



<http://economix.u-paris10.fr/>

Document de Travail  
Working Paper  
2010-14

A factor-augmented probit model  
for business cycle analysis

Christophe Bellégo  
Laurent Ferrara



FRE 3257

Université de Paris Ouest Nanterre La Défense  
(bâtiments K et G)  
200, Avenue de la République  
92001 NANTERRE CEDEX

Tél et Fax : 33.(0)1.40.97.59.07  
Email : [secretariat-economix@u-paris10.fr](mailto:secretariat-economix@u-paris10.fr)

université  
Paris | Ouest  
  
Nanterre La Défense

# A factor-augmented probit model for business cycle analysis

C. Bellégo <sup>\*</sup>, L. Ferrara <sup>†</sup>

May 17, 2010

## Abstract

Dimension reduction of large data sets has been recently the topic of interest of many research papers dealing with macroeconomic modelling. Especially dynamic factor models have been proved to be useful for GDP nowcasting or short-term forecasting. In this paper, we put forward an innovative factor-augmented probit model in order to analyze the business cycle. Factor estimation is carried either by standard statistical methods or by allowing a richer dynamic behaviour. An application is provided on euro area data in order to point out the ability of the model to detect recessions over the period 1974-2008.

---

<sup>\*</sup>Ecole Normale Supérieure de Cachan - ENSAE, email: christophe.bellego@ensae.fr

<sup>†</sup>Corresponding author, Banque de France and EconomiX, University Paris Ouest Nanterre, email: laurent.ferrara@banque-france.fr. We would like to thank M. Billio and participants of the internal seminar of the Econometrics Department at University Ca' Foscari in Venice. The views expressed herein are those of the authors and do not necessarily reflect those of the Banque de France.

# 1 Introduction

In business cycle analysis, parametric modelling is often used by economists in central banks or governmental institutions in order to assess periodically the probability of being in a given phase of the business cycle, defined here in the NBER sense by the alternation of expansion and recession phases (see for example NBER, 2008). In this respect, binary response models, such as logit and probit models, have been proved to be of great interest for practitioners aiming at monitoring the business cycle (see Layton and Katsuura, 2001, for a comparison of various parametric models). The empirical literature on this topic is huge but among the recent papers we refer for example to Estrella, Rodrigues and Schich (2003), Chauvet and Potter (2002, 2005), Kauppi and Saikkonen (2008) or Nyberg (2010) for the US and to Moneta (2005), Duarte et al. (2005) or Bellégo and Ferrara (2009) for the euro area as a whole. Binary response models establish a relationship between a low number of explanatory variables and a binary dependent variable. As an output, a probability of being in a given phase of the business cycle is estimated according to the values of the explanatory variables. Explanatory variables are generally chosen among a wide range of financial (interest rate term spread, asset prices ...) and macroeconomic (industrial production, employment, ...) time series. The dependent variable, denoted  $(r_t)_t$ , is binary and takes the value 1 if the economy is in recession and 0 otherwise.

Recently, another strand of the literature in applied macroeconomics has focused on econometric modelling in a data-rich environment. Specifically, econometric models have been put forward to reduce the dimension of databases in order to use the most meaningful information, in a given sense. In this respect, dynamic factor models have proved their usefulness, specially in order to nowcast or forecast macroeconomic variables. Among the recent empirical literature on this topic, we refer for example to the papers of Boivin and Ng (2006), Schumacher (2007, 2010), Giannone, Reichlin and Small (2008) or Barhoumi, Darné and Ferrara (2010). Generally, the factor estimation step is carried out either by using standard statistical methods of dimension reduction (Stock and Watson, 2002) or by using estimation methods that allow a richer dynamic behaviour for the factor (Forni et al., 2004, 2005, or Doz, Giannone and Reichlin, 2007). According to the empirical literature, there is yet no consensus about the preference towards dynamic factors in modelling.

In this paper, we reconcile both strands of the literature by putting forward a factor-augmented probit model that allows to get for each date a probability of recession conditionally to a large informational set. This conditional probability is estimated from a probit model that includes as explanatory variables the estimated dynamic or static factors. This model contributes to the recent generalisation of econometric models by incorporating synthetic factors into the standard equations, enabling thus to estimate tractable low-dimensional econometric models conditionally to a large informational set. For example, we refer to the factor-augmented VAR (FAVAR) of Bernanke, Boivin and Elias (2005) or to the factor-augmented error correction model by Banerjee and Marcellino (2008). As an application, we point out the ability of the factor-augmented probit model to replicate business cycles in the euro area from 1974 to 2008 by using a set of macroeconomic and financial variables that have been already proved to be reliable in-

dicators of recession. Section 2 presents the model and section 3 contains the details of the application that we have carried out in order to replicate the euro area business cycle over the period January 1974 - December 2008.

## 2 Modelling

In this section we introduce the factor-augmented probit model that we put forward to handle large databases in order estimate the occurrence of turning points in the business cycle through a probit model, conditionally to this large information.

Assume we observe a vector of  $n$  stationary zero mean time series  $x_t = [x_{1t}, \dots, x_{nt}]'$ , for  $t = 1, \dots, T$ . The objective of dynamic factor modelling is to decompose  $(x_t)_t$  into a sum of two mutually orthogonal unobservable components: a common component of low dimension  $(\chi_t)_t$ , summarizing the dynamics common to all the series, and an idiosyncratic component  $(\xi_t)_t$ , specific to each series. The common component  $(\chi_t)_t$  is supposed to linearly summarize the common behaviour of the  $n$  series. For  $t = 1, \dots, T$ , the static factor model is defined by

$$x_t = \Lambda f_t + \xi_t, \quad (1)$$

where  $\Lambda$  is the loading matrix of dimension  $(n \times r)$ , the common component  $\chi_t = \Lambda f_t$  is driven by a small number  $r$  of factors  $f_t$  common to all the variables in the model such that  $f_t = [f_{1t}, \dots, f_{rt}]$ , and  $\xi_t = [\xi_{1t}, \dots, \xi_{nt}]'$  is a vector of  $n$  idiosyncratic mutually uncorrelated components, driven by variable-specific shocks.

To take dynamics into account in modelling, an alternative specification integrates explicitly the dynamics of the factors  $f_t$ . Specifically the dynamic factor representation is supposed to be given by the following equation

$$x_t = A(L)f_t + \xi_t, \quad (2)$$

where the common component  $\chi_t = A(L)f_t$  integrates a linear dynamics where  $A(L)$  is a  $(n \times r)$  matrix describing the autoregressive form of the  $r$  factors. If we assume that there exists a  $(n \times q)$  matrix  $B(L)$  such that  $B(L) = A(L)N(L)$  with  $N(L)$  of dimension  $(r \times q)$ , then the dynamic factor is such that  $f_t = N(L)U_t$  where  $U_t$  is a  $(q \times 1)$  independent vector containing the dynamic shocks. It follows that the factor dynamics are described by

$$A(L)f_t = B(L)U_t, \quad (3)$$

which specifies a VAR model for the factor  $f_t$ .

We assume now that we observe the values of a binary variable  $(r_t)_t$  that takes for value 1 when the economy is in recession at date  $t$  and 0 otherwise. Binary response models rely on the assumption that the values of the binary dependent variable,  $(r_t)_t$  stem from a latent continuous variable, denoted  $(y_t)_t$ , defined by the following general linear equation, for all  $t$ ,:

$$y_t = \alpha + \beta'_0 f_t + \dots + \beta'_k f_{t-k} + \varepsilon_t, \quad (4)$$

where for  $j = 0, \dots, k$ ,  $f_{t-j} = (f_{t-j}^1, \dots, f_{t-j}^r)'$  is a  $r$ -vector of lagged dynamic factors used as explanatory variables,  $\beta_j = (\beta_j^1, \dots, \beta_j^r)'$  is a  $r$ -vector parameter and  $(\varepsilon_t)_t$  is the error term supposed to be a strong white noise process with finite variance  $\sigma_\varepsilon^2$ .

The binary response modelling relies on the following relationship between  $(r_t)_t$  and  $(y_t)_t$

$$r_t = \begin{cases} 1 & \text{if } y_t \leq 0, \\ 0 & \text{if } y_t > 0 \end{cases} \quad (5)$$

For each date  $t$ , it can be easily proved that the conditional probability of recession is given by

$$P(r_t = 1 | I_t) = F(-\alpha - \beta'_0 f_t - \dots - \beta'_k f_{t-k}), \quad (6)$$

where  $I_t$  represents the large informational set available at date  $t$  and  $F(\cdot)$  is the cumulative density function of  $(\varepsilon_t)_t$ . We assume here that the distribution of  $(\varepsilon_t)_t$  is supposed to be Gaussian in order to get the probit specification of the model (the logit specification is obtained when  $F(\cdot)$  is the logistic function).

Parameter estimation of equations (1) to (6) is carried out using a two-step procedure. First, dynamic factors  $(\hat{f}_t)_t$  are estimated, then those estimated factors are put into equation (4) in order to estimate the conditional probability of being in recession given in (6). Note that this two-step procedure implicitly incorporates the uncertainty inherent to the measure of factors.

Regarding factor estimation in the static framework of equation (1), we implement the procedure of Stock and Watson (2002) using static principal component analysis (PCA) to estimate the factors  $\hat{f}_t$ . An eigenvalue decomposition of the empirically estimated covariance matrix,  $T^{-1} \sum_{t=1}^T x_t x_t'$ , provides the  $(n \times r)$  eigenvector matrix  $\hat{S} = (\hat{S}_1, \dots, \hat{S}_r)$  containing the eigenvectors  $\hat{S}_j$  corresponding to the  $r$  largest eigenvalues for  $j = 1, \dots, r$ . The factor estimates are the first  $r$  principal components of  $x_t$  defined as  $\hat{f}_t^{sta} = \hat{S}' x_t$ .

In the dynamic framework of equations (2) and (3), factor estimation is carried out by using a 2-step Kalman filter to compute ML estimates as proposed by Doz, Giannone and Reichlin (2007). This approach consists in estimating first the parameters by PCA, then, in the second step, the model is put into a state-space form and factors are estimated via Kalman smoothing. We note  $\hat{f}_t^{dyn}$  the estimated factors.

The choice of the number of factors  $r$  to include in equation (4) is determined by using the Bai and Ng (2002) test. We implement the version of the test that determines  $r$  by minimizing the following penalized information criterion

$$IC(k) = \ln [V(k, F)] + k \times \left( \frac{n+T}{nT} \right) \ln \left( \frac{nT}{n+T} \right), \quad (7)$$

where  $V(r, F)$  is the sum of squared residuals such that:  $V(k, F) = (nT)^{-1} \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \Lambda f_t)^2$ . Note that Bai and Ng (2002) also put forward a variety of information criteria that vary

according to the penalization function.

To assess the goodness-of-fit of binary response models, Estrella (1998) proposes a *Pseudo- $R^2$*  measure given by the following expression

$$Pseudo - R^2 = 1 - \frac{L_u}{L_c}^{-2/n} L_c \quad (8)$$

where  $L_u$  is the log-likelihood of the considered model and  $L_c$  is the log-likelihood of a reference nested model to which  $L_u$  is compared. By construction, the nested model must have a lower log-likelihood value than the basic model. This *Pseudo- $R^2$*  measure is often used in the literature for comparison between models (see for example Estrella et al., 2003).

We also decide to focus on a general goodness-of-fit criterion to assess the quality of the models, often used in business cycle analysis (see, among others, Anas et al., 2008), namely the quadratic probability score (QPS) given by

$$QPS = \frac{1}{T} \sum_{t=1}^T (r_t - \hat{P}(r_t = 1|I_t))^2, \quad (9)$$

where  $\hat{P}(r_t = 1|I_t)$  is the estimated conditional probability of being in recession, conditionally to current large information and estimated parameters.

### 3 Empirical aspects

In this section we apply the previous factor-augmented probit model to euro area data to assess the reliability of the model to replicate recessions in the zone, from January 1974 to December 2008.

#### 3.1 Data description

We focus only on data that have been proved to be useful to detect recession and that are available since the early seventies. In this respect, we first consider industrial production and unemployment rate, for both the US and the euro area, which are generally assessed to estimate recession dates, as for example by the NBER Dating Committee. As other useful series, we include the consumer prices index. Note that, because of a lack of historical availability, we do not consider business surveys. Moreover, business surveys have not proved so far their reliability for recession detection.

Many research papers have pointed out the leading ability of financial variables (see for example Stock and Watson, 2003, Estrella, Rodrigues and Schich, 2003, or King, Levin and Perli, 2007, and the references therein). Thus in this paper, we test several variables from financial markets in the euro area, such as stock market index, oil prices, various

term spread, monetary aggregates ..., with relatively long sample (we refer to Bellégo and Ferrara, 2009, for a detailed description of this database).

Finally, we get a medium-size database of 18 variables from January 1974 to December 2008, 5 from real economy and 13 from financial markets (see Table 1 for details). For each variable, we determine whether the variable is stationary or not by using standard tests. When the variable is stationary, we use it without any transformation (i.e. in level), otherwise we transformate it by taking the growth rate, or the differences, over one month.

Concerning the benchmark dating chronology  $(r_t)_t$ , we use the business cycle turning point chronology proposed by Anas et al. (2007). They provide both a quarterly chronology based on GDP and a monthly chronology based on IPI. As we need a monthly reference for our exercise, we choose the IPI chronology except for the year 2001 in which for the first time in euro area countries an industrial recession occurred without a global recession. Since 1974, we retain thus four recession periods until 2006 from peak to trough: March 1974 - March 1975, January 1980 - December 1980, September 1981 - December 1982 and February 1992 - February 1993. Regarding the last recession, we assume that a peak occurred in March 2008 (see also CEPR, 2009) and that the economy is still in recession in December 2008, the last point of the sample.

## 3.2 Results

First we carry out a test to determine the number  $r$  of factors to include in the factor-augmented probit model given in equation (6). The Bai and Ng (2002) test leads us to assume that  $r = 2$  factors, denoted  $(f_{1t})_t$  and  $(f_{2t})_t$ , are sufficient to provide a good description of the database. Then the two factors are estimated over the complete period using both the static model, denoted by  $(\hat{f}_{1t}^{sta})_t$  and  $(\hat{f}_{2t}^{sta})_t$ , and the dynamic model, denoted by  $(\hat{f}_{1t}^{dyn})_t$  and  $(\hat{f}_{2t}^{dyn})_t$ . Note that the two first principal factors represent around 30 percent of the total variance and that other factors possess an idiosyncratic variance lower than 10 percent.

The first estimated factor-augmented probit model is the model (6) without any lags ( $k = 0$ ) whose conditional probability of being in recession is given by

$$P(r_t = 1|I_t) = F(-\alpha - \beta_0^1 \hat{f}_{1t} - \beta_0^2 \hat{f}_{2t}), \quad (10)$$

where  $I_t$  is the information set available at time  $t$ . Parameter estimation results are presented in Table 2, for both static and dynamic factors. The first factor always possesses a negative coefficient, meaning thus that a decrease in this factor leads to an improvement of the recession probability, while the second one presents a positive coefficient. The dynamic version of the model presents better goodness-of-fit criteria than the static version (lower QPS and higher Pseudo-R<sup>2</sup>). Estimated probabilities are presented in Figure 1. Both approaches estimate similar conditional probabilities of being in recession, although the dynamic approach provides a smoother probability. This fact is related to the more persistent behaviour of the dynamic factor that integrates autoregressive lags by definition.

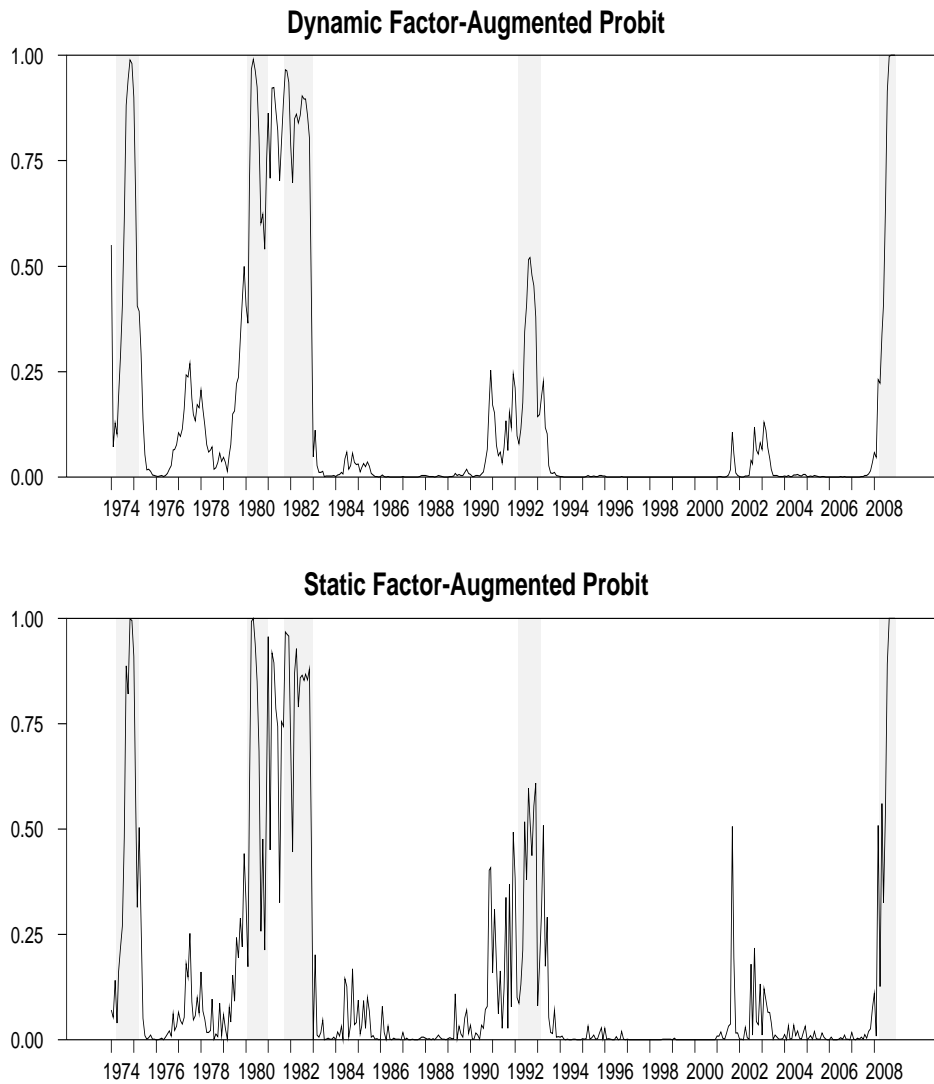


Figure 1: Estimated probabilities with  $k = 0$  of being in recession in the euro area stemming from the factor-augmented probit model using both dynamic (top) and a static (bottom) factors, from Feb. 1974 to Dec. 2008. Shaded areas correspond to the benchmark chronology of recessions (source Anas et al., 2007).

It turns out that upward movements in the conditional probability generally correspond to benchmark recession phases. However, the double-dip movement in early eighties is not clearly reproduced by the models and in addition the conditional probabilities barely cross the critical value of 0.50 during the 1992-93 recession. This latter point reflects the fact that the 1992-93 recession belongs to the *Great Moderation* period characterized by a lower amplitude in macroeconomic fluctuations of industrialized countries. The maximum value of 1.0 is only reached briefly during the first oil shock and during the last 2008 recession. We also note that the year 1977 is characterized by an upward movement in the estimated probabilities that stay however largely below the threshold of 0.50.



In a second step, we take advantage of the leading behaviour of financial variables over recession periods. Indeed it is well known that financial data possess a forward-looking property and can be thus taken as leading indicators of recession with an average lead ranging from 3 to 12 months over the business cycle (see for example Estrella et al., 2003, Moneta, 2005, King et al. 2007, or Bellégo and Ferrara, 2009). Moreover, real data are well known by practitioners to reflect economic activity in a coincident manner. In this respect, we allow the model to reflect both coincident and leading properties in the dataset by imposing the constraint that  $\beta_1 = \dots = \beta_{k-1} = 0$  in equation (6), for several values of  $k$  such that  $k \in \{3, 6, 9, 12\}$ . Thus, the conditional probability of being in recession at any date  $t$  is given by

$$P(r_t = 1|I_t) = F(-\alpha - \beta_0^1 \hat{f}_{1t} - \beta_0^2 \hat{f}_{2t} - \beta_k^1 \hat{f}_{1,t-k} - \beta_k^2 \hat{f}_{2,t-k}), \quad (11)$$

where  $k = 3, 6, 9, 12$  months.

Parameter estimation results are presented in Table 2 for the various lags  $k$ . From Table 2, we note that the coincident factors always have a significant impact on the business cycle according to the  $t$ -stat statistics. For dynamic factors, the first factor lagged  $k$  months,  $\hat{f}_{1,t-k}^{dyn}$ , is not significant for lags  $k = 9$  and  $k = 12$ . In opposition, the influence of the second factor tends to increase with  $k$ , for both static and dynamic approaches. This evolution may reflect the growing impact of financial variables when the horizon increases. This could mean that the second factor is related to the financial variables while the first one reflects rather the contemporaneous behaviour of hard data. It turns out that the inclusion of a lag in the factor-augmented equation enables to improve the goodness-of-fit of the models. The optimal goodness-of-fit criterion is reached for  $k = 12$  months, which can be interpreted as the estimated lead of financial variables over recessions.

Estimated probabilities of recession for  $k = 12$  are presented in Figure 2. We observe that conditional probability values have been overall increased and reach now the maximum value of 1.0 at several consecutive dates, leading thus to a more understandable signal. The critical value of 0.50 is now largely crossed during the 1992-93 recession, but a double-dip signal now appears, reflecting perhaps the US 1990-91 recession. Again, the dynamic factor-augmented probit model provides a smoother signal, easier to interpret. For example, regarding the timing of the last recession, the dynamic version of the model with  $k = 12$  identifies a peak in February 2008 (benchmark peak is located in March 2008), when the probability crosses durably the critical value of 0.50, while the static version is noisier and only identifies a readable peak in June 2008.

## 4 Conclusions

In this paper, we have put forward an innovative factor-augmented probit model able to replicate the business cycle. This model enables to summarize a large database into few estimated factors that are then included into a probit model. As an output we get an estimated conditional probability of being in recession each month. An application is provided on euro area data in order to point out the ability of the model to detect

recession. We provide a model specification that enables to take the leading pattern of financial variables contained in the data set into account. As further research, it would be worthwhile to check the robustness of the model using real-time data in order to monitor the euro area business cycle.

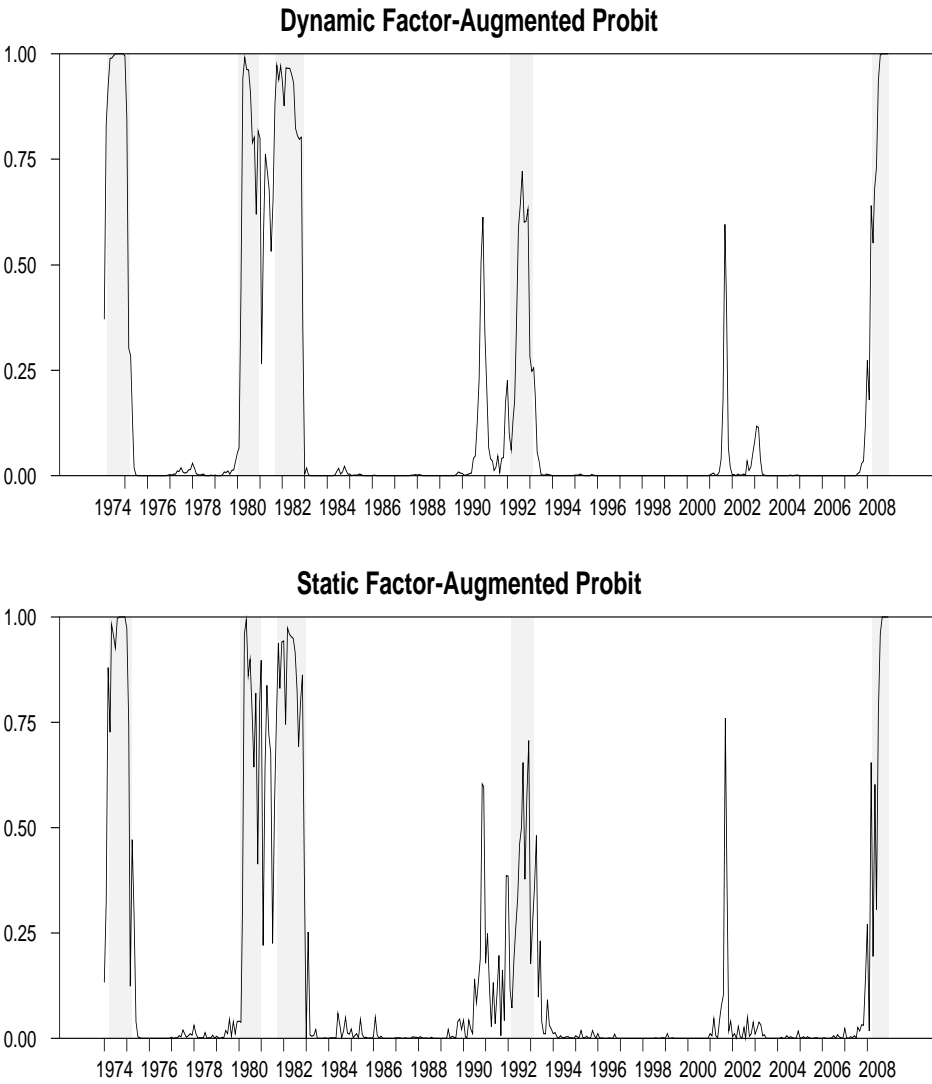


Figure 2: Estimated probabilities with  $k = 12$  of being in recession in the euro area stemming from the factor-augmented probit model using both dynamic (top) and a static (bottom) factors, from Feb. 1974 to Dec. 2008. Shaded areas correspond to the benchmark chronology of recessions (source Anas et al., 2007).

## References

- [1] Anas J., Billio M., Ferrara L. and LoDuca M. (2007), A turning point chronology for the Euro-zone classical and growth cycles, in *Growth and Cycles in the Euro-zone*, G.L. Mazzi and G. Savio (eds.), Palgrave-MacMillan, New-York.
- [2] Anas J., Billio M., Ferrara L. and Mazzi G.-L. (2008), A system for dating and detecting turning points in the euro area, *The Manchester School*, 76, 5, 549-577.
- [3] Bai, J. and S. Ng (2002), Determining the number of factors in approximate factor models, *Econometrica*, 70, 191-221.
- [4] Banerjee, A. and M. Marcellino (2008), Factor-augmented error correction model, CEPR Discussion Paper No. 6707.
- [5] Barhoumi, K., O. Darné and L. Ferrara (2010), Are disaggregate data useful for forecasting French GDP with dynamic factor models ?, *Journal of Forecasting*, 29, 1-2, 132-144
- [6] Bernanke, B., J. Boivin and P. Eliasch (2005), Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach, *Quarterly Journal of Economics*, 120, 1, 387-422.
- [7] Bellégo C. and L. Ferrara (2009), Forecasting euro area recessions using time-varying binary response models for financial variables, Working Paper, Banque de France, No. 259
- [8] Boivin J. and Ng S. (2005), Understanding and comparing factor-based forecasts, *International Journal of Central Banking*, 1, 117-151.
- [9] CEPR (2009), Euro Area Business Cycle Dating Committee: Determination of the 2008 Q1 Peak in Economic Activity, CEPR, March 2009.
- [10] Chauvet, M. and Potter, S. (2002), Predicting recessions: Evidence of the yield curve in the presence of structural breaks, *Economics Letters*, 77,2, 245-253.
- [11] Chauvet, M. and Potter, S. (2005), Forecasting recessions using the yield curve, *Journal of Forecasting*, 24, 2, 77-103.
- [12] Doz C., Giannone D. and Reichlin L. (2007), A two-step estimator for large approximate dynamic factor models based on Kalman filtering, CEPR Discussion Paper No 6043.
- [13] Duarte, A., Venetis, I.A. and Paya, I. (2005), Predicting real growth and the probability of recession in the Euro-area using the yield spread, *International Journal of Forecasting*, 21, 2, 261-277.
- [14] Estrella A. (1998), A new measure of fit for equations with dichotomous dependent variables, *Journal of Business and Economic Statistics*, 16, 2, 198-205.

- [15] Estrella A. and Mishkin F. S. (1998), Predicting US recessions: Financial variables as leading indicators, *Review of Economics and Statistics*, vol. 80, pp. 45-61.
- [16] Estrella, A., Rodrigues, A. P., and Schich, S. (2003), How stable is the predictive power of the yield curve ? Evidence from Germany and the United States, *Review of Economics and Statistics*, 85, 3, 629-644.
- [17] Forni M., Hallin M., Lippi M. and Reichlin L. (2004), The generalized dynamic factor model: Consistency and rates, *Journal of Econometrics*, 119, 231-255.
- [18] Forni M., Hallin M., Lippi M. and Reichlin L. (2005), The generalized dynamic factor model: One-sided estimation and forecasting, *Journal of the American Statistical Association*, 100, 830-840.
- [19] Giannone D., Reichlin L. and Small D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55, 665-676.
- [20] Hamilton, J. (1994), *Time Series Analysis*, Princeton University Press.
- [21] Kauppi, H. and Saikkonen, P. (2008), Predicting US recessions with dynamic binary response models, *The Review of Economics and Statistics*, 90, 4, 777-791.
- [22] King, T.B., Levin, A.T., and Perli, R. (2007), Financial market perceptions of recession risk, FEDS Working Paper No. 2007-57.
- [23] Layton, A. and Katsuura, M. (2001), Comparison of regime switching, probit and logit models in dating and forecasting US business cycles, *International Journal of Forecasting*, 17, 403-417.
- [24] Moneta, F. (2005), Does the yield spread predict recessions in the Euro area?, *International Finance*, 8, 2, 263-301.
- [25] NBER (2008), Determination of the December 2007 peak in economic activity, Business Cycle Dating Committee, NBER, December 2008.
- [26] Schumacher C. (2007), Forecasting German GDP using alternative factor models based on large datasets, *Journal of Forecasting*, 26, 271-302.
- [27] Schumacher C. (2010), Factor forecasting using international targeted predictors: The case of German GDP, *Economics Letters*, 107, 95-98.
- [28] Stock J. and Watson M. (2002), Macroeconomic forecasting using diffusion indexes, *Journal of Business and Economic Statistics*, 20, 147-162.
- [29] Stock, J.H. and M.W. Watson (2003), Forecasting output and inflation: The role of asset prices, *Journal of Economic Literature*, 41, 3, 788-829.

## APPENDIX

<b>Variable</b>	<b>Transformation</b>	<b>Source</b>
EA dividend yield	level	Datastream
EA P/E ratio	level	Datastream
EA stock index	1-month growth	Datastream
German spread 10y-3m	level	Bundesbank
German spread 3m-1y	level	Bundesbank
German spread corporate-1y	level	Bundesbank
German spread corporate-3m	level	Bundesbank
US spread corporate-1y	level	US treasury and Moody's
US default rate	1-month growth	Moody's
CRB spot index	1-month growth	Commodity Research Bureau
Oil spot Dollar	1-month growth	Bloomberg
Gold Dollar	1-month growth	gold Bullion LBM
EA M1	1-month growth	OCDE
EA IPI	1-month growth	Eurostat
US IPI	1-month growth	BEA
EA unemployment rate	1-month difference	Eurostat
US unemployment rate	1-month difference	Census Bureau
EA inflation rate	1-month difference	Eurostat

Table 1: Database description (EA= Euro Area)

Method	$k$	QPS	Pseudo- $R^2$	$\alpha$	$\beta_0^1$	$\beta_0^2$	$\beta_k^1$	$\beta_k^2$
Static	0	0.055	0.504	1.980 (10.7)	-0.592 (-7.9)	0.718 (7.0)		
	3	0.049	0.558	2.194 (9.9)	-0.239 (-2.1)	0.704 (5.4)	-0.445 (-4.2)	0.030 (0.3)
	6	0.044	0.582	2.203 (9.7)	-0.325 (-3.3)	0.692 (6.3)	-0.393 (-4.5)	-0.124 (-1.3)
	9	0.044	0.592	2.229 (9.7)	-0.403 (-3.9)	0.790 (6.3)	-0.306 (-3.5)	-0.328 (-3.1)
	12	0.043	0.607	2.305 (9.2)	-0.437 (-4.0)	0.791 (6.2)	-0.304 (3.3)	-0.404 (-3.6)
Dynamic	0	0.051	0.539	2.221 (10.0)	-0.633 (-8.2)	0.872 (7.1)		
	3	0.045	0.580	2.357 (9.5)	-0.302 (-2.2)	1.074 (5.4)	-0.383 (-3.1)	-0.252 (-1.5)
	6	0.041	0.612	2.380 (9.4)	-0.435 (-3.9)	1.077 (6.4)	-0.276 (-2.9)	-0.369 (-2.8)
	9	0.037	0.652	2.678 (8.1)	-0.602 (-4.2)	1.289 (6.1)	-0.182 (-1.6)	-0.637 (-4.1)
	12	0.035	0.670	2.284 (7.6)	-0.654 (-4.2)	1.310 (6.1)	-0.183 (-1.5)	-0.722 (-4.4)

Table 2: Estimated parameters ( $t$ -stat in parenthesis) and criteria for the factor-augmented probit model with 2 coincident factors  $f_{1t}$  and  $f_{2t}$  and their lagged versions for a given lag  $k$ ,  $f_{1,t-k}$  and  $f_{2,t-k}$