UNIVERSITY OF ALBERTA
FACULTY OF ARTS
Department of Economics

# Working Paper No. 2011-05

# Dating U.S. Business Cycles with Macro Factors

**Sebastian Fossati**
**University of Alberta**

February 2012

# Dating U.S. Business Cycles with Macro Factors

Sebastian Fossati*
University of Alberta

This version: February 27, 2012

**Abstract**

Latent common factors estimated from a set of macroeconomic time series are used to generate recession probabilities for the U.S. economy. The focus is on current (rather than future) business conditions. Two macro factors are considered: (1) a dynamic factor estimated by maximum likelihood from a set of 4 monthly time series; (2) the first of 8 static factors estimated by principal components using a panel of 102 monthly time series. Based on these macro factors, recession probabilities are generated using a traditional probit model, an autoregressive probit model, and a Markov-switching model. In- and out-of-sample results show that the macro factors have important predictive power for subsequently declared NBER recession dates. Two main results emerge from this work. First, a simple Markov-switching model based on the small data dynamic factor generates the sequence of class predictions that better approximates NBER recession dates. Second, a standard probit model based on the big data macro factor exhibits the most accurate performance during recessions at the cost of some false positives during expansions.

*Keywords*: Business Cycle, Forecasting, Factors, Probit Model, Bayesian Methods.
*JEL Codes*: E32, E37, C01, 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 thank Eric Zivot, Drew Creal, Dante Amengual, German Cubas, Ana Galvao, Jeremy Piger, and Byron Tsang for helpful comments.

# 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 (or shortly after) is not an easy task since business conditions are not observable and NBER announcements are issued long after the fact.[2]

In this context, a common strategy among those interested in modeling business cycles consists in treating business conditions as an unobserved variable and using a factor model to extract a common component (a latent factor) from a panel of macroeconomic indicators. Initial contributions to this literature favored a small data approach where a dynamic factor is estimated by maximum likelihood from a few time series; see, e.g., Stock and Watson (1991), Chauvet (1998), Kim and Nelson (1998), and the more recent contribution of Aruoba et al. (2009). Recently, however, the big data approach where static factors are estimated by principal components from a large number of time series has been found useful in many forecasting exercises; see, e.g.,

[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 has taken between 6 to 20 months to announce peaks and troughs.

Stock and Watson (2002a,b, 2006), Giannone et al. (2008), and Ludvigson and Ng (2009a,b).

In this paper, I use latent common factors estimated from small and large data sets of macroeconomic indicators (macro factors) to predict NBER recession dates. Traditionally, two alternative approaches have been used to generate recession probabilities: (1) binary class models; (2) Markov-switching models. Therefore, the first approach considered in this paper consists in using probit regressions for NBER recession dates. 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 similar to Chauvet and Potter (2010) who consider the performance of four monthly coincident macroeconomic indicators as predictors of current (rather than future) business conditions. However, instead of relying on a small number of observed variables, in this paper I consider the information contained in small data and big data macro factors. The second approach considered in this paper consists in generating recession probabilities using Markov-switching models; see, e.g., Hamilton (1989), Chauvet (1998), Chauvet and Hamilton (2006), and Chauvet and Piger (2008). In this case, I use a framework similar to Diebold and Rudebusch (1996) and Camacho et al. (2011) where recession probabilities are generated directly using simple Markov-switching models for the macro factors.[3]

The main results of this paper can be summarized as follows. First, in-sample results show that while standard probit models generate recession probabilities that consistently rise during NBER recession dates, these probabilities are relatively volatile and the models exhibit several false positives during expansions. Autoregressive probit

[3] A nice review of the different approaches to dating business cycle turning points is provided by Hamilton (2011).

models and simple Markov-switching models, on the other hand, generate in-sample recession probabilities that are smooth and high during NBER recession dates and exhibit few false positives. Second, a pseudo out-of-sample forecasting exercise shows that end-of-sample predicted recession probabilities from the standard probit models also rise during subsequently declared NBER recession dates and the probit model based on the big data static factor fits NBER recession dates better than the probit model based on the small data dynamic factor. The autoregressive probit models, on the other hand, exhibit a poor out-of-sample performance, generating probabilities that are low during NBER recession dates and yielding significantly delayed recession calls. As a result, autoregressive probit models appear to offer no out-of-sample improvements over standard probit models. Finally, the Markov-switching models generate out-of-sample recession probabilities that are smooth and high during NBER recession dates and exhibit few false positives. Overall, a simple Markov-switching model based on the small data macro factor generates the sequence of class predictions that better approximates NBER recession dates. In addition, a standard probit model based on the big data macro factor exhibits the most accurate performance during recessions at the cost of some false positives during expansions.

This paper is organized as follows. Section 2 discusses the estimation of macro factors from small and large data sets. Section 3 presents the econometric models used to generate recession probabilities. Section 4 presents in- and out-of-sample results and discusses the classification of recession probabilities into class predictions. Section 5 concludes.

# 2 Estimation of Macro Factors

In this section, I consider the estimation of macro factors from small and large data sets. First, I discuss the use of maximum likelihood to estimate a dynamic factor from four macroeconomic indicators commonly used in the literature. Subsequently, I discuss the use of principal components to estimate static common factors from a large number of macroeconomic indicators.

## 2.1 A Small Data Macro Factor

Consider the case where we observe a $T \times N$ panel of macroeconomic data $x_t$ where $N$ is small, typically $N = 4$. Assume $x_{it}$, $i = 1, \ldots, N$, $t = 1, \ldots, T$, has a factor structure of the form

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

where $g_t$ is an unobserved common factor, $\lambda_i$ is the factor loading, and $e_{it}$ is the idiosyncratic error. The dynamics of the common factor are driven by $\phi(L)g_t = \eta_t$ with $\eta_t \sim i.i.d.\, N(0,1)$, while the dynamics of the idiosyncratic errors are driven by $\psi_i(L)e_{it} = \nu_{it}$ with $\nu_{it} \sim i.i.d.\, N(0,\sigma_i^2)$ for $i = 1, \ldots, N$. As in Stock and Watson (1991), identification is achieved by assuming that all shocks are independent.

A set of monthly economic indicators previously used in Stock and Watson (1991), Diebold and Rudebusch (1996), Chauvet (1998), Chauvet and Piger (2008), and Camacho et al. (2011), among others, includes industrial production, real manufacturing sales, real personal income less transfer payments, and employment.[4] In this paper, the common factor is estimated for the period 1960:1 – 2010:12 with the data in $x_t$

[4] A data set of these variables was generously provided by Jeremy Piger.

transformed to ensure stationarity, standardized prior to estimation, and with all autoregressive processes including two lags.[5] Figure 1 presents the estimated dynamic factor along with the (standardized) index of capacity utilization. The series are similar, with major troughs corresponding closely to NBER recession dates (shaded areas).

[ FIGURE 1 ABOUT HERE ]

To obtain an initial assessment of the power of the estimated dynamic factor, $\hat{g}_t$, as predictor of U.S. recessions, I estimate a standard probit model for the sample 1961:1 – 2010:12. 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

$$y_t^* = \alpha + \delta \hat{g}_t + \epsilon_t, \tag{2}$$

where $\alpha$ and $\delta$ are regression coefficients and $\epsilon_t \,|\, \hat{g}_t \sim i.i.d.\, N(0,1)$.[6] We do not observe $y_t^*$ but rather $y_t$, which represents the observable recession indicator according to the following rule

$$y_t = \begin{cases} 1 & \text{if } y_t^* \geqslant 0 \\ 0 & \text{if } y_t^* < 0 \end{cases}, \tag{3}$$

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

$$p_t = P(y_t = 1 \,|\, \hat{g}_t) = P(y_t^* \geqslant 0 \,|\, \hat{g}_t) = \Phi(\alpha + \delta \hat{g}_t), \tag{4}$$

where $\Phi(\cdot)$ is the distribution function of the standard normal. The first column in Table 1 reports parameter estimates, McFadden's pseudo-$R^2$, the value of the log

[5] The model is written in state-space form and estimated via maximum likelihood following Kim and Nelson (1999).

[6] Note that since $y_t^*$ is not observable, if $\epsilon_t \,|\, \hat{g}_t \sim i.i.d.\, N(0, \sigma^2)$ is assumed, the regression coefficients $\alpha$, $\delta$, and $\sigma$ are not separately identified. As a result, it is standard to normalize $\sigma$ to 1.

likelihood, and the likelihood ratio (LR) test statistic for the hypothesis that $\delta = 0$ with its associated probability value. The estimated dynamic factor exhibits important (in-sample) predictive power for $y_t$ with a pseudo-$R^2$ of 0.47. Figure 2 presents recession probabilities estimated from the standard probit model. 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.

[ TABLE 1 ABOUT HERE ]

[ FIGURE 2 ABOUT HERE ]

## 2.2 Big Data Macro Factors

In this section, instead of relying on a small number of indicators, 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 (2009a,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 $T$. Assume $x_{it}$, $i = 1, \ldots, N$, $t = 1, \ldots, T$, has a factor structure of the form

$$x_{it} = \lambda_i' f_t + e_{it}, \tag{5}$$

where $f_t$ is a $r \times 1$ vector of static common factors, $\lambda_i$ is a $r \times 1$ vector of factor loadings, and $e_{it}$ is the idiosyncratic error. Stock and Watson (2002a) show that, when $N, T \to \infty$, $f_t$ can be consistently estimated by principal components analysis. The number of latent common factors, $r$, to be estimated by principal components can be determined using model selection criteria as in Bai and Ng (2002).

Common factors are estimated from a balanced panel of 102 monthly U.S. macroeconomic time series spanning the period 1960:1 – 2010:12. The data set is similar to the one used in Stock and Watson (2002b, 2006) and Ludvigson and Ng (2009a,b). 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.[7]

As in Ludvigson and Ng (2009a,b), eight static common factors are estimated by principal components. 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.[8] 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 standard probit models where

$$y_t^* = \alpha + \delta \hat{f}_{it} + \epsilon_t, \tag{6}$$

for $i = 1, \ldots, 8$ and the sample 1961:1 – 2010:12. Note that the normalization imposed for identification purposes implies that the estimated factors are mutually orthogonal. 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. Although several factors appear to be significant (p-value $< 0.1$), the estimated first factor not only explains most of the variation in the

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

[8] "Total variation" is the sum of the variances of the variables in the panel $x$.

panel $x_t$, but also has the largest (in-sample) predictive power for $y_t$ with a pseudo-$R^2$ of 0.54. The other significant factors exhibit very low values of pseudo-$R^2$.

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 (2009a, Figure 1) show that the first factor loads heavily on real variables such as employment, production, capacity utilization, and manufacturing orders. Figure 3 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 (2009a,b), the first factor appears to be an index of real economic activity. Figure 4 presents recession probabilities estimated from the standard probit model using the first estimated factor, $\hat{f}_{1t}$, as predictor. As in the case of the dynamic factor, recession probabilities consistently rise during NBER recession dates and the model signals recessions with high probability values. In addition, while the model also shows probabilities that are relatively volatile, the first static factor exhibits fewer false positives during expansions.

[ FIGURE 3 ABOUT HERE ]

[ FIGURE 4 ABOUT HERE ]

# 3 Econometric Framework

I consider three alternative models for generating recession probabilities from macro factors. The first model is a factor-augmented probit regression for $y_t$ given by

$$y_t^* = \alpha + \delta \hat{h}_t + \epsilon_t, \tag{7}$$

where $\hat{h}_t$ is a one-dimensional estimated common factor (either $\hat{g}_t$ or $\hat{f}_{1t}$). Classical or bayesian implementations of this model are straightforward.[9]

The standard probit model defined above ignores that $y_t^*$ is a time series variable which exhibits strong persistence. To account for this autocorrelation in the latent variable, the second probit model introduces a lag of $y_t^*$ in (7). As a consequence, the second model is an autoregressive probit given by

$$y_t^* = \alpha + \delta \hat{h}_t + \theta y_{t-1}^* + \epsilon_t, \tag{8}$$

where $|\theta| < 1$. This model is similar to the models considered in Dueker (1999) and Chauvet and Potter (2005, 2010). As in the case of the standard probit, the conditional probability of recession is given by

$$p_t = P(y_t = 1 \,|\, \hat{h}_t, y_{t-1}^*) = P(y_t^* \geqslant 0 \,|\, \hat{h}_t, y_{t-1}^*) = \Phi\big(\alpha + \delta \hat{h}_t + \theta y_{t-1}^*\big). \tag{9}$$

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, on the other hand, can greatly simplify the problem. The approach I take consists on 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 for the autoregressive probit model is similar to that of Dueker (1999) and Chauvet and Potter (2005, 2010) and is discussed in appendix A.

Finally, instead of using binary class models to generate recession probabilities, we can use Markov-switching models to generate probabilities directly from the macro

[9] In this paper, the bayesian implementation of traditional probit models follows Koop (2003) and is not discussed here.

factors as in Diebold and Rudebusch (1996) and Camacho et al. (2011). Therefore, for the third model it is assumed that the factor $\hat{h}_t$ switches between expansion and contraction regimes following a mean plus noise specification given by

$$\hat{h}_t = \mu_{s_t} + \epsilon_t, \tag{10}$$

where $s_t$ is defined such that $s_t = 0$ during expansions and $s_t = 1$ during recessions, and $\epsilon_t \sim i.i.d.\, N(0, \sigma^2_\epsilon)$. In addition, $s_t$ is an unobserved two-state first-order Markov process with transition probabilities given by

$$p(s_t = j \,|\, s_{t-1} = i) = p_{ij}, \tag{11}$$

where $i, j = 0, 1$. The regime-switching mean plus noise model can be estimated by maximum likelihood following Kim and Nelson (1999).[10]

A formal evaluation of the predicted probabilities of recession requires the selection of a loss function. A commonly used function is the Quadratic Probability Score (QPS, hereafter) which is given by

$$QPS = \frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{p}_t)^2, \tag{12}$$

where $N$ is the number of forecasts and $\hat{p}_t$ is the fitted probability of recession for a given model (see, for example, Katayama, 2009). Smaller QPS values indicate more accurate predictions.

# 4 Empirical Results

In this section, the two macro factors are combined with each of the three models discussed in the previous section to generate recession probabilities. The first group

[10]Autoregressive Markov-switching models of order one (as in Diebold and Rudebusch, 1996) and up to four lags were also considered but did not yield better results than the mean plus noise model considered in Camacho et al. (2011).

of models use the small data macro factor $\hat{g}_t$. The three models considered are: (1) a standard probit model with (the dynamic factor) $\hat{g}_t$ as predictor (DF-SP, hereafter); (2) an autoregressive probit model with $\hat{g}_t$ as predictor (DF-AP, hereafter); (3) a Markov-switching mean plus noise model for $\hat{g}_t$ (DF-MS, hereafter). The second group of models use the big data macro factor $\hat{f}_{1t}$. The three models considered are: (4) a standard probit model with (the static factor) $\hat{f}_{1t}$ as predictor (SF-SP, hereafter); (5) an autoregressive probit model with $\hat{f}_{1t}$ as predictor (SF-AP, hereafter); (6) a Markov-switching mean plus noise model for $\hat{f}_{1t}$ (SF-MS, hereafter).

For the in-sample results, both macro factors are estimated using the full sample of time series information, and 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 from a pseudo real-time exercise. In this case, the factors are estimated recursively, each period using data only up to time $t$.[11] 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 twelve months such that $y_{t-i} = y_{t-12}$ for $i = 0, 1, ..., 11$. Further details are given below. Finally, sections 4.3 and 4.4 consider the classification of recession probabilities into binary class predictions using a sequence of observed forecasts.

## 4.1 In-Sample Results

To estimate the probit models, the Gibbs sampler was run with 25,000 iterations. After discarding the first 5,000 draws (burn-in period), posterior means are computed

[11]It should be noted that the out-of-sample exercise uses revised data instead of data as it was available at the time. Unfortunately, a real-time data set for all the series considered in this paper is not available. However, as discussed in Berge and Jorda (2011), although data revisions can be considerable for individual series, changes are usually smaller for indices.

using a thinning factor of 20, i.e. computed from every 20th draw. As a result, the subsequent analysis is based on the means of these 1000 draws. Table 2 (top panel) reports posterior means and standard deviations for the four probit regressions. All models show parameter posterior distributions that are concentrated away from zero and the macro factors are clearly important. Even though Bayes factors are the main tool of Bayesian model selection, with improper priors, Bayes factors are not well defined. As a consequence, I compute standard frequentist goodness-of-fit statistics using the posterior means (Table 2, bottom panel). These statistics can be directly compared with those computed from maximum likelihood estimates. In the case of the standard probit, the first static factor $\hat{f}_{1t}$ exhibits more predictive power for $y_t$ than the dynamic factor $\hat{g}_t$. Specifically, model SF-SP shows a pseudo-$R^2$ of 0.54 while model DF-SP yields 0.47. In addition, the inclusion of the autoregressive term yields large improvements in pseudo-$R^2$ in both cases. Finally, the inclusion of additional static factors does not yield important improvements (results not reported).

[ TABLE 2 ABOUT HERE ]

Figure 5 plots the estimated probabilities of recession for the four probit models (top and center rows). The standard probit models produce probabilities that consistently rise during NBER recession dates and signal recessions with high probability. While both standard probit models show probabilities that are relatively volatile during recessions and exhibit some false positives during expansions, model SF-SP has a lower QPS and, hence, fits NBER recession dates better than model DF-SP. Comparing the estimated recession probabilities from these models with the ones from the autoregressive probits can be useful to understand the effect of including the autoregressive term in the regression. The autoregressive probit models generate recession probabilities that are smooth and eliminate, for the most part, false alarms. As a consequence, the inclusion of an autoregressive term in the probit regressions generates large improvements in QPS. In contrast, the inclusion of additional static factors in the probit models improves the overall fit by generating recession probabilities that are only marginally closer to 1 during recessions and to 0 during expansions (results not reported).

[ FIGURE 5 ABOUT HERE ]

Maximum likelihood estimates of the two Markov-switching models are reported in Table 3 (top panel). Both models exhibit very persistent transition probabilities (close to 1) and the parameter estimates for the dynamic factor Markov-switching model are very close to those reported in Camacho et al. (2011). Estimated probabilities of recession for the two Markov-switching models are plotted in Figure 5 (bottom row). The Markov-switching models also generate recession probabilities that are smooth and high during NBER recession dates and exhibit very few false positives. Overall, the QPS suggests that model DF-MS fits NBER recession dates better than model SF-MS and the in-sample fit is similar to that of the autoregressive probit models.

[ TABLE 3 ABOUT HERE ]

## 4.2 Out-of-Sample Results

To provide a more realistic assessment of the 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 $t$.[12] Second, since recent 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 twelve months. This implies that, at time $t$, each probit model is estimated assuming that $y_{t-i} = y_{t-12}$ for $i = 0, 1, ..., 11$.[13] Since end-of-sample 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 1988:1 – 2010:12 to generate the end-of-sample forecasts $\hat{p}_{t,t}$. The models are estimated recursively, expanding the estimation window by one observation each month. In the case of the probit models, at each time $t$, the Gibbs sampler was run 6,000 iterations and, after discarding the first 1,000 draws to allow the sampler to converge, results are computed using a thinning factor of 10. Figure 6 presents end-of-sample probabilities of recession for the six models considered. The standard probit models generate recession probabilities that consistently rise during subsequently declared NBER recession dates and exhibit few false positives. Model DF-SP, however, shows lower and more volatile probabilities during recessions and also exhibits more noise during expansions. The autoregressive probit models, on the other hand, exhibit the worst performance, generating probabilities that are smooth but very low during NBER recession dates and yielding significantly delayed recession calls. As a result, the autoregressive probit models fail to identify the 1990 and 2001 recessions with high probabilities and only identify the 2007 recession with an important lag. Finally, the Markov-switching models generate recession probabilities that are smooth

[12] 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.

[13] As a result, the probit models are estimated without imposing the sign condition on $y_t^*$ on these last twelve observations when generating draws of the latent variable in the Gibbs sampler (see appendix A for more details).

and high during NBER recession dates and exhibit few false positives. Table 4 (top panel) reports out-of-sample QPS results for the six models. Model DF-MS generates the series of end-of-sample recession probabilities that better fit subsequently declared NBER recession dates. Since the predictive performance of the models is different for expansion and recession periods, Table 4 also reports the QPS for these sub-periods. Compared to model DF-MS, model SF-SP exhibits a lower loss during recessions at the cost of a larger loss during expansions due to some false positives.

[ FIGURE 6 ABOUT HERE ]

Figure 7 presents the full paths of recession probabilities from which the end-of-sample probabilities are obtained (tentacle plot). In the case of the standard probit models and the Markov-switching models, the probability paths do not exhibit much variation as more data is incorporated and, as a result, in- and end-of-sample estimated probabilities are similar. The results for the autoregressive probit models, on the other hand, are quite different. In this case, the paths exhibit important changes as additional observations are added to the sample and this issue is particularly evident during recession dates. In this case, the probability paths appear to be highly persistent, explaining the delayed recession calls noticed above.

[ FIGURE 7 ABOUT HERE ]

## 4.3 Binary Class Predictions

A formal evaluation of the end-of-sample probabilities as predictors of NBER recession dates requires the selection of a classification rule and 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 (ML, hereafter) 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 ML function is given by

$$ML = \frac{1}{N} \sum_{t=1}^{N} \big( (1-q) y_t (1-\hat{y}_{t,t}) + q(1-y_t)\hat{y}_{t,t} \big), \tag{13}$$

where $N$ is the number of end-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. To compute (13) 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. Finally, we need to select a classification rule that translates the end-of-sample recession probabilities into class predictions. A simple rule is given by

$$\hat{y}_{t,t} = \begin{cases} 1 & \text{if } \hat{p}_{t,t} \geq c \\ 0 & \text{otherwise} \end{cases}, \tag{14}$$

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, 2010), an alternative cut-off considered in the literature consists on setting $c$ equal to the sample proportion of recession periods $\bar{p}$ (see, e.g., Birchenhall et al., 1999), yielding the rule $c = 0.13$. 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 these two decision rules: $c = 0.13$ and $c = 0.5$. The results show that setting the cut-off at 0.5 delivers a conservative rule where the probit models struggle to recognize some of the recessions. Setting the cut-off at $\bar{p}$ substantially reduces the number of false negatives, but implies too many false positives for most probit models. In comparison, the two Markov-switching models recognize subsequently declared NBER recession dates relatively well with few false positives. Table 4 presents the ML for $c = 0.5$ and $c = \bar{p}$. With these classification rules, model DF-MS generates the sequence of class predictions that better approximate subsequently declared NBER recession dates.

## 4.4 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).[14] The optimal cut-off $c^*$ can be estimated by minimizing the cost-weighted misclassification loss (13) such that

$$c^* = \arg\min_c \frac{1}{N} \sum_{t=1}^{N} \big((1-q)y_t(1-\hat{y}_{t,t}(c)) + q(1-y_t)\hat{y}_{t,t}(c)\big), \tag{15}$$

with $\hat{y}_{t,t}(c)$ given by (14). 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 $(1-q) = 2/3$. 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

[14]Berge and Jorda (2011) consider a similar approach to determine optimal thresholds to classify economic activity into recessions and expansions directly from indices of business conditions.

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 three recessions and, as a consequence, I only perform the calibration exercise.

Table 4 (bottom panel) presents the optimal cut-off and ML for each model for the hold-out sample 1988:1 – 2010:12. For the four probit models, the optimal cut-off falls between the two thresholds considered in the previous section. On the other hand, for both Markov-switching models the optimal cut-off is larger than 0.5. The horizontal solid lines on Figure 6 represent the optimal threshold for each model. Looking at the ML for $c = c^*$ on Table 4, we observe that model DF-MS exhibits a smaller overall loss at the cost of some missed signals and a larger loss for recession periods (specifically, a delayed call of the 2007 recession). On the other hand, model SF-SP shows a smaller loss during recessions, at the cost of a larger loss during expansions due to some false positives. In sum, with the classification rule $c = c^*$, model DF-MS is more conservative (fewer false positives at the cost of some missed signals) while model SF-SP detects most peaks and throughs accurately at the cost of some false positives. Therefore, the selection of the best performing model depends on the preferences of the forecaster.

# 5 Conclusion

This paper provides an assessment of the predictive power of macro factors for U.S. recession dates using both binary class models and Markov-switching models. Instead of relying on a small number of observed variables, these models are built around latent common factors estimated from small and large data sets of macroeconomic indicators. The results show that macro factors have important predictive power for

NBER recession dates and can be used to assess current business conditions. Overall, a simple Markov-switching model based on the small data macro factor generates the sequence of (out-of-sample) class predictions that better approximates NBER recession dates. In addition, a standard probit model based on the big data macro factor exhibits the most accurate performance during recessions at the cost of some false positives during expansions.

# A Autoregressive Probit Model Estimation

The regression equation for the factor-augmented autoregressive probit model is

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

where $\gamma = (\alpha, \delta)'$ and $z_t = (1, \hat{h}_t)'$, and the likelihood function for the model is

$$L\left(y|z, \gamma, \theta, y_0\right) = \prod_{t=1}^{T} \left[\Phi(\gamma' z_t + \theta y_{t-1}^*)\right]^{y_t} \left[1 - \Phi(\gamma' z_t + \theta y_{t-1}^*)\right]^{1-y_t}. \tag{A.2}$$

The implementation of the Gibbs sampler is similar to that of Dueker (1999) and Chauvet and Potter (2005, 2010). 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 $(\gamma', \theta)$ and the observed data; (ii) generate draws of $\gamma'$ conditional on $(y_t^*, \theta)$ and the observed data; (iii) generate draws of $\theta$ conditional on $(y_t^*, \gamma')$ and the observed data. Prior and posterior distributions are discussed next.

## A.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 (A.1) and substituting backwards for lagged $y^*$'s on the right side, the following results can be derived:

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

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

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

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

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 $i$th cycle of the Gibbs sampler. As in Chauvet and Potter (2005, 2010), 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.

## A.2 Prior and Posterior for $\gamma$

Following Albert and Chib (1993) and Dueker (1999), I use a flat non-informative prior for $\gamma$. Initial values for $\gamma$ 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 $\gamma$ are generated from the multivariate normal distribution $\gamma \,|\, y^*, \theta, y \sim N\left(\hat{\gamma}, (z'z)^{-1}\right)$ where $\hat{\gamma} = (z'z)^{-1} z' W^\gamma$.

## A.3 Prior and Posterior for $\theta$

Similarly, I use a flat non-informative prior for the autoregressive parameter $\theta$. The initial value of $\theta$ 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 $\theta$ are generated from $\theta \,|\, y^*, \gamma, y \sim N\big(\hat{\theta}, (W^{y\prime} W^y)^{-1}\big)$ 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.

### A.4 Recession Probabilities

Conditional recession probabilities are generated at each draw of the Gibbs sampler such that

$$p_t^{(i)} = \Phi\left(\gamma^{(i)\prime} z_t + \theta^{(i)} y_{t-1}^{*(i)}\right), \tag{A.3}$$

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

$$\hat{p}_t = \frac{1}{I} \sum_{i=1}^{I} p_t^{(i)}, \tag{A.4}$$

where $I$ denotes the total number of draws.

## B Data Appendix

This appendix lists the 102 time series included in the balanced panel. The table lists the short name of each series, the transformation applied, and a brief data description. All series are from FRED – St. Louis Fed –, unless the source is listed as ECON (Economagic), GFD (Global Financial Data), or AC (author's calculation). The transformation codes are: 1 = no transformation; 2 = first difference; 3 = second difference; 4 = logarithm; 5 = first difference of logarithms; 6 = second difference of logarithms.

| | Short Name | Trans. | Description |
|---|---|---|---|
| 1 | PI | 5 | Personal Income (Bil. Chain 2005 $) |
| 2 | PILT | 5 | Personal Income Less Transfer Payments (AC) |
| 3 | CONS | 5 | Real Consumption (Bil. Chain 2005 $) |
| 4 | IP | 5 | Industrial Production Index - Total Index |
| 5 | IPP | 5 | Industrial Production Index - Products, Total (ECON) |
| 6 | IPF | 5 | Industrial Production Index - Final Products |
| 7 | IPCG | 5 | Industrial Production Index - Consumer Goods |
| 8 | IPDCG | 5 | Industrial Production Index - Durable Consumer Goods |
| 9 | IPNDCG | 5 | Industrial Production Index - Nondurable Consumer Goods |
| 10 | IPBE | 5 | Industrial Production Index - Business Equipment |
| 11 | IPM | 5 | Industrial Production Index - Materials |
| 12 | IPDM | 5 | Industrial Production Index - Durable Goods Materials |
| 13 | IPNDM | 5 | Industrial Production Index - Nondurable Goods Materials |
| 14 | IPMAN | 5 | Industrial Production Index - Manufacturing |
| 15 | NAPMPI | 1 | Napm Production Index (%) |
| 16 | MCUMFN | 2 | Capacity Utilization |
| 17 | CLFT | 5 | Civilian Labor Force: Employed, Total (Thous.,sa) |
| 18 | CLFNAI | 5 | Civilian Labor Force: Employed, Nonagric. Industries (Thous.,sa) (ECON) |
| 19 | U: all | 2 | Unemployment Rate: All Workers, 16 Years & Over (%,sa) |
| 20 | U: duration | 2 | Unempl. By Duration: Average Duration In Weeks (sa) |
| 21 | U <5 wks | 5 | Unempl. By Duration: Persons Unempl. Less Than 5 Wks (Thous.,sa) |
| 22 | U 5–14 wks | 5 | Unempl. By Duration: Persons Unempl. 5 To 14 Wks (Thous.,sa) |
| 23 | U 15+ wks | 5 | Unempl. By Duration: Persons Unempl. 15 Wks + (Thous.,sa) |
| 24 | U 15–26 wks | 5 | Unempl. By Duration: Persons Unempl. 15 To 26 Wks (Thous.,sa) |
| 25 | U 27+ wks | 5 | Unempl. By Duration: Persons Unempl. 27 Wks + (Thous,sa) |
| 26 | UI claims | 5 | Average Weekly Initial Claims, Unempl. Insurance |
| 27 | Emp: total | 5 | Employees On Nonfarm Payrolls: Total Private |
| 28 | Emp: gds prod | 5 | Employees On Nonfarm Payrolls - Goods-Producing |
| 29 | Emp: mining | 5 | Employees On Nonfarm Payrolls - Mining |
| 30 | Emp: const | 5 | Employees On Nonfarm Payrolls - Construction |
| 31 | Emp: mfg | 5 | Employees On Nonfarm Payrolls - Manufacturing |
| 32 | Emp: dble gds | 5 | Employees On Nonfarm Payrolls - Durable Goods |
| 33 | Emp: nondbles | 5 | Employees On Nonfarm Payrolls - Nondurable Goods |
| 34 | Emp: serv | 5 | Employees On Nonfarm Payrolls - Service-Providing |
| 35 | Emp: TTU | 5 | Employees On Nonfarm Payrolls - Trade, Transportation, And Utilities |
| 36 | Emp: wholesale | 5 | Employees On Nonfarm Payrolls - Wholesale Trade |

| | Short Name | Trans. | Description |
|---|---|---|---|
| 37 | Emp: retail | 5 | Employees On Nonfarm Payrolls - Retail Trade |
| 38 | Emp: fin | 5 | Employees On Nonfarm Payrolls - Financial Activities |
| 39 | Emp: govt | 5 | Employees On Nonfarm Payrolls - Government |
| 40 | Avg hrs | 2 | Avg Weekly Hrs, Private Nonfarm Payrolls - Goods-Producing |
| 41 | Overtime | 1 | Avg Weekly Hrs, Private Nonfarm Payrolls - Mfg Overtime Hours |
| 42 | Avg hrs mfg | 1 | Average Weekly Hours, Mfg. (Hours) |
| 43 | NAPM emp | 1 | NAPM Employment Index (%) |
| 44 | Starts: nonfarm | 4 | Housing Starts: Total (Thous.,saar) |
| 45 | Starts: NE | 4 | Housing Starts: Northeast (Thous.U.,sa) |
| 46 | Starts: MW | 4 | Housing Starts: Midwest(Thous.U.,sa) |
| 47 | Starts: S | 4 | Housing Starts: South (Thous.U.,sa) |
| 48 | Starts: W | 4 | Housing Starts: West (Thous.U.,sa) |
| 49 | BP: total | 4 | Housing Authorized: Total New Priv Housing Units (Thous.,saar) |
| 50 | NAPM new ords | 1 | NAPM New Orders Index (%) |
| 51 | NAPM vend del | 1 | NAPM Vendor Deliveries Index (%) |
| 52 | NAPM invent | 1 | NAPM Inventories Index (%) |
| 53 | M1 | 6 | Money Stock: M1 (Bil $,sa) |
| 54 | M2 | 6 | Money Stock: M2 (Bil $,sa) |
| 55 | MB | 6 | Monetary Base, Adj For Reserve Requirement Changes (Mil $,sa) |
| 56 | Rsrv tot | 3 | Depository Inst Reserves: Total, Adj For Reserve Req Chgs (Mil $,sa) |
| 57 | Rsrv nonbor | 3 | Depository Inst Reserves: Nonborrowed, Adj Res Req Chgs (Mil $,sa) |
| 58 | Cons credit | 6 | Consumer Credit Outstanding - Nonrevolving |
| 59 | S&P 500 | 5 | S&P's Common Stock Price Index: Composite (1941-43=10) (GFD) |
| 60 | S&P indst | 5 | S&P's Common Stock Price Index: Industrials (1941-43=10) (GFD) |
| 61 | S&P div yield | 5 | S&P's Composite Common Stock: Dividend Yield (% per annum) (GFD) |
| 62 | S&P PE ratio | 5 | S&P's Composite Common Stock: Price-Earnings Ratio (%) (GFD) |
| 63 | Fed Funds | 2 | Interest Rate: Federal Funds (Effective) (% per annum) |
| 64 | Comm paper | 2 | Commercial Paper Rate |
| 65 | 3-m T-bill | 2 | Interest Rate: U.S.Treasury Bills, Sec Mkt, 3-Mo. (% per annum) |
| 66 | 6-m T-bill | 2 | Interest Rate: U.S.Treasury Bills, Sec Mkt, 6-Mo. (% per annum) |
| 67 | 1-y T-bond | 2 | Interest Rate: U.S.Treasury Const Maturities, 1-Yr. (% per annum) |
| 68 | 5-y T-bond | 2 | Interest Rate: U.S.Treasury Const Maturities, 5-Yr. (% per annum) |
| 69 | 10-y T-bond | 2 | Interest Rate: U.S.Treasury Const Maturities, 10-Yr. (% per annum) |
| 70 | AAA bond | 2 | Bond Yield: Moody's AAA Corporate (% per annum) (GFD) |
| 71 | BAA bond | 2 | Bond Yield: Moody's BAA Corporate (% per annum) (GFD) |
| 72 | CP spread | 1 | Comm paper – Fed Funds (AC) |

| | Short Name | Trans. | Description |
|---|---|---|---|
| 73 | 3-m spread | 1 | 3-m T-bill – Fed Funds (AC) |
| 74 | 6-m spread | 1 | 6-m T-bill – Fed Funds (AC) |
| 75 | 1-y spread | 1 | 1-y T-bond – Fed Funds (AC) |
| 76 | 5-y spread | 1 | 5-y T-bond – Fed Funds (AC) |
| 77 | 10-y spread | 1 | 10-y T-bond – Fed Funds (AC) |
| 78 | AAA spread | 1 | AAA bond – Fed Funds (AC) |
| 79 | BAA spread | 1 | BAA bond – Fed Funds (AC) |
| 80 | Ex rate: index | 5 | Exchange Rate Index (Index No.) (GFD) |
| 81 | Ex rate: Swit | 5 | Foreign Exchange Rate: Switzerland (Swiss Franc per U.S.$) |
| 82 | Ex rate: Jap | 5 | Foreign Exchange Rate: Japan (Yen per U.S.$) |
| 83 | Ex rate: U.K. | 5 | Foreign Exchange Rate: United Kingdom (Cents per Pound) |
| 84 | Ex rate: Can | 5 | Foreign Exchange Rate: Canada (Canadian$ per U.S.$) |
| 85 | PPI: fin gds | 6 | Producer Price Index: Finished Goods (82=100,sa) |
| 86 | PPI: cons gds | 6 | Producer Price Index: Finished Consumer Goods (82=100,sa) |
| 87 | PPI: int mat | 6 | Producer Price Index: Intermed. Mat. Supplies & Components (82=100,sa) |
| 88 | PPI: crude mat | 6 | Producer Price Index: Crude Materials (82=100,sa) |
| 89 | Spot Mrk Price | 6 | Spot market price index: all commodities (GFD) |
| 90 | CPI-U: all | 6 | Cpi-U: All Items (82-84=100,sa) |
| 91 | CPI-U: app | 6 | Cpi-U: Apparel & Upkeep (82-84=100,sa) |
| 92 | CPI-U: transp | 6 | Cpi-U: Transportation (82-84=100,sa) |
| 93 | CPI-U: med | 6 | Cpi-U: Medical Care (82-84=100,sa) |
| 94 | CPI-U: comm | 6 | Cpi-U: Commodities (82-84=100,sa) (ECON) |
| 95 | CPI-U: dbles | 6 | Cpi-U: Durables (82-84=100,sa) (ECON) |
| 96 | CPI-U: serv | 6 | Cpi-U: Services (82-84=100,sa) (ECON) |
| 97 | CPI-U: ex food | 6 | Cpi-U: All Items Less Food (82-84=100,sa) |
| 98 | CPI-U: ex shelter | 6 | Cpi-U: All Items Less Shelter (82-84=100,sa) (ECON) |
| 99 | CPI-U: ex med | 6 | Cpi-U: All Items Less Medical Care (82-84=100,sa) (ECON) |
| 100 | PCE defl | 6 | PCE, Implicit Price Deflator: PCE (1987=100) |
| 101 | AHE: const | 6 | Avg Hourly Earnings - Construction |
| 102 | AHE: mfg | 6 | Avg Hourly Earnings - Manufacturing |

# 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", *Journal of Business and Economic Statistics*, 27(4), 417–427.

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

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

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.

Camacho, M., Perez-Quiros, G., and Poncela, P. (2011): "Extracting Nonlinear Signals from Several Economic Indicators", Documento de Trabajo 1202, 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 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. (2010): "Business Cycle Monitoring with Structural Changes", *International Journal of Forecasting*, 26 (4), 777–793.

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.

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.

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.

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.

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

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

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.

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.

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 14221.

Ludvigson, S.C., and Ng, S. (2009): “Macro Factors in Bond Risk Premia”, *The Review of Financial Studies*, 22(12), 5027–5067.

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

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.

Table 1: Single-Factor Probit Models for $y_t$

| Regressor | $\hat{g}_t$ | $\hat{f}_{1t}$ | $\hat{f}_{2t}$ | $\hat{f}_{3t}$ | $\hat{f}_{4t}$ | $\hat{f}_{5t}$ | $\hat{f}_{6t}$ | $\hat{f}_{7t}$ | $\hat{f}_{8t}$ |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | -1.59 | -1.67 | -1.04 | -1.03 | -1.03 | -1.06 | -1.02 | -1.03 | -1.03 |
| | (0.11) | (0.12) | (0.06) | (0.06) | (0.06) | (0.06) | (0.06) | (0.06) | (0.06) |
| $\delta$ | -1.30 | -1.66 | -0.19 | 0.12 | 0.14 | 0.25 | -0.02 | 0.09 | -0.11 |
| | (0.12) | (0.16) | (0.06) | (0.06) | (0.05) | (0.06) | (0.06) | (0.06) | (0.06) |
| $R^2$ | 0.47 | 0.54 | 0.02 | 0.01 | 0.01 | 0.03 | 0.00 | 0.00 | 0.01 |
| $\ln \hat{L}$ | -137.07 | -117.19 | -251.68 | -254.67 | -253.81 | -248.54 | -256.98 | -256.07 | -255.35 |
| $LR$ | 240.00 | 279.76 | 10.78 | 4.80 | 6.52 | 17.06 | 0.18 | 1.99 | 3.44 |
| p-value | 0.00 | 0.00 | 0.00 | 0.03 | 0.01 | 0.00 | 0.67 | 0.16 | 0.06 |

Note: Probit models where $y_t^* = \alpha + \delta \hat{h}_t + \epsilon_t$ and $\hat{h}_t$ is either $\hat{g}_t$ or $\hat{f}_{it}$ for $i = 1, \ldots, 8$ are estimated by maximum likelihood. Top panel reports parameter estimates and standard errors (in parentheses). $R^2 = 1 - \ln \hat{L} / \ln L_0$ 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 $\ln L_0$ is the log likelihood computed only with a constant term. $LR = -2(\ln \hat{L} - \ln L_0)$ is the likelihood ratio test statistic and p-value is the associated probability value.

Table 2: Bayesian Probit Models for $y_t$

| | DF-SP | | DF-AP | | SF-SP | | SF-AP | |
|---|---|---|---|---|---|---|---|---|
| $\alpha$ | -1.60 | (0.11) | -0.59 | (0.13) | -1.68 | (0.12) | -0.58 | (0.14) |
| $\delta$ | -1.31 | (0.11) | -0.53 | (0.13) | -1.67 | (0.15) | -0.47 | (0.15) |
| $\theta$ | | | 0.78 | (0.06) | | | 0.77 | (0.07) |
| $R^2$ | 0.47 | | 0.86 | | 0.54 | | 0.84 | |
| $\ln \hat{L}$ | -137.08 | | -36.97 | | -117.19 | | -40.55 | |
| $QPS$ | 0.06 | | 0.02 | | 0.05 | | 0.02 | |

Note: Top panel reports the parameters' posterior means and standard deviations (in parentheses). $R^2 = 1 - \ln \hat{L} / \ln L_0$ is McFadden's pseudo-$R^2$, where $\ln \hat{L}$ is the value of the log likelihood function evaluated at the posterior means and $\ln L_0$ is the log likelihood computed only with a constant term. $QPS = \frac{1}{N} \sum_{t=1}^{N} (y_t - \hat{p}_t)^2$ where $N$ is the number of forecasts and $\hat{p}_t$ is the fitted probability of recession.

Table 3: MLE Markov-Switching Models

| | DF-MS | | SF-MS | |
|---|---|---|---|---|
| $\mu_0$ | 0.30 | (0.04) | 0.31 | (0.03) |
| $\mu_1$ | -1.67 | (0.13) | -1.69 | (0.10) |
| $\sigma^2$ | 0.71 | (0.05) | 0.49 | (0.03) |
| $p_{00}$ | 0.98 | (0.01) | 0.99 | (0.01) |
| $p_{11}$ | 0.89 | (0.04) | 0.91 | (0.03) |
| $\ln \hat{L}$ | -801.23 | | -686.26 | |
| $QPS$ | 0.02 | | 0.04 | |

Note: Top panel reports parameter estimates and standard errors (in parentheses). $\ln \hat{L}$ is the value of the log likelihood function evaluated at the parameter estimates. $QPS = \frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{p}_t)^2$ where $N$ is the number of forecasts and $\hat{p}_t$ is the fitted probability of recession.

Table 4: Out-of-Sample Loss

| | | DF-SP | DF-AP | DF-MS | SF-SP | SF-AP | SF-MS |
|---|---|---|---|---|---|---|---|
| $QPS$ | Hold-out sample | 0.06 | 0.06 | **0.03** | 0.05 | 0.07 | 0.05 |
| | Recessions | 0.27 | 0.36 | 0.15 | **0.12** | 0.29 | 0.12 |
| | Expansions | 0.03 | 0.02 | **0.01** | 0.03 | 0.03 | 0.04 |
| $ML(c = 0.5)$ | Hold-out sample | 0.05 | 0.07 | **0.02** | 0.03 | 0.06 | 0.03 |
| | Recessions | 0.34 | 0.47 | 0.13 | 0.14 | 0.38 | **0.11** |
| | Expansions | 0.00 | **0.00** | 0.01 | 0.01 | 0.01 | 0.02 |
| $ML(c = \bar{p})$ | Hold-out sample | 0.08 | 0.06 | **0.03** | 0.08 | 0.07 | 0.04 |
| | Recessions | **0.00** | 0.04 | 0.09 | **0.00** | **0.00** | 0.07 |
| | Expansions | 0.09 | 0.06 | **0.02** | 0.09 | 0.09 | 0.03 |
| $ML(c = c^*)$ | $c^*$ | 0.35 | 0.25 | 0.72 | 0.40 | 0.31 | 0.86 |
| | Hold-out sample | 0.04 | 0.04 | **0.02** | 0.02 | 0.04 | 0.02 |
| | Recessions | 0.14 | 0.16 | 0.13 | **0.02** | 0.13 | 0.11 |
| | Expansions | 0.02 | 0.02 | **0.00** | 0.02 | 0.02 | 0.01 |

Note: $QPS = \frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{p}_{t,t})^2$ where $N$ is the number of out-of-sample forecasts and $\hat{p}_{t,t}$ is the end-of-sample probability of recession. $ML = \frac{1}{N}\sum_{t=1}^{N}[(1-q)y_t(1-\hat{y}_{t,t}) + q(1-y_t)\hat{y}_{t,t}]$ where $\hat{y}_{t,t} = 1(\hat{p}_{t,t} \geq c)$, $c$ is some threshold such that $0 < c < 1$, and $q$ is the cost of a false positive. $\bar{p} = 0.13$ for all models.

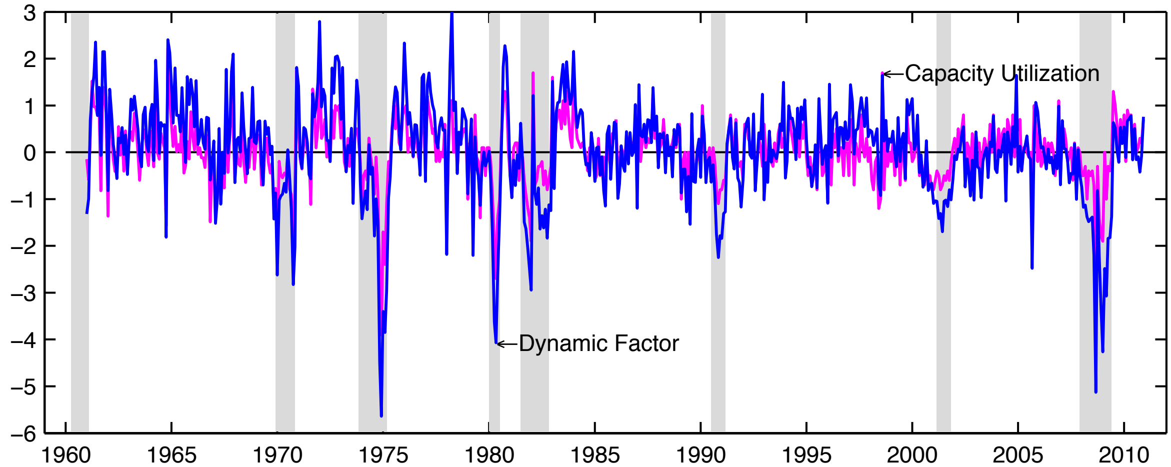


Figure 1: Dynamic factor ($\hat{g}_t$) and capacity utilization. Standardized units are reported. Shaded areas denote NBER recession months.

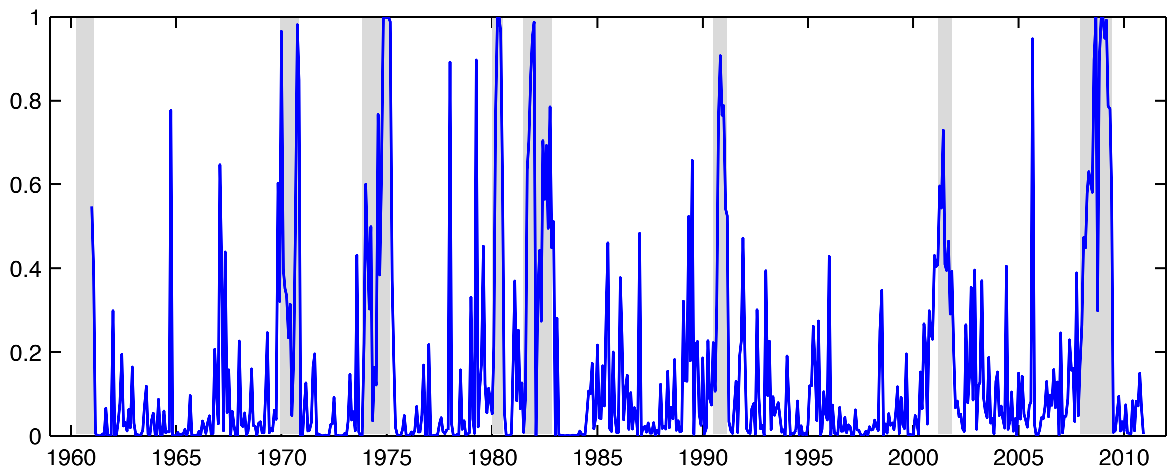

Figure 2: In-sample probabilities of recession from the single-factor probit model using $\hat{g}_t$ as predictor. Shaded areas denote NBER recession months.

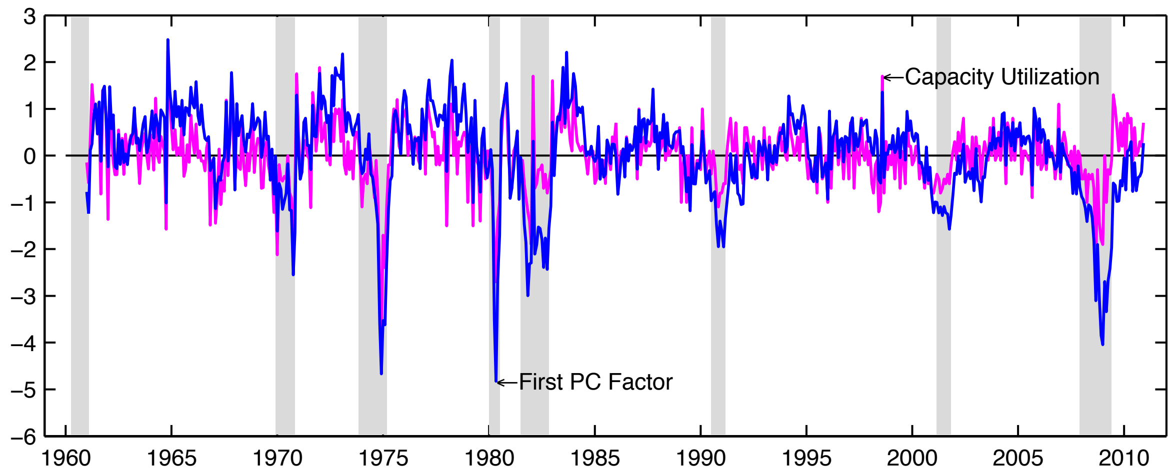


Figure 3: First principal components factor ($\hat{f}_{1t}$) and capacity utilization. Standardized units are reported. Shaded areas denote NBER recession months.

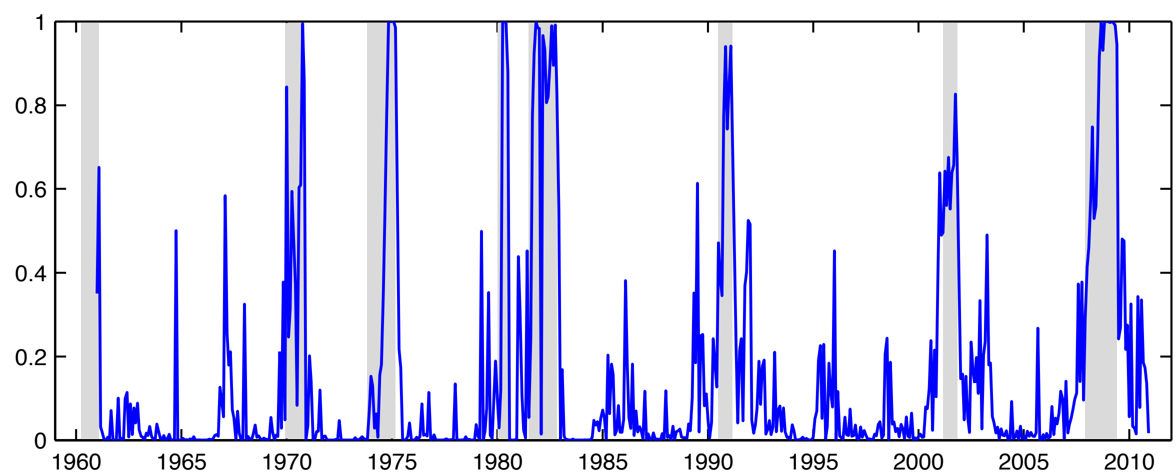

Figure 4: In-sample probabilities of recession from the single-factor probit model using $\hat{f}_{1t}$ as predictor. Shaded areas denote NBER recession months.

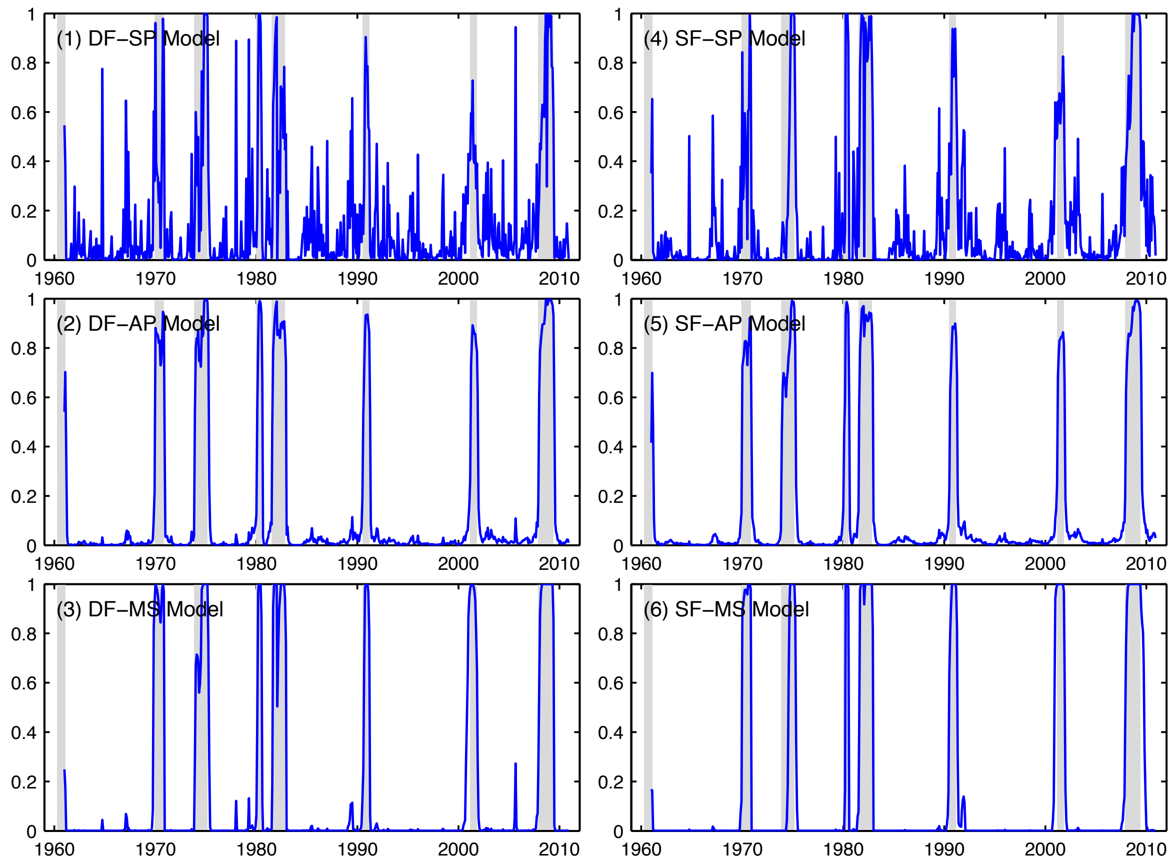


Figure 5: In-sample probabilities of recession ($\hat{p}_t$). Shaded areas denote NBER recession months.

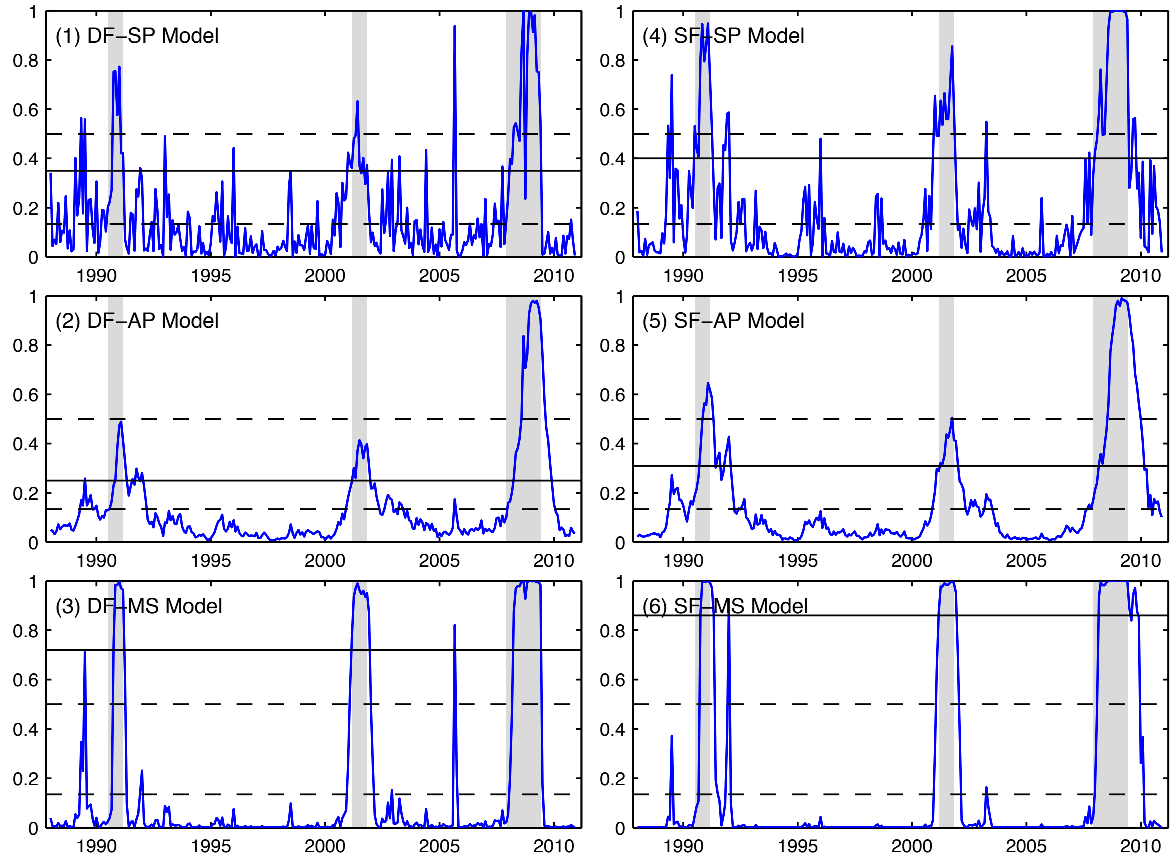


Figure 6: End-of-sample probabilities of recession ($\hat{p}_{t,t}$) for the hold-out sample 1988:1 – 2010:12. Shaded areas denote NBER recession months.

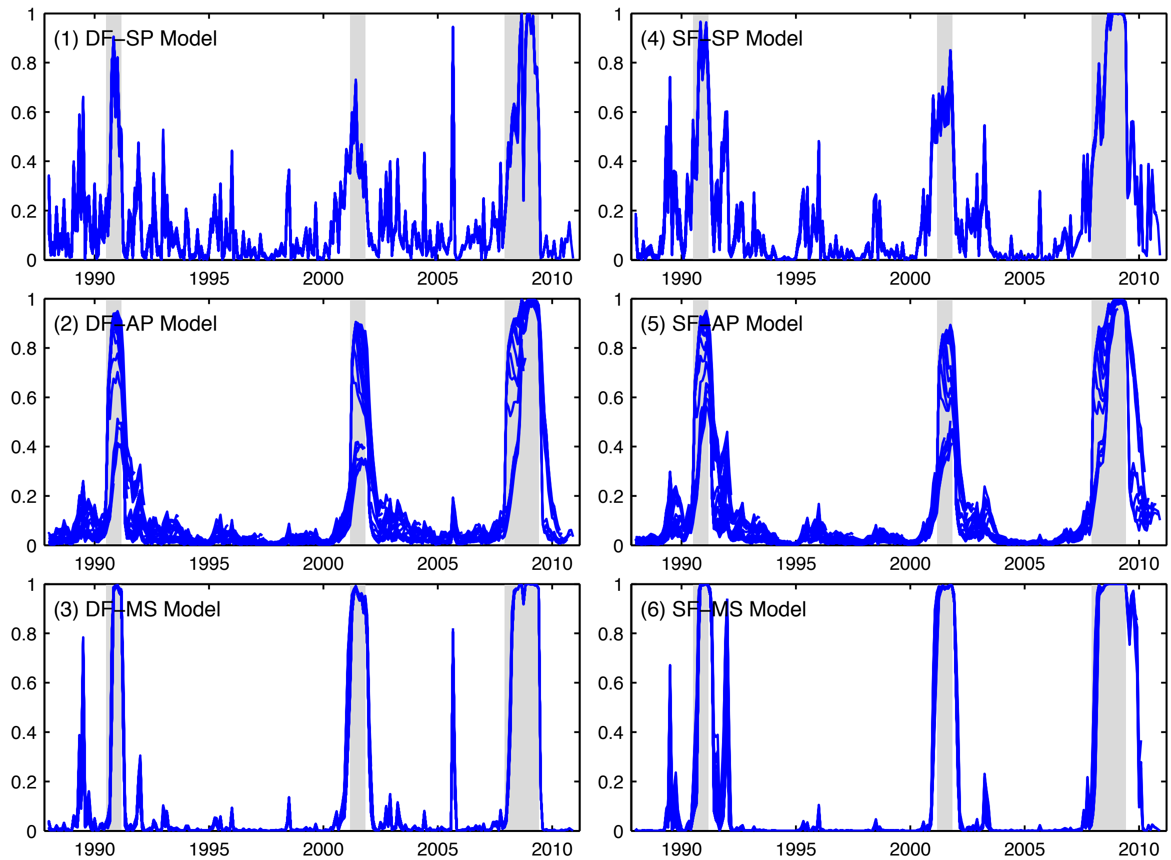


Figure 7: Probabilities of recession (paths) for the hold-out sample period 1988:1 – 2010:12. Shaded areas denote NBER recession months.

**Department of Economics, University of Alberta**
**Working Paper Series**

http://www.economics.ualberta.ca/WorkingPapers.aspx

| |
|---|
| **2011-04:** Covariate Unit Root Tests with Good Size Power – **Fossati** |
| **2011-03:** New measures of the costs of unemployment: Evidence from the subjective well-being of 2.3 million Americans – Helliwell, **Huang** |
| **2011-02:** Childhood Determinants of Risk Aversion: The Long Shadow of Compulsory Education – **Hryshko,** Luengo-Prado, |
| **2011-01:** Will Biofuel Mandates Raise Food Prices? – **Chakravorty,** Hubert, Moreaux, Nostbakken |
| **2010-20:** Does the Retirement Consumption Puzzle Differ Across the Distribution: - Fisher, **Marchand** |
| **2010-19:** A Test of Monopoly Price Dispersion Under Demand Uncertainty – **Humphreys,** Soebbing |
| **2010-18:** Split Incentives and Energy Efficiency in Canadian Multi-Family Dwellings – **Maruejols, Young** |
| **2010-17:** Local Labor Market Impacts of Energy Boom-Bust-Boom in Western Canada - **Marchand** |
| **2010-16:** House Prices and Risk Sharing – **Hryshko**, Luengo-Prado, Sørensen |
| **2010-15:** Education vs. Optimal Taxation: The Cost of Equalizing Opportunities – **Stephens** |
| **2010-14:** The Economic Choice of Participation and Time Spent in Physical Activity and Sport in Canada – **Humphreys, Ruseski** |
| **2010-13:** Do Gamblers Think That Teams Tank? Evidence from the NBA – Soebbing, **Humphreys** |
| **2010-12:** Would Hotelling Kill the Electric Car? – **Chakravorty,** Leach, Moreaux |
| **2010-11:** Residential Land Use Regulation and the US Housing Price Cycle Between 2000 and 2009 – **Huang,** Tang |
| **2010-10:** Government Revenue Volatility in Alberta – **Landon, Smith C.** |
| **2010-09:** Sports Participation and Happiness: Evidence from U.S. Micro Data – **Huang, Humphreys** |
| **2010-08:** Does Financial and Goods Market Integration Matter for the External Balance? A Comparison of OECD Countries and Canadian Provinces – **Smith, C** |
| **2010-07:** Consumption Benefits and Gambling: Evidence from the NCAA Basketball Betting Market – **Humphreys**, Paul, Weinbach |
| **2010-06:** On Properties of Royalty and Tax Regimes in Alberta's Oil Sands – **Plourde** |
| **2010-05:** Prices, Point Spreads and Profits: Evidence from the National Football League - **Humphreys** |
| **2010-04:** State-dependent congestion pricing with reference-dependent preferences - **Lindsey** |
| **2010-03:** Nonlinear Pricing on Private Roads with Congestion and Toll Collection Costs – Wang, **Lindsey**, Yang |
| **2010-02:** Think Globally, Act Locally? Stock vs Flow Regulation of a Fossil Fuel – Amigues, **Chakravorty,** Moreaux |
| **2010-01:** Oil and Gas in the Canadian Federation – **Plourde** |

**Please see above working papers link for earlier papers**



FACULTY OF ARTS