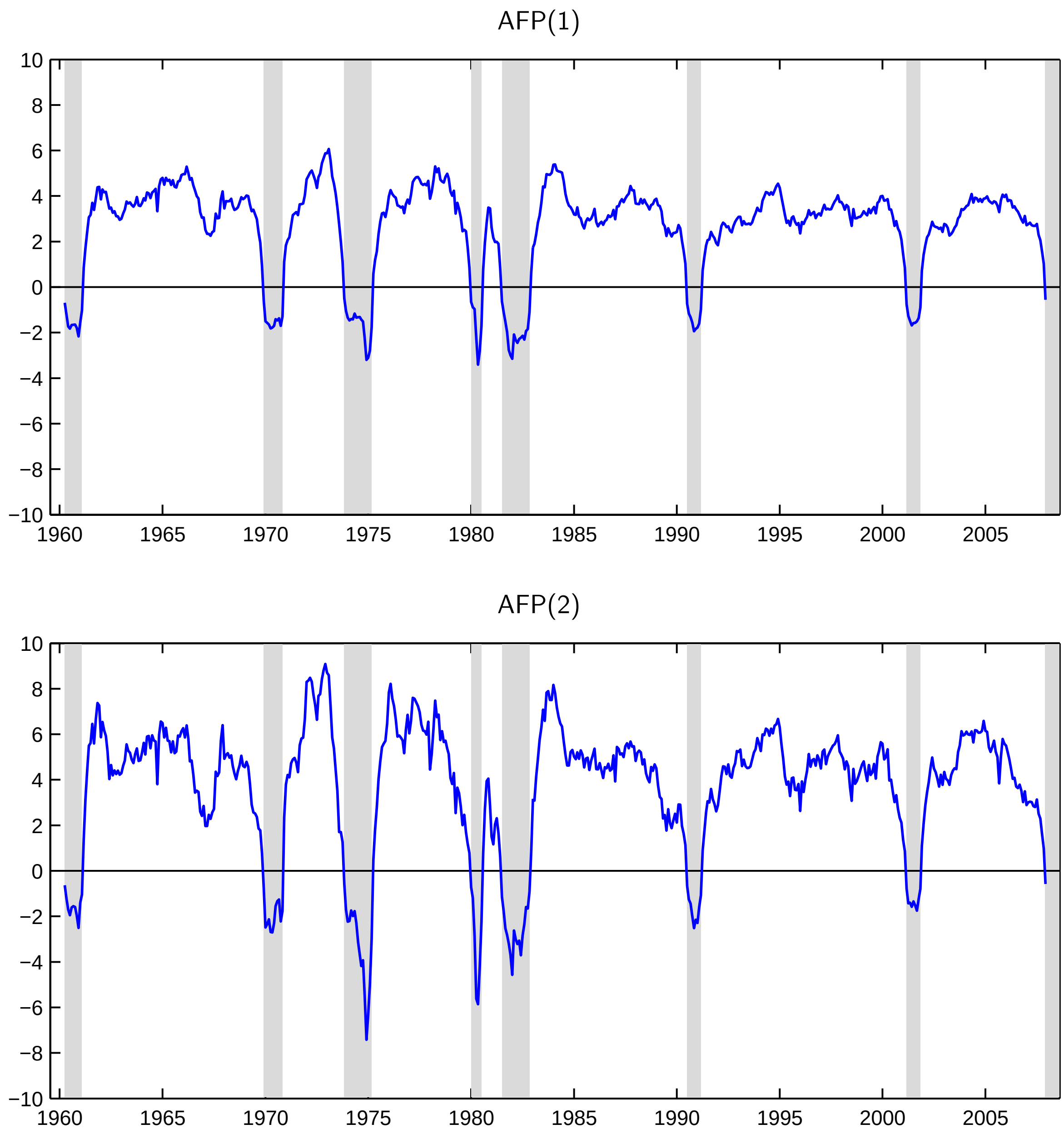

Figure 3: In-sample negative posterior mean latent variable from the 1- and 2-factor autoregressive probit models. Shaded areas denote NBER recession months.

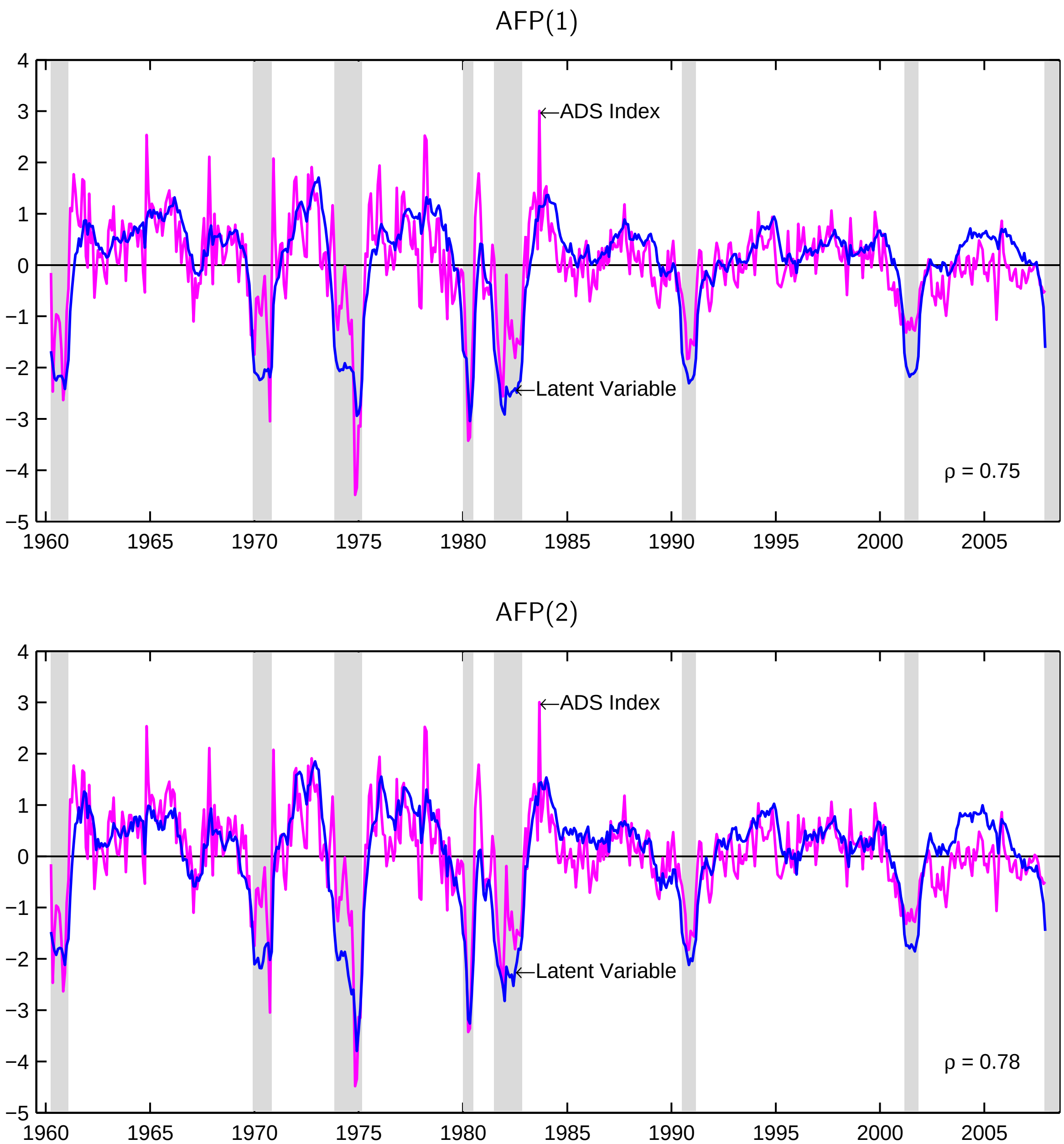

Figure 4: Standardized negative posterior mean latent variable from the 1- and 2-factor autoregressive probit models for the full sample and the Aruoba-Diebold-Scotti (ADS) business conditions index. ρ is the correlation between $-y_t^*$ and the ADS index. Shaded areas denote NBER recession months.

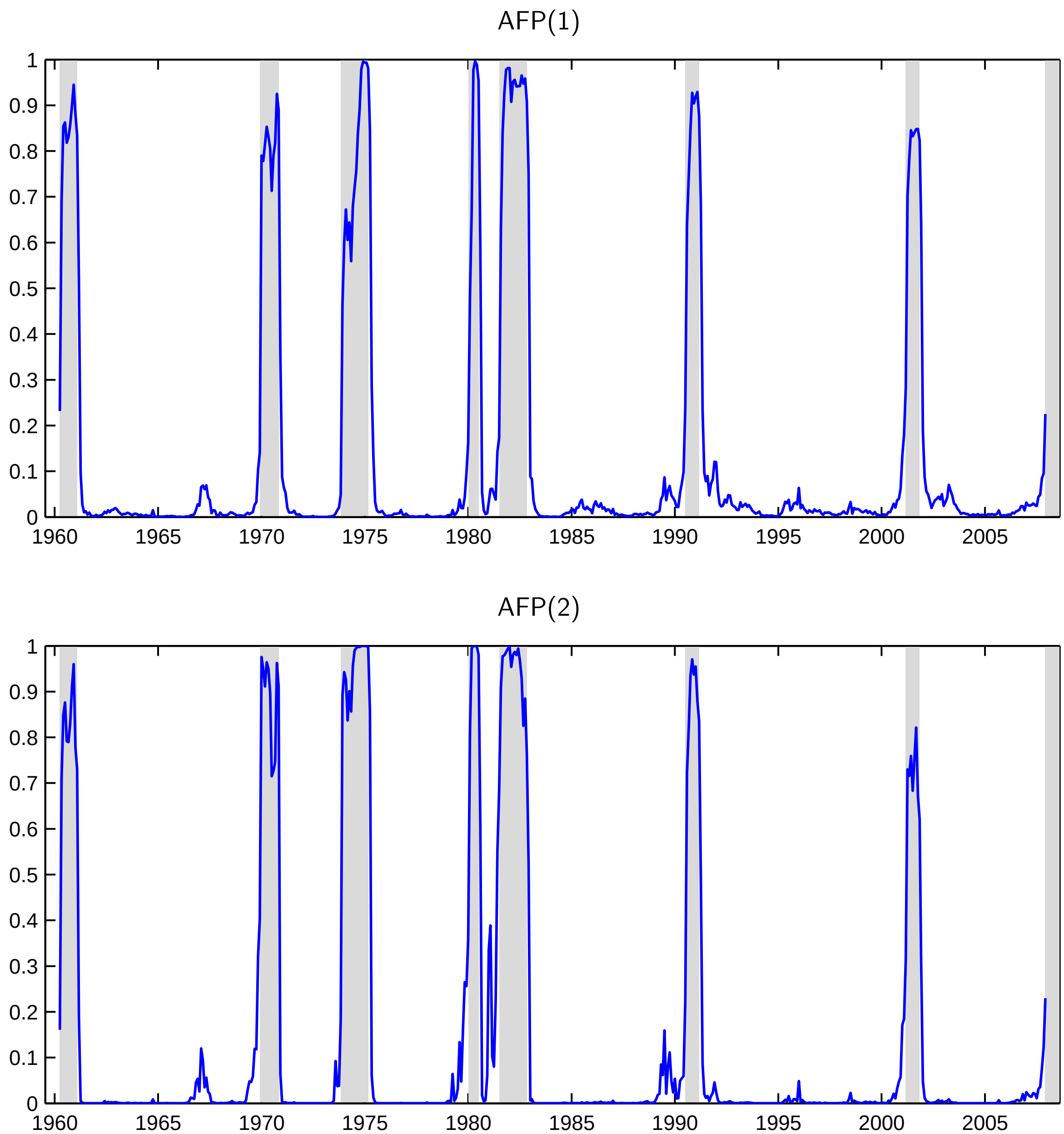

Figure 5: In-sample posterior mean probabilities of recession ($\hat{p}_t$) from the 1- and 2-factor autoregressive probit models. Shaded areas denote NBER recession months.

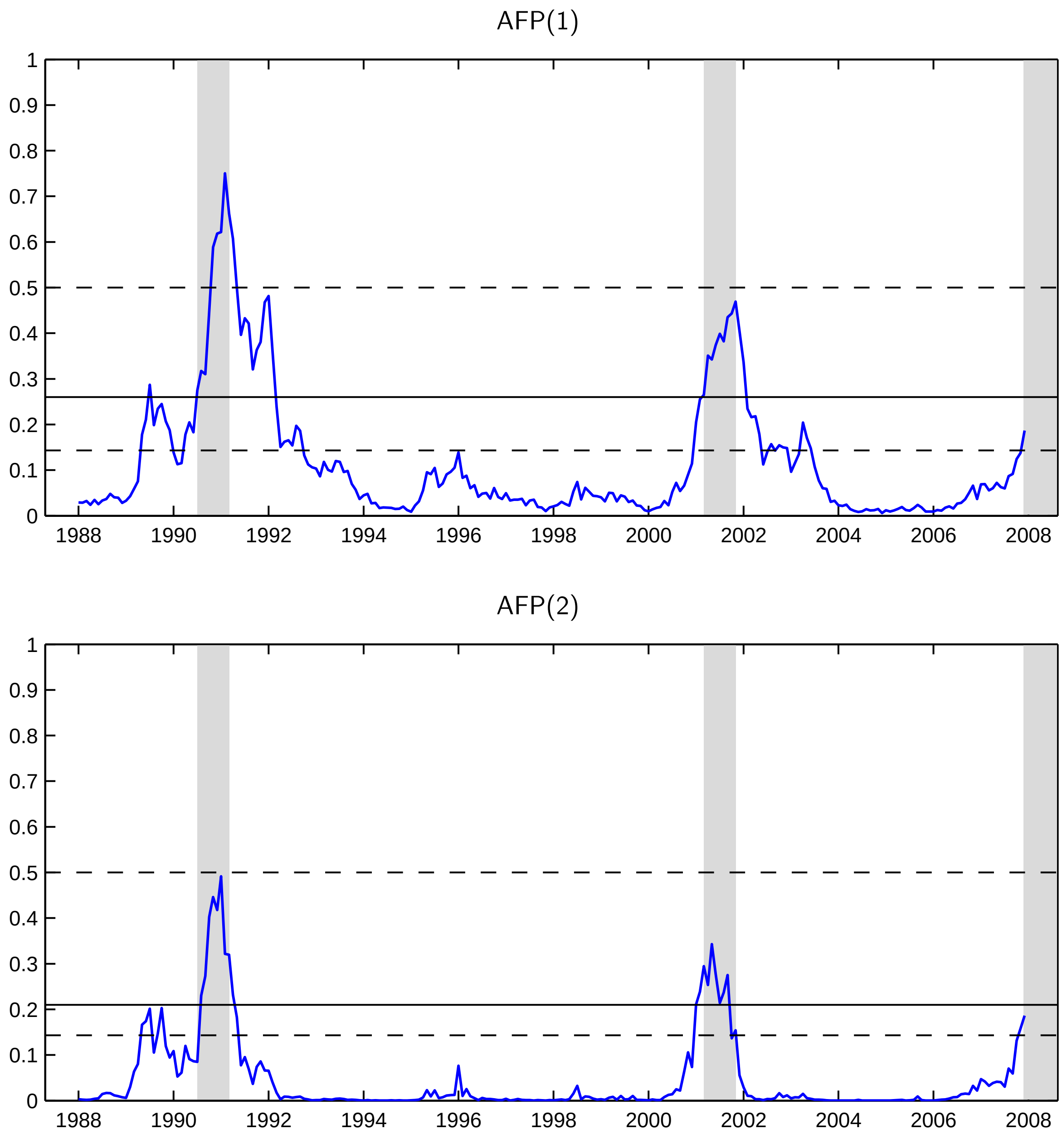

Figure 6: Out-of-sample posterior mean probabilities of recession ($\hat{p}_{t,t}$) from the 1- and 2-factor autoregressive probit models. The horizontal dashed lines represent the decision rules $c = 0.1431$ and $c = 0.5$. The horizontal solid lines represent the decision rule $c = c^*$. Shaded areas denote NBER recession months.