Forecasting the state of the Finnish business cycle

Harri Pönkä and Markku Stenborg

Ministry of Finance, Finland

12 October 2018

Forecasting the state of the Finnish business cycle*

Harri Pönkä† Markku Stenborg‡

Abstract

We employ probit models to study the predictability of recession periods in Finland using a set of commonly used variables based on previous literature. The findings point out that individual predictors, including the term spread and the real housing prices from the capital area, are useful predictors of recession periods. However, the best in-sample fit is found using combinations of variables. The pseudo out-of-sample forecasting results are generally in line with the in-sample results, and suggest that in the one-quarter ahead forecasts a model combining the term spread, the unemployment expectation component of the consumer confidence index, and the consumer confidence index performs the best based on the area under the receiver operating characteristic curve. An autoregressive specification improves the in-sample fit of the models compared to the static probit model, but findings from pseudo out-of-sample forecasts vary between forecasting horizons.

Keywords: Business cycle, Recession period, Probit model
JEL classification: C22, G12, G17

*The authors would like to thank Henri Nyberg, David Turner, and Yi Zheng for useful discussions and comments on the paper. This paper is related to an ongoing research by Henri Nyberg and Harri Pönkä entitled "Business cycle transmission from large countries small open economies." Disclaimer: The opinions expressed in this article are the authors' own and do not reflect the views of the Finnish Ministry of Finance.

†Address: Ministry of Finance, Economics Department, Snellmaninkatu 1 A, Helsinki. PO Box 28, FI-00023 Government, Finland, E-mail: firstname.lastname@vm.fi.

‡ Ministry of Finance, Economics Department, Snellmaninkatu 1 A, Helsinki. PO Box 28, FI-00023 Government, Finland.

1 Introduction

In the theoretical and applied literature on business cycle analysis the identification of reliable early warning indicators for future economic downturns is of upmost importance. Nonlinear probability models, such as probit and logit models, have been a standard tool in modelling the conditional probability of recessions based on potential predictive variables ever since Estrella and Hardouvelis [1991] used them to study business cycle fluctuations in the US. Empirical findings based on these models have suggested a number of such leading indicators and particular attention has been paid on the term spread and stock market returns (see, e.g., Estrella and Mishkin [1998], Chauvet and Potter [2005], Nyberg [2010], and Ng [2012]). Further research has highlighted the role of sentiment (Christiansen et al. [2014]) and credit variables (Pönkä [2017]) as useful predictors of US recession periods.

In addition to progress in the identification of potential predictors, several methodological developments have been made in the past years. Kauppi and Saikkonen [2008] introduced dynamic and autoregressive extensions to the conventional static probit models and found them able to improve forecasts of US recession periods. Factor-augmented probit models, based on the use of principal components (Stock and Watson [2002]), were first introduced by Chen et al. [2011] and later employed by e.g. Bellégo and Ferrara [2012], Christiansen et al. [2014], and Pönkä [2017]. Other methods employed in recession forecasting include e.g. model averaging and boosting (Berge [2015]) as well as Bayesian shrinkage (Fornaro [2016]). Although newly developed methods have in some cases been found to outperform the results obtained from more traditional probit models, the probit model has remained the standard tool in the literature, partly due to the ease of interpretability of the findings.

The purpose of this paper is to apply probit models to predict business cycle fluctuations in Finland. The majority of academic research on business cycle forecasting has focused on the US, including all of the aforementioned studies with the exception of Bellégo and Ferrara [2012].[1] Therefore, one of the contributions of this paper is to see how well the findings in the literature hold for a small and open economy, such as Finland. We focus on a small set of commonly used recession predictors, based on previous academic research. The purpose of this paper is illustrative and we hope that it will motivate further work on the topic.

Our findings suggest that individual predictors, including the term spread, the real housing prices from the capital area, and measures of consumer confidence, are useful predictors of Finnish recession periods. However, the best in-sample fit is found using multiple variables. The pseudo out-of-sample forecasting results are generally in line with the in-sample results, and suggest that in the one-quarter ahead forecasts a model combining the term spread, the unemployment expectation component of the consumer confidence index, and the consumer confidence index performs the best based on the area under the receiver operating characteristic curve (AUC). An autoregressive specification improves the in-sample fit of the models compared to the conventional and commonly used static probit model, but

[1] Examples of cross-country studies include Bernard and Gerlach [1998] and Sensier et al. [2004].

findings from pseudo out-of-sample forecasts give mixed results between different forecasting horizons.

The remainder of this paper is organised as follows. In Section 2, we discuss the econometric framework the goodness-of-fit measures and statistical tests. Section 3 describes the data, including the employed business cycle chronology and the predictive variables. In Section 4, we report in-sample and out-of-sample forecasting results. Finally, in Section 5 we conclude and discuss possible extensions of this study.

2 Econometric methodology

In this section we briefly present the econometric framework and discuss goodness-of-fit measures related to the binary response models.

2.1 The probit model

We are interested in predicting the state of the Finnish economy, defined as a binary indicator

$$y_t = \begin{cases} 1, \text{if the economy is in a recession,} \\ 0, \text{if the economy is in an expansion.} \end{cases} \tag{1}$$

The business cycle chronology y_t used in this paper will be discussed further in Section 3. The conditional probability of the economy being in a state of recession (p_t) is modelled using a univariate probit model

$$p_t = P_{t-1}(y_t = 1) = \Phi(\pi_t), \tag{2}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution and π_t is a linear function of the variables in the information set Ω_{t-1}. In the most commonly used model, the so-called static probit model, π_t is specified as

$$\pi_t = \omega + \boldsymbol{x}'_{t-k}\boldsymbol{\beta}, \tag{3}$$

where ω is a constant term and $\boldsymbol{x}_{t-k}$ includes the k:th lagged values of the explanatory variables. The parameters of the probit model can be estimated using the maximum likelihood method, and Newey-West-type robust standard errors are typically used. For more details on the estimation of the probit model and the standard errors, we refer to Kauppi and Saikkonen [2008] and de Jong and Woutersen [2011].

As an extension, we consider the autoregressive probit model proposed by Kauppi and Saikkonen [2008]. In this specification, an autoregressive structure is introduced using the lagged value of the linear function π_t, as follows

$$\pi_t = \omega + \alpha_1 \pi_{t-1} + \boldsymbol{x}'_{t-k}\boldsymbol{\beta}. \tag{4}$$

Further research has found evidence in favor of the autoregressive probit model (4) over the static model (3). Nyberg [2010, 2014] found them to outperform static probit models in predicting US and German recessions, whereas Pönkä [2017] found minor improvements over the static probit model.

2.2 Goodness-of-fit Measures

There are several measures to evaluate the goodness-of-fit of binary dependent variable models. The most obvious one is simply the percentage of correct predictions, typically referred to as the success ratio (SR). Formally, a signal forecast for the state of the economy y_t may be defined as

$$\hat{y}_t = \mathbf{1}(p_t > \xi), \tag{5}$$

where the conditional probability of recession p_t is obtained from a probit model, as defined in equation (2). If p_t is larger than a prespecified threshold ξ, we get a signal forecast $\hat{y}_t = 1$ (i.e. recession), and vice versa $\hat{y}_t = 0$ if $p_t \leq \xi$.

In this paper, we employ the threshold $\xi = 0.5$ for SR, which can be seen as natural threshold in (5). However, this is not a fully objective selection, and in some previous studies lower values for ξ have also been used (see, e.g. Nyberg [2010]). The success ratio is certainly important from the practical forecasters point of view, as it is a simple and easily interpretible goodness-of-fit measure. However, as recession periods are uncommon compared to expansion periods, the success ratios of relatively uninformative models might turn out high. To test whether the value of the success ratio is higher than that obtained when the realized values y_t and the forecasts $\hat{y}_t$ are independent, we employ the predictability test (PT) of Pesaran and Timmermann [2009].

The Receiver Operating Characteristic (ROC) curve is an alternative method to assess the goodness-of-fit of binary dependent variable models, and it has recently gained popularity in economic applications (see, e.g., Berge and Jorda [2011]; Schularick and Taylor [2012]; Christiansen et al. [2014]). The ROC curve is a mapping of the true positive rate

$$TP(\xi) = P_{t-1}(p_t > \xi | y_t = 1) \tag{6}$$

and the false positive rate

$$FP(\xi) = P_{t-1}(p_t > \xi | y_t = 0), \tag{7}$$

for all possible thresholds $0 \leq \xi \leq 1$, described as an increasing function in $[0, 1] \times [0, 1]$ space, with $TP(\xi)$ plotted on the Y-axis and $FP(\xi)$ on the X-axis. A ROC curve above the 45-degree line indicates forecast accuracy superior to a coin toss. Given that it takes into account all possible thresholds ξ, the ROC curve is a more robust method to evaluate the goodness-of-fit of a model than the success ratio.

The information in the ROC curve is typically summarized by the area under the ROC curve (AUC), which is simply the integral of the ROC curve between zero and one. Therefore, the AUC also gets values between 0 and 1, with the value of 0.5 corresponding a coin toss and the value 1 to a perfect forecast. Any improvement over the AUC=0.5 indicates statistical predictability. We test the null hypothesis of AUC= 0.5 implying no predictability using standard techniques [see Hanley and McNeil, 1982].

Another commonly used measure is the pseudo-R^2 of Estrella [1998], which is a counterpart of the coefficient of determination (R^2) designed for binary response models. The measure is defined as

$$psR^2 = 1 - \Big(\frac{logL_u}{logL_c}\Big)^{-(2/T)logL_c}, \tag{8}$$

where $logL_u$ and $logL_c$ are the maximum values of the constrained and unconstrained log-likelihood functions respectively, and T is the sample size. This measure takes on values between 0 and 1, and can be interpreted in the same way as the coefficient of determination in the usual linear predictive regression model. In Section 4, we also report the adjusted form of (8) (see Estrella [1998]) that takes into account the trade-off between improvement in model fit and the use of additional estimated parameters. Finally, for model selection purposes, we also report the Bayesian information criterion (BIC).

3 Data

In this Section we discuss the data employed in this paper, including the business cycle chronology and the predictive variables. The employed data is in quarterly frequency and the sample ranges from the first quarter of 1988 to the end of 2017.

3.1 The Finnish Business Cycle

One of the key issues in terms of data is the selection of business cycle chronology, as defined in equation (1). Unlike in the U.S., where the National Bureau of Economic Research (NBER) determines the official turning points[2], in Finland there is no such official chronology of recessions and expansions. However, there are a number of ways to determine the turning points based on data. In this paper, we define the turning points for Finnish business cycles using the Bry-Boschan algorithm (Bry and Boschan [1971]), which is a commonly used method in the literature. The dating is based on the algorithm used for seasonally adjusted real GDP data for the period 1988Q1–2017Q4.[3] The turning point chronology is given in Table I.

Table I: Turning points for the Finnish GDP

Peaks	Troughs
1990Q1	1993Q2
2007Q4	2009Q2
2012Q1	2013Q1
2013Q3	2015Q1

There are several interesting observations one can make from the chronology. First of all, the severe recession in the 1990's lasted three years based on this specification, ending only in mid 1993. Second, the period following the tech bubble in the early 2000's is not classified as a recession based on the chronology[4].

[2] http://www.nber.org/cycles/

[3] Alternatively, the dating could be based on monthly indicators of production, such as the trend indicator of output. In this exercise we used the latest available vintage of the GDP series in April 2018.

[4] Alternative dating methods might identify a recession period in the early 2000's, as discussed in Lanne and Nyberg [2009].

From a classical business cycle point of view, this was not a recession, as output did not decline, although from a growth cycle perspective the period could be viewed as a downturn. Third, the period between 2012Q2–2015Q1 includes two recessions based on this classification. In 2013 there were two quarters of expansion followed by another recession.

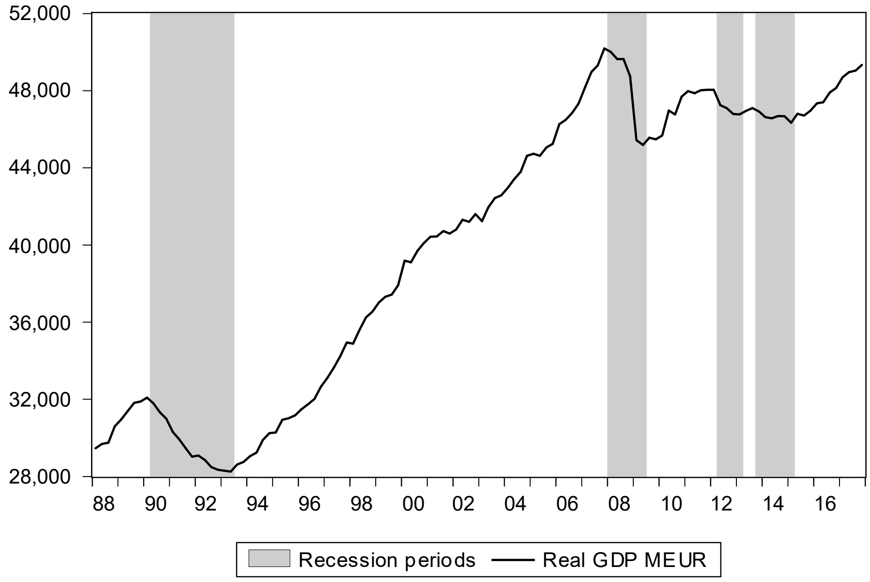

Figure 1: Finnish recession periods determined with the Bry-Boschan algorithm

3.2 Predictive Variables

As discussed in the introduction, the previous literature has suggested a number of leading indicators of recession periods for the US and other major economies. In this study, we focus on a set of variables based on previous literature. A common finding has been that an inverted yield curve, i.e. a situation where short-term yields are higher than long term yields, has a tendency to precede recessions. Theoretically this link between can be explained by the expectations hypothesis of interest rates (see, e.g. Estrella [2005]), as expected declines in future short rates would tend to decrease long-term interest rates and in extreme situations invert the yield curve. Based on the previous findings of e.g. Estrella and Mishkin [1998] and Nyberg [2010], it is natural to include the term spread (TS) as a predictor in our analysis.

Declines in asset markets are often associated with economic downturns. Equity returns, in particular, have commonly been used as predictors in the literature. Equity prices reflect discounted values of expected future dividends, and are therefore forward-looking in nature. Apart from financial assets, household wealth consists mainly of housing assets, and they typically have a higher share of the household wealth than financial assets. In the Finnish case, in 2013, the share of housing wealth relative to the total assets of households was just under 69%, whereas the the share of financial wealth was 21%[5]. Given that the sources of the 2007 great recession were in a housing bubble, recent studies have also included housing market variables as predictors (see, e.g. Christiansen et al. [2017] and references therein) of US recession periods. In our study we include the logarithmic first difference of the Finnish stock market index (RET) and the first difference in a real house price index of the capital area (HPI) as predictors.

Sentiment variables, such as consumer confidence indices, are a particularly interesting group of variables. They are often designed for the particular purpose to convey expectations of current and future economic developments. Christiansen et al. [2014] found that that the consumer confidence and purchasing managers' indices are useful predictors of US recession periods, even after controlling for the information contained in classic recession predictors and common factors based on a large panel of economic and financial variables. Based on these findings, we include the consumer confidence index (CCI) and the component related to the risk of unemployment (CUE) as predictors in our set of variables.

In Table II, we present the correlation coefficients between the employed variables. The correlation between the term spread and the real house price variable is rather high (0.616), as is the correlations between the real house price growth and the consumer confidence (0.623). The risk of unemployment component of the consumer confidence index is negatively correlated with the consumer confidence index (-0.609), as one would expect.

Table II: Correlations between employed variables

	TS_t	RET_t	HPI_t	CCI_t	CUE_t
TS_t	1	0.313	0.616	0.508	-0.286
RET_t		1	0.423	0.262	-0.257
HPI_t			1	0.623	-0.279
CCI_t				1	-0.609
CUE_t					1

Notes: This table presents the correlation coefficients between the employed variables.

4 Empirical findings

In this section, we present the empirical findings of the study. Following the typical convention, we first present in-sample findings from estimated models. After this, we test the findings in a pseudo out-of-sample setting.

4.1 Estimation results for single-predictor probit models

We start off by presenting findings based on single-predictor models in Table III. Previous research has suggested that the predictive ability of different leading

[5]Source: Official Statistics of Finland (OSF): Households' assets [e-publication]. ISSN=2242-3230. 2013. Helsinki: Statistics Finland [referred: 24.5.2018]. Access method: http://www.stat.fi/til/vtutk/2013/vtutk_2013_2015-04-01_tie_001_en.html

indicators vary between different forecasting horizons (or lag lengths). Especially the term spread has shown predictive ability for future downturns with longer lags. Therefore, in Table III, we present the findings for each variable using lags of one to four quarters.

Table III: In-sample results for single-predictor probit models

	Variable	Coeff.	adj.psR^2	BIC	SR	AUC
First lags						
1	TS_{t-1}	-0.624***	0.252	54.192	0.861**	0.796***
2	RET_{t-1}	-2.852**	0.047	66.019	0.739	0.666***
3	HPI_{t-1}	-0.144***	0.326	49.851	0.826	0.854***
4	CCI_{t-1}	-0.100***	0.383	46.413	0.835**	0.889***
5	CUE_{t-1}	0.044***	0.237	55.073	0.783	0.809***
Second lags						
6	TS_{t-2}	-0.865***	0.362	47.559	0.877***	0.834***
7	RET_{t-2}	-3.452***	0.074	64.227	0.728	0.703***
8	HPI_{t-2}	-0.162***	0.368	47.211	0.860***	0.845***
9	CCI_{t-2}	-0.087***	0.315	50.322	0.825**	0.849***
10	CUE_{t-2}	0.032***	0.142	60.334	0.746	0.753***
Third lags						
11	TS_{t-3}	-0.825***	0.347	48.284	0.858	0.829***
12	RET_{t-3}	-3.281***	0.066	64.378	0.735	0.696***
13	HPI_{t-3}	-0.137***	0.304	50.798	0.841	0.827***
14	CCI_{t-3}	-0.070***	0.218	55.786	0.796	0.799***
15	CUE_{t-3}	0.021*	0.054	65.085	0.743*	0.681***
Fourth lags						
16	TS_{t-4}	-0.632***	0.264	52.921	0.821	0.806***
17	RET_{t-4}	-2.139*	0.019	66.722	0.723	0.635**
18	HPI_{t-4}	-0.112	0.228	55.022	0.804	0.791***
19	CCI_{t-4}	-0.062***	0.175	57.983	0.786	0.769***
20	CUE_{t-4}	0.016	0.026	66.322	0.732	0.642**

Notes: This table presents the findings from single-predictor probit models for Finnish recessions. The goodness-of-fit measures are described in detail in Section 2. In the table, *, **, and *** denote the statistical significance of the estimated coefficients using robust standard errors, the Pesaran and Timmermann [2009] (PT) predictability test for the success ratio, and the AUC at 10%, 5% and 1% significance levels, respectively.

The findings for the single-predictor probit models indicate that the term spread, real house prices and consumer confidence measures are potentially useful predictors of Finnish recession periods. Overall, the highest success ratio is obtained using the first lag of the term spread (TS), whereas the best fit in terms of the adjusted pseudo-R^2 and the highest AUC are obtained with the model using the first lag of the consumer confidence index (CCI). As previous studies have found using US data, the term spread has predictive ability even with longer lags. Similar findings also hold for the real house price variable and the consumer confidence index.

The least favorable results are found for the stock return variable (RET). The estimated coefficients are statistically significant at the lags of 1 to 3 quarters, but the fit of the model and the classification ability remain modest at best. One reason for this may be that we use quarterly returns, which still contain a substantial amount of variation that may be regarded as noise. To test whether this may have an effect on the findings, we also experiment with models using 4-quarter returns. These findings are presented in Table IV, along those with the 4-quarter changes of the real house price index. The findings for the AUC illustrate that the 4-quarter stock returns are in fact more useful in classifying recessions. Similar findings are also made for the real house prices for the first two quarterly lags.

Table IV: In-sample results for single-predictor models using 4-quarter changes

	Variable	Coeff.	adj.psR^2	BIC	SR	AUC
First lags						
21	YRET$_{t-1}$	-1.891***	0.147	60.275	0.748	0.753***
22	YHPI$_{t-1}$	-0.063***	0.490	39.981	0.826*	0.902***
Second lags						
23	YRET$_{t-2}$	-1.651***	0.116	61.7895	0.711	0.736***
24	YHPI$_{t-2}$	-0.048***	0.391	45.826	0.833	0.870***
Third lags						
25	YRET$_{t-3}$	-1.381**	0.081	63.567	0.699	0.713***
26	YHPI$_{t-3}$	-0.034***	0.248	54.020	0.805	0.798***
Fourth lags						
27	YRET$_{t-4}$	-1.019*	0.038	65.675	0.714	0.667***
28	YHPI$_{t-4}$	-0.022***	0.120	61.085	0.804	0.717***

Notes: This table presents the findings from single-predictor probit models for Finnish recessions. For details on the goodness-of-fit measures, see Section 2. In the table, *, **, and *** denote the statistical significance of the estimated coefficients using robust standard errors, the Pesaran and Timmermann [2009] (PT) predictability test for the success ratio, and the AUC at 10%, 5% and 1% significance levels, respectively.

As an overall conclusion on the single-predictor models, the individual variables fair rather well in predicting recession periods in Finland. Some additional remarks may be made for the individual variables. First, the coefficients of each variable are of expected sign; higher term spread, stock returns, growth in real house prices, and consumer confidence are negatively associated with recession risk, whereas a higher risk of unemployment has a positive coefficient. The stock return and housing price variables generally give a false alarm at the beginning of the 2000's after the IT boom, when asset prices collapsed. The term spread classified the 1990's recession as well as the latter part of the 2008–2009 recession, but did not catch the recessions of 2011 and 2013. This finding is interesting, as it suggests that the classification power of the term spread has declined when the interest rates are close to the zero lower bound.

4.2 Findings from multi-predictor probit models

Although the findings from the single-predictor models were rather promising, it is not reasonable to expect that any single indicator would contain comprehensive information about the future state of the economy. In Table V, we present findings of selected multi-predictor probit models, based on combinations of the variables used in the previous section. We run all possible combinations of the seven variables used in the previous section, including up to four lags. For each variable, we allow for only one lag in our illustration. We present models with the strongest findings based on the AUC and SR using two, three and four predictors. In order to keep the models relatively parsimonious, we limit the maximum number of explanatory variables to four. Models 29, 31, and 33 are ones that yield the highest AUC:s, whereas models 30, 32, and 34 give the highest success ratios.

The results in Table V indicate that models including the second lag of the term spread and the first lag of the unemployment expectation component of the consumer confidence index produce the highest in-sample fit among the two-predictor models based on the adjusted pseudo-R^2 and the AUC (model 29). In the three variable case (model 31), the first lag of the real house price index is selected in addition to the aforementioned variables. This model improves over the two-predictor model in terms of the AUC, and yields a value of 0.952 (compared to 0.937). In the four-variable case (model 33), the first lag of the annual stock return variable (YRET) is selected into the model. It is noteworthy, that it would not be necessary for the same variables and lags to be selected when an additional variable is included, but this is how the results turn out in case of the AUC. The improvement of model 33 over model 31 is more modest, as the AUC increases to 0.955 from 0.952. However, the model fit, as measured by the adjusted pseudo-R^2, actually decreases slightly to 0.602 from 0.604, which implies potential overfitting.

Table V: Estimation results for in-sample predictive models

Variable	29	30	31	32	33	34
TS_{t-2}	-1.205***	-0.832***	-0.869**		-0.893**	
$YRET_{t-1}$					0.326	
HPI_{t-1}			-0.074	-0.134***	-0.078	-0.219***
$YHPI_{t-4}$		-0.018**				-0.035***
CCI_{t-3}				-0.033		0.034
CUE_{t-1}	0.066***		0.059***	0.037**	0.059***	0.040***
CONST	-0.286	0.145	-0.511	-0.792***	-0.501	-1.316***
psR^2	0.593	0.436	0.618	0.529	0.620	0.566
$adj.psR^2$	0.582	0.421	0.604	0.512	0.602	0.546
BIC	36.272	45.887	37.025	42.579	39.277	42.646
SR	0.875***	0.893***	0.884***	0.902***	0.875***	0.920***
AUC	0.937***	0.862***	0.952***	0.923***	0.955***	0.925***

Notes: This table presents the findings from two-, three-, and four-predictor probit models for Finnish recessions. In the table, *, **, and *** denote the statistical significance of the estimated coefficients using robust standard errors, the Pesaran and Timmermann [2009] (PT) predictability test for the success ratio, and the AUC at 10%, 5% and 1% significance levels, respectively.

In Table V, models 30, 32, and 34 yield the highest success ratios based on the 50% probability threshold ($\xi = 0.5$). Among the two-predictor models, model 30, including the second lag of the term spread and the first lag of the annual change in the real house price index, yields the highest success ratio (0.893). In contrast to the models based on the AUC, neither of the variables in model 30 are included in three-predictor model (32) that yields the highest success ratio (0.902) among three-predictor models. Finally, the four-predictor model (34) gives a success ratio of 0.920, improving over the two- and three-predictor models.

4.3 Findings from autoregressive probit models

We extend the in-sample analysis by employing autoregressive probit models (cf. equation (4)). These findings are presented in Table VI. Similarly to model 30 in the case of static probit models, the best performing two-predictor model (35) based on the AUC includes the second lag of TS and the first lag of CUE. The coefficient for the autoregressive term (π_{t-1}) is not statistically significant, but the AUC (0.939) is slightly higher than for the static probit model (0.937) in Table V. Similarly, the best performing two-predictor autoregressive model (36) in terms of the success ratio outperforms the static counterpart (model 31), by yielding a success ratio of 0.902 (vs. 0.893).

Table VI: Estimation results for autoregressive probit models

Variable	35	36	37	38	39	40
TS_{t-1}		-0.515***				-0.830***
TS_{t-2}	-1.224***				-0.444	
TS_{t-3}			-0.927**			
$YRET_{t-1}$		-0.664				
HPI_{t-1}				-0.209***	-0.093	
HPI_{t-2}			-0.130***			
$YHPI_{t-3}$					0.083	0.103***
CCI_{t-1}					-0.153	-0.219***
CCI_{t-3}				0.021		
CUE_{t-1}	0.063***		0.064***			
CUE_{t-2}				0.027**		
CUE_{t-4}						0.012*
π_{t-1}	0.093	0.472***	-0.279*	0.510***	0.755	0.726***
CONST	-0.159	-0.172	-0.665	-0.785***	1.580	2.334***
psR^2	0.614	0.438	0.617	0.635	0.712	0.698
adj.psR^2	0.600	0.417	0.599	0.618	0.695	0.681
BIC	37.304	48.146	39.486	38.333	35.751	36.644
SR	0.875***	0.902***	0.866***	0.911***	0.911***	0.920***
AUC	0.939***	0.849***	0.946***	0.939***	0.956***	0.951***

Notes: This table presents the findings from two-, three-, and four-predictor autoregressive probit models for Finnish recessions. In the table, *, **, and *** denote the statistical significance of the estimated coefficients, the Pesaran and Timmermann [2009] (PT) predictability test for the success ratio, and the AUC at 10%, 5% and 1% significance levels, respectively.

4.4 Out-of-sample findings

As previous forecasting literature has shown, good in-sample fit does not necessarily imply good out-of-sample performance. Therefore, in this section, we will examine the pseudo out-of-sample forecasting performance of our models. We use an expanding window forecasting approach with estimation samples ranging from 1990Q1–1995Q4 to 1990Q1–2017Q3 and report the results of one- to four-quarter ahead forecasting horizons.

Similarly to the previous section, we first present the findings based on single-predictor models (Table VII). Among these models, the ones including the term spread perform the best based on the success ratio in all four forecast horizons, and also based on the AUC in three-and-four-quarter horizons. This finding is in line with previous literature that emphasizes the role of the inverted yield curve as a recession predictor (Estrella and Hardouvelis [1991] and subsequent literature).

Table VII: Out-of-sample results for single-predictor models

Forecast horizon: 1 quarter				
Model	Variable	oos.psR^2	SR	AUC
1	TS_{t-1}	0.177	0.852	0.681**
2	RET_{t-1}	0.023	0.772	0.539
21	YRET_{t-1}	Neg.	0.761	0.648**
3	HPI_{t-1}	0.329	0.830	0.811***
22	YHPI_{t-1}	0.218	0.841**	0.852***
4	CCI_{t-1}	0.232	0.841**	0.886***
5	CUE_{t-1}	0.197	0.750	0.751***
Forecast horizon: 2 quarters				
6	TS_{t-2}	0.134	0.875***	0.775***
7	RET_{t-2}	Neg.	0.727	0.557
23	YRET_{t-2}	Neg.	0.739	0.595
8	HPI_{t-2}	0.251	0.841***	0.808***
24	YHPI_{t-2}	0.044	0.818	0.801***
9	CCI_{t-2}	Neg.	0.830	0.806***
10	CUE_{t-2}	0.083	0.716**	0.640**
Forecast horizon: 3 quarters				
11	TS_{t-3}	0.089	0.852	0.756***
12	RET_{t-3}	Neg.	0.693	0.523
25	YRET_{t-3}	Neg.	0.705*	0.543
13	HPI_{t-3}	0.053	0.807	0.728***
26	YHPI_{t-3}	0.156	0.795***	0.659**
14	CCI_{t-3}	Neg.	0.807***	0.709***
15	CUE_{t-3}	Neg.	0.727***	0.480
Forecast horizon: 4 quarters				
16	TS_{t-4}	0.166	0.830	0.661**
17	RET_{t-4}	Neg.	0.693	0.441
27	YRET_{t-4}	Neg.	0.670	0.454
18	HPI_{t-4}	0.009	0.773**	0.648**
28	YHPI_{t-4}	0.112	0.773	0.493
19	CCI_{t-4}	Neg.	0.807	0.638**
20	CUE_{t-4}	Neg.	0.750	0.389

Notes: This table presents the one- to four-quarter ahead forecasting results from static probit models for Finnish recession periods for 1996Q1-2017Q4. Model numbers refer to those used in Table III and IV.

The out-of-sample findings for multi-predictor static probit models are presented in Table VIII. The findings indicate that model 31, including TS, CUE and CCI as predictors, yields the highest AUC (0.912) in the one-quarter ahead forecasts (and overall). Similarly, model 32 yields the highest success ratio (0.864) in the one-quarter ahead forecasts. Both models include three predictors and were selected based on the in-sample performance of the corresponding measures they yield best results for also in the out-of-sample exercise.

In the two-quarter ahead forecasts, the highest AUC is given by model 34

(0.845) and the highest success ratios by models 29 and 32 (0.841). It should be noted, that the model based only on the term spread (see Table VII) gave a higher success ratio (0.875) than the aforementioned models. Model 29 yields the highest success ratios among the studied models for the 3- and 4-quarter ahead forecasts, whereas the highest AUC:s are given by models 30 and 34, respectively.

Table VIII: Out-of-sample results from static probit models

Forecast horizon: 1 quarter						
Model	29	30	31	32	33	34
oos.psR^2	0.423	0.183	0.462	0.009	0.445	Neg.
SR	0.852**	0.841*	0.852**	0.864*	0.852**	0.841
AUC	0.878***	0.829***	0.912***	0.886***	0.901***	0.867***
Forecast horizon: 2 quarters						
oos.psR^2	0.323	0.124	0.122	0.127	0.080	0.198
SR	0.841	0.830	0.739	0.841	0.739	0.807
AUC	0.832***	0.805***	0.782***	0.826***	0.780***	0.845***
Forecast horizon: 3 quarters						
oos.psR^2	0.004	0.103	Neg.	Neg.	Neg.	Neg.
SR	0.852*	0.841	0.841	0.818	0.830	0.818
AUC	0.731***	0.766***	0.723***	0.760***	0.733***	0.753***
Forecast horizon: 4 quarters						
oos.psR^2	0.054	0.144	Neg.	Neg.	Neg.	Neg.
SR	0.830	0.830	0.784***	0.807	0.750	0.807
AUC	0.602	0.652**	0.583	0.672**	0.587	0.691***

Notes: This table presents the one- to four-quarter ahead forecasting results from static probit models for Finnish recession periods for 1996Q1-2017Q4. Model numbers refer to those used in Table V.

The out-of-sample findings from multi-predictor autoregressive probit models are presented in Table IX[6]. There is little variation in the success ratios between most models in the one-period ahead forecasts, and the highest success ratio (0.852) is yielded by models 35 and 39. Model 39 produces the highest AUC (0.905). Both the highest success ratio and AUC among the autoregressive models are actually slightly lower than for static models presented in Table VIII. However, this finding changes when we move to longer forecasting horizons. The highest success ratio among all forecasting horizons, 0.875, is obtained using model 36 for the two-quarter ahead forecasts.

5 Conclusions

This paper has presented an empirical application of probit models in forecasting the probability of recessions in Finland. Our in-sample findings indicate that from our set of predictive variables, it turns out that real house price indices are the most useful individual predictors of Finnish recession periods at short lags. The term spread and other classic recession predictors have also significant classification

[6]The out-of-sample findings for the single-predictor autoregressive models are included in the appendix.

Table IX: Out-of-sample results from autoregressive probit models

Forecast horizon: 1 quarter						
Model	35	36	37	38	39	40
oos.psR^2	0.455	0.274	Neg.	0.397	0.469	0.433
SR	0.852**	0.852**	0.750	0.841	0.852	0.830
AUC	0.887***	0.758***	0.720***	0.867***	0.905***	0.903***
Forecast horizon: 2 quarters						
oos.psR^2	0.281	0.149	Neg.	Neg.	Neg.	0.292
SR	0.818	0.875***	0.807	0.841	0.852	0.807
AUC	0.805***	0.778***	0.663***	0.846***	0.815***	0.858***
Forecast horizon: 3 quarters						
oos.psR^2	Neg.	Neg.	Neg.	Neg.	Neg.	Neg.
SR	0.852**	0.796	0.841*	0.773	0.830	0.784
AUC	0.734***	0.714***	0.727***	0.771***	0.705***	0.843***
Forecast horizon: 4 quarters						
oos.psR^2	Neg.	Neg.	Neg.	Neg.	Neg.	Neg.
SR	0.784	0.693	0.773**	0.807	0.795	0.830
AUC	0.539	0.490	0.615**	0.661***	0.660***	0.763***

Notes: This table presents the one- to four-quarter ahead forecasting results from autoregressive probit models for Finnish recession periods for 1996Q1-2017Q4. Model numbers refer to those used in Table VI.

ability. While using higher lags of variables, the term spread becomes the strongest predictor of future recessions, a finding that is in line with previous literature.

Multi-predictor models improve the in-sample fit of the models. A model that combines the term spread, the unemployment expectation component of the consumer confidence index, the consumer confidence index, and the real house price variable yields the strongest in-sample results based on the area under the receiver operating characteristic curve (AUC).

The out-of-sample results generally confirm the in-sample findings, although the strongest findings are obtained for slightly more parsimonious models than in the in-sample case, as is commonly found in forecasting applications. Furthermore, the autoregressive specification of the probit model generally yields higher values of the AUC than the static probit model in the in-sample estimations, but findings from pseudo out-of-sample forecasts give mixed results between different forecasting horizons.

The purpose of this study has been illustrative. The idea has been to test whether previous findings in from other countries also hold for the Finnish business cycle, and that goal has been achieved. However, there are many ways to extend the analysis in this paper. Firstly, the use of a larger set of variables could be considered. Related to this, it would be interesting to see how recently developed models, designed specifically for data rich environments, would compare in performance to the more traditional probit models on Finnish data. Further empirical work could also consist of studies focusing on spillover effects of recession probabilities from large countries to small open economies (such as Finland) using bivariate probit models (see, e.g., Nyberg [2014]).

References

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

T.J. Berge. Predicting recessions with leading indicators: Model averaging and selection over the business cycle. *Journal of Forecasting*, 34:455–471, 2015.

T.J. Berge and O. Jorda. Evaluating the classification of economic activity into recessions and expansions. *American Economic Journal: Macroeconomics*, 3: 246–277, 2011.

H. Bernard and S. Gerlach. Does the term structure predict recessions? The international evidence. *International Journal of Finance and Economics*, 3:195–215, 1998.

G. Bry and C. Boschan. *Evaluating the Classification of Economic Activity into Recessions and Expansions.* NBER Books. National Bureau of Economic Research, Inc, 1971.

M. Chauvet and S. Potter. Forecasting recessions using the yield curve. *Journal of Forecasting*, 24:77–103, 2005.

Z. Chen, A. Iqbal, and H. Lai. Forecasting the probability of US recessions: A probit and dynamic factor modelling approach. *Canadian Journal of Economics*, 44:651–672, 2011.

C. Christiansen, J.N. Eriksen, and S.T. Møller. Forecasting US recessions: The role of sentiment. *Journal of Banking and Finance*, 49:459–468, 2014.

C. Christiansen, J.N. Eriksen, and S.T. Møller. Metro area common house price declines and US recessions. *Available at SSRN*, 2017.

R.M. de Jong and T. Woutersen. Dynamic time series binary choice. *Econometric Theory*, 27:673–702, 2011.

A. Estrella. A new measure of fit for equations with dichotomous dependent variables. *Journal of Business and Economic Statistics*, 16:198–205, 1998.

A. Estrella. Why does the yield curve predict output and inflation? *Economic Journal*, 115:722–744, 2005.

A. Estrella and G.A. Hardouvelis. The term structure as a predictor of real economic activity. *Journal of Finance*, 46:555–576, 1991.

A. Estrella and F.S. Mishkin. Predicting U.S. recessions: Financial variables as leading indicators. *Review of Economics and Statistics*, 80:45–61, 1998.

P. Fornaro. Forecasting US recessions with a large set of predictors. *Journal of Forecasting*, 35:477–492, 2016.

J.A. Hanley and B.J. McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143:29–36, 1982.

H. Kauppi and P. Saikkonen. Predicting U.S. recessions with dynamic binary response models. *Review of Economics and Statistics*, 90:777–791, 2008.

M. Lanne and H. Nyberg. Suomen kansantalouden suhdanneindeksi. *Kansantaloudellinen aikakauskirja*, 105:422–432, 2009.

E.C.Y. Ng. Forecasting US recessions with various risk factors and dynamic probit models. *Journal of Macroeconomics*, 34:112–125, 2012.

H. Nyberg. Dynamic probit models and financial variables in recession forecasting. *Journal of Forecasting*, 29:215–230, 2010.

H. Nyberg. A bivariate autoregressive probit model: Business cycle linkages and transmission of recession probabilities. *Macroeconomic Dynamics*, 18:838–862, 2014.

M.H. Pesaran and A. Timmermann. Testing dependence among serially correlated multi-category variables. *Journal of the American Statistical Association*, 485: 325–337, 2009.

H. Pönkä. The role of credit in predicting U.S. recessions. *Journal of Forecasting*, 36:221–247, 2017.

M. Schularick and A.M. Taylor. Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008. *American Economic Review*, 102:1029–1061, 2012.

M. Sensier, M. Artis, D.R. Osborn, and C. Birchenhall. Domestic and international influences on business cycle regimes in Europe. *International Journal of Forecasting*, 20:343–357, 2004.

J.H. Stock and M.W. Watson. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97: 1167–1197, 2002.

6 Appendix

Table X: Out-of-sample results for autoregressive single-predictor models

Model	Variable	oos.psR^2	SR	AUC
Forecast horizon: 1 quarter				
1	TS_{t-1}	0.277	0.841	0.756***
2	RET_{t-1}	Neg.	0.614	0.668***
21	$YRET_{t-1}$	Neg.	0.727	0.676***
3	HPI_{t-1}	0.253	0.772	0.843***
22	$YHPI_{t-1}$	0.132	0.841**	0.871***
4	CCI_{t-1}	0.044	0.841**	0.893***
5	CUE_{t-1}	0.131	0.716	0.723***
Forecast horizon: 2 quarters				
6	TS_{t-2}	0.206	0.875***	0.777***
7	RET_{t-2}	Neg.	0.648	0.663***
23	$YRET_{t-2}$	Neg.	0.693	0.569
8	HPI_{t-2}	0.315	0.818	0.810***
24	$YHPI_{t-2}$	0.127	0.830	0.793***
9	CCI_{t-2}	0.050	0.829	0.819***
10	CUE_{t-2}	0.071	0.727	0.631**
Forecast horizon: 3 quarters				
11	TS_{t-3}	Neg.	0.841*	0.724***
12	RET_{t-3}	Neg.	0.670	0.599*
25	$YRET_{t-3}$	Neg.	0.659	0.525
13	HPI_{t-3}	Neg.	0.807	0.774***
26	$YHPI_{t-3}$	0.146	0.807***	0.667***
14	CCI_{t-3}	Neg.	0.807***	0.718***
15	CUE_{t-3}	Neg.	0.602	0.391
Forecast horizon: 4 quarters				
16	TS_{t-4}	0.148	0.818	0.670***
17	RET_{t-4}	Neg.	0.602	0.521
27	$YRET_{t-4}$	Neg.	0.625	0.333
18	HPI_{t-4}	Neg.	0.796***	0.659***
28	$YHPI_{t-4}$	0.112	0.773	0.493
19	CCI_{t-4}	Neg.	0.807	0.648
20	CUE_{t-4}	Neg.	0.602	0.340

Notes: This table presents the one- to four-quarter ahead forecasting results from static probit models for Finnish recession periods for 1996Q1-2017Q4. Model numbers refer to those used in Table III and IV.