Bank of Finland

BoF Economics Review

4 • 2022

A model for predicting Finnish household loan stocks

Juho Nyholm
Aino Silvo

Abstract

We propose a new Bayesian VAR model for forecasting household loan stocks in Finland. The model is designed to work as a satellite model of a larger DSGE model for the Finnish economy, the Aino 2.0 model. The forecasts produced with the BVAR model can be conditioned on projections of several macro variables obtained from the Aino 2.0 model. We study several specifications for the set of variables and lags included in the BVAR, and evaluate their out-of-sample forecast accuracy with root mean squared forecasting errors (RMSFEs). We then select a preferred specification that performs best in predicting the loan stocks over forecast horizons ranging from one to twelve quarters ahead. The model adds to the existing toolkit of forecast models currently in use at the Bank of Finland and improves our understanding of household debt trends in Finland.

Keywords: household debt, Bayesian estimation, conditional forecasting

JEL codes: C11, C32, E37

We thank Adam Gulan and seminar participants at the Bank of Finland for valuable comments.

Juho Nyholm, Economist, Macroprudential Analysis Division, Bank of Finland, juho.nyholm@bof.fi, jtnyholm@gmail.com
Aino Silvo, Senior Economist, Forecasting Division, Bank of Finland, aino.silvo@bof.fi

A model for predicting Finnish household loan stocks*

Juho Nyholm†
Aino Silvo‡

Abstract

We propose a new Bayesian VAR model for forecasting household loan stocks in Finland. The model is designed to work as a satellite model of a larger DSGE model for the Finnish economy, the Aino 2.0 model. The forecasts produced with the BVAR model can be conditioned on projections of several macro variables obtained from the Aino 2.0 model. We study several specifications for the set of variables and lags included in the BVAR, and evaluate their out-of-sample forecast accuracy with root mean squared forecasting errors (RMSFEs). We then select a preferred specification that performs best in predicting the loan stocks over forecast horizons ranging from one to twelve quarters ahead. The model adds to the existing toolkit of forecast models currently in use at the Bank of Finland and improves our understanding of household debt trends in Finland.

JEL classification: C11, C32, E37

Keywords: *household debt, Bayesian estimation, conditional forecasting*

---

*We thank Adam Gulan and seminar participants at the Bank of Finland for valuable comments.

†Economist, Macroprudential Analysis Division, Bank of Finland. juho.nyholm@bof.fi, jtnyholm@gmail.com.

‡Senior Economist, Forecasting Division, Bank of Finland. aino.silvo@bof.fi.

1 Introduction

We introduce a Bayesian vector autoregressive (BVAR) model for forecasting household loan stocks in Finland. The model is specified such that the forecasts can be conditioned on projections of several macroeconomic variables obtained from a large-scale macroeconomic DSGE model, the Aino 2.0 model of the Finnish economy (Kilponen, Orjasniemi, Ripatti, and Verona, 2016). The BVAR model is designed to be used as a satellite model extending the Aino 2.0 model that does not contain household loans, but it can also serve as a stand-alone projection model.

In the balance sheet statistics of Monetary Financial Institutions (MFIs) published by the Eurosystem, household loans are classified in three categories: loans for house purchase, credit for consumption, and other loans. Our interest here focuses on producing projections for the three loan stock categories used in the Eurosystem staff Broad Macroeconomic Projection Exercises (BMPE) and in the context of macroprudential policy analyses conducted at the Bank of Finland. The projection horizon of interest is three years or 12 quarters ahead, which roughly corresponds to the BMPE forecast horizon.

In the estimation, we use a Minnesota prior of the Normal-Inverse-Wishart form, an analytically convenient conjugate prior. The model is estimated using the optimal prior shrinkage algorithm proposed by Giannone, Lenza, and Primiceri (2015). This procedure is based on a hierarchical Bayesian approach, where hyperparameters of the prior are chosen to maximize the marginal likelihood of the model. This coincides with maximizing the one-step ahead forecasting performance of the model, which is well suited for our purposes. The algorithm is particularly useful when the dimensionality of the model is large and, as in our case, there is a limited number of data observations.

We then compare the forecasting performance of the model estimated using the Giannone et al. (2015) algorithm with commonly used rule-of-thumb hyperparameter values suggested in the literature for the Minnesota prior. The Giannone et al. (2015) algorithm is designed to minimize one-step ahead forecasting errors, but we find that it also consistently outperforms estimating the models with rule-of-thumb hyperparameter values for the Minnesota prior at longer forecasting horizons.

In specifying the model, we focus on a set of variables for which projections can be obtained from the Aino 2.0 model and that we expect to be useful for forecasting household loan stocks based on macroeconomic theory. This choice is motivated by the nature of our projection model, designed to work as a satellite model for the Aino 2.0 model. These variables are the euro area short-term interest rate (3-month Euribor), the Finnish real GDP, the Finnish Harmonized Consumer Price Index, nominal household disposable income, the nominal house price index, the net saving rate of households, and the stock of corporate loans issued by Finnish monetary financial institutions (MFIs).

These core variables are then complemented with other variables that we judge are potentially useful for predicting household loan stocks: the average lending spread for Finnish household loans, a credit condition index calculated using information from the Finnish Bank Lending Survey (BLS) and the CISS index of financial stress for the euro area. All variables are listed in Table 2 in the Appendix.

All models are estimated in (log-)levels over the quarterly data sample 2003Q1–2021Q4. We estimate the model over an expanding window starting from 2003Q1 and up to 2018Q4, and produce both unconditional and conditional forecasts for the three household loan stock variables at each sub-sample. The conditioning variables are listed in Table 2.

Model selection is based on evaluating out-of-sample root mean squared forecast errors (RMSFE) of different model specifications at horizons one, four, eight, and twelve quarters ahead. We compare different sets of endogenous variables and models estimated with different lag lengths of up to $p = 4$.

We find that a BVAR model with ten variables (including the three household loan stock variables, GDP volume, price level, the euro area short-term interest rate, loans to non-financial corporations, household disposable income, lending spread, and the CISS stress index) estimated with two lags ($p = 2$) performs best overall at various projection horizons for both unconditional and conditional forecasts. Model specifications are listed in Table 1. Our preferred model is Model 9 with $p = 2$. In addition, we find that conditioning on the observed core variables substantially improves on forecasting accuracy for our selected model.

In the following sections, we first describe the data and our estimation methodology in greater detail. We then describe our model selection criteria and discuss the forecasting performance of various specifications. The final section concludes.

2 Debt and savings trends of Finnish households

Loans to households issued by MFIs can be classified into three categories: loans for house purchase, consumer credit, and other loans. The vast majority of Finnish household loans consists of house purchase loans. This stock amounted to €108 billion at the end of 2021, accounting for about three-quarters of the total household loan stock (Figure 1, upper right panel). Consumer credit consists of both collateralized and non-collateralized consumer credit, including credit card debt and overdrafts. Other loans include, among others, student loans as well as loans to sole proprietors. At the end of 2021, the stocks of consumer credit and other loans were €17 billion and €19 billion, respectively. The total stocks of loans to households issued by Finnish MFIs amounted to €143 billion.

Finland experienced a deep recession in 1991–1993. The growth rates of loan stocks turned negative for several years as the economy experienced a dramatic deleveraging episode following the credit boom of the late 1980s (Figure 1, lower panel). Real GDP growth started to recover in 1994. Loan growth picked up speed again in the late 1990s, a few years after the start of the economic recovery from the depression. This coincides with an economic boom that lasted until the outbreak of the global financial

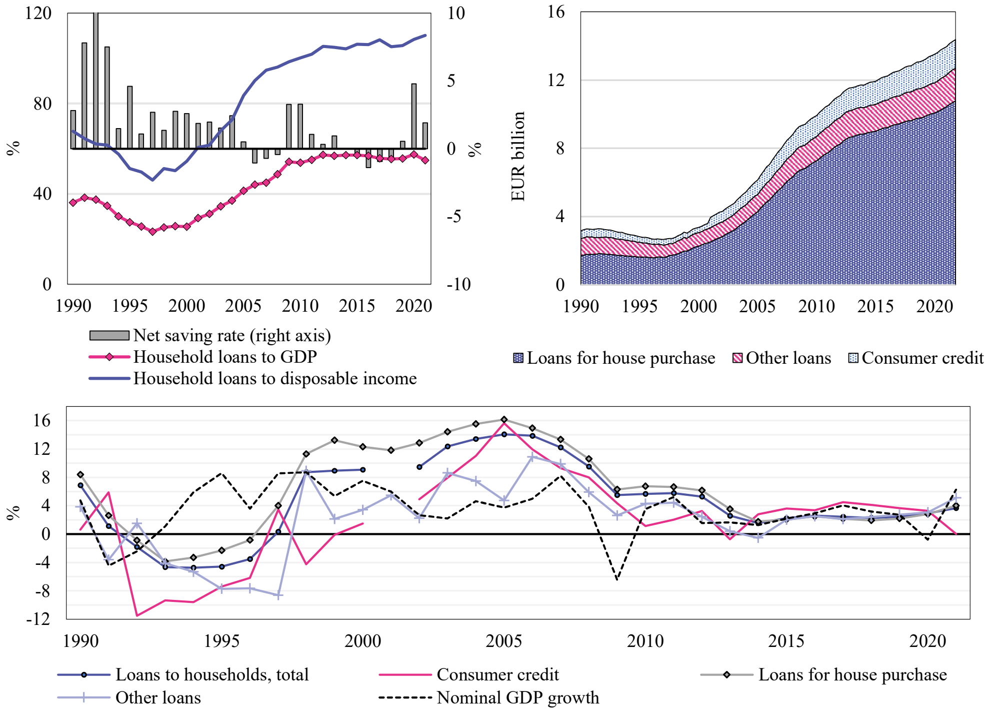

Figure 1: Top left: Evolution of total household loan stock relative to nominal disposable income and nominal GDP (left axis), and the net household saving rate (right axis). Top right: Evolution of consumer credit, loans for house purchase, and other loans (in euros). Bottom: Growth rates of the three components of the household loan stock and the growth rate of nominal GDP.

crisis in 2008. In the late 1990s, Finland recorded its highest GDP growth rates in three decades.

The loans-to-GDP ratio recovered to its 1990 level in 2004 (Figure 1, upper left panel). The fast growth rate of household loan stocks lasted until the aftermath of the global financial crisis and the euro crisis that followed. The economic downturn started in 2008 in Finland, and in 2009, Finnish GDP contracted by 8%. The accompanying decline in loan growth was gradual and lasted longer, but Finnish households did not de-lever in this crisis. Instead, slow growth in household loan stocks continued through the recession. In the 15 years since the global financial crisis, growth rates of loan stocks have been moderate compared to the boom of the early 2000s.

Over the past decades, loans for house purchase have grown the most in relative terms. Between 1990 and 2021, the stock of loans for house purchase grew six-fold in nominal terms, while the stock of other loans doubled and the stock of consumer credit grew less than four-fold. Following the long episode of household de-leveraging, in the early 2000s, household loan growth picked up again. In our estimation sample running from 2003 to 2021, the stock of loans for house purchase more than triples, while the stocks of consumer credit and other loans doubles.

The household indebtedness ratio can be defined as the total stock of household loans over a four-quarter moving sum of household nominal disposable income. As seen in the upper left panel of Figure 1, the ratio reached a record of 110% at the end of 2021. At the same time, the stock of household loans was 56% relative to GDP. Since early 1990s, this corresponds to a growth of 55% and 62% in these ratios, respectively.

In recent decades, Finnish households have taken on more debt and their saving rate has declined. The saving rate is defined as the ratio of savings (disposable income less consumption expenditures) to disposable income. The net saving rate peaked historically during the Finnish depression in 1992 at 10%. This peak in the saving rate was soon dampened. Between 1996 and 2005, the saving rate was rather stable at around 2%. This coincides with the period of the fastest growth in household indebtedness. The global financial crisis saw another peak in the saving rate, but since 2010, it has fluctuated around zero. During the early phase of the COVID-19 pandemic, the rate of household saving again increased.

Generally speaking, Finnish household saving tends to increase in recessions and decrease during economic expansions, suggesting the existence of certain precautionary motives driving household consumption and saving patterns over business cycles. Households may also be constrained by various credit constraints. If those credit constraints depend on their income or wealth, they may be countercyclical, *i.e.* relax during economic booms as income and asset prices climb, and tighten in downturns. While the indebtedness ratio, defined in terms of a stock variable, moves much more slowly, it nonetheless exhibits a degree of cyclical behavior.

In addition to these contemporaneous correlation patterns, changes in loan growth have historically foreshadowed business cycle fluctuations.[1] This is illustrated in Figure 2, which shows cross correlations between changes in the household loans-to-income ratio and selected macroeconomic variables. The loans-to-income ratio is divided in three categories, depending on the purpose for which the loan was acquired.

The upper left panel displays cross correlations between changes in the loans-to-income ratio and real GDP growth. Growth in loans relative to disposable income is negatively correlated with real GDP

[1] A negative association between household debt growth and subsequent GDP growth has been indentified by *e.g.* Mian, Sufi, and Verner (2017) using a panel of countries and Mian and Sufi (2010) using state-level data from the US. A prominent piece of evidence supporting a causal effect of indebtedness on a subsequent economic slowdown is provided in Verner and Gyöngyösi (2020). There is a large body of literature on the predictive power of household debt growth for financial crises; see *e.g.* Borio and Drehmann (2009), Schularick and Taylor (2012), and Jordà, Schularick, and Taylor (2013), and references therein. In the Finnish context, similar results have been reported by Lainà, Nyholm, and Sarlin (2015) and Nyholm and Voutilainen (2021).

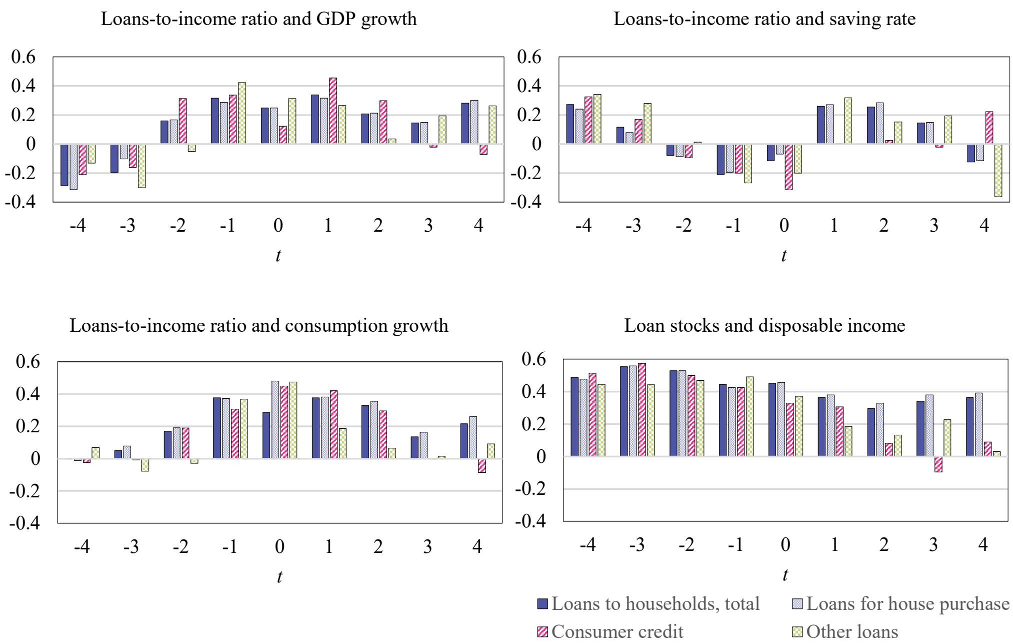

Figure 2: Cross correlation coefficients between components of the loans-to-income ratio or loan stock growth, and other macro variables. Correlations are computed as correlations between a given macro variable in year zero and loan stock growth, or changes in the household loans-to-income ratio, in year $t = -4, \ldots, 4$, and estimated using data from 2003Q1 to 2021Q4. Macro variables are the year-on-year real GDP growth (upper left), net saving rate (upper right), year-on-year real private consumption growth (lower left), and year-on-year growth in nominal disposable income of households (lower right). Changes in the loans-to-income ratio are calculated as *growth in loan stock – growth in disposable income*. In the lower right panel, the correlations have been calculated using loan stock growth rates instead of the loans-to-income ratio, because disposable income is present in the denominator of the definition for the loans-to-income ratio.

growth three to four years in the future. In other words, rapid indebtedness growth today predicts lower economic growth in three to four years. Contemporaneously, the two variables are positively correlated, which means that indebtedness tends to grow simultaneously with GDP. High GDP growth today also anticipate rapid growth in the loans-to-income in one to two years.

The upper right panel in Figure 2 displays the correlation between growth in the loans-to-income ratio and the net saving rate. Their contemporaneous correlation is negative, which means that indebtedness growth – especially in consumer credit – is associated with negative saving.

The lower left panel of Figure 2 shows that the contemporaneous correlation between real private consumption growth and changes in the household loans-to-income ratio is strongly positive. Fast consumption growth today is also associated with indebtedness growth in one to two years. Rapid growth in the loans-to-income ratio, however, does not necessarily indicate future consumption growth.

Finally, the lower right panel displays correlation patterns between the growth in loan stocks and nominal disposable income growth. For total loan stock growth, we find a clear positive relationship between income growth and loan growth both contemporaneously and at all lags. In addition, there is clear positive correlation with disposable income growth today and growth in housing loans three to four years subsequent.

3 Methodology

This section briefly describes the methodology used in the model estimation and conditional forecasting of household loan stocks.

3.1 Model estimation

Vector autoregressive (VAR) models are highly flexible tools for modeling the dynamics between macroeconomic variables, but the large amount of variables combined with a short estimation period and a limited number of observations creates challenges to the precision of forecasts produced with such VAR models.

The estimation of such large systems involves a lot of uncertainty, which can also lead to inaccurate out-of-sample properties of the model. Bayesian estimation methods may provide a solution to this curse of dimensionality by shrinking the parameter space of the model toward one that admits models that are compatible with some prior information. We therefore use Bayesian methods in estimating the VAR model. In choosing the informativeness of the priors in our model, we rely on the theoretically grounded and data-driven prior selection method proposed by Giannone et al. (2015).

Following Giannone et al. (2015), we specify a hierarchical BVAR model, *i.e.* we specify a hyperprior distribution for the parameters of the prior. This hierarchical Bayesian approach is based on the idea that hyperparameters, which in the standard Bayesian approach define the prior distribution of the parameters of interest, are also subject to uncertainty. This allows to conduct sensitivity analysis on the prior.

Additionally, the Giannone et al. (2015) algorithm relies, essentially, on an empirical Bayesian approach, where the parameters of the hyperprior – and hence the informativeness of the prior – are chosen by exploiting observed data. The algorithm chooses a given model's hyperparameters to maximize the posterior of the hyperparameter distribution, conditional on observed data. The optimization algorithm trades off goodness of the in-sample fit of the model with a penalty for model complexity that leads to imprecise out-of-sample forecasts. While this approach is not purely Bayesian in the sense that observed data are used in specifying the prior hyperparameters, it works quite well in forecasting applications.

The VAR model is specified as:

$$y_t = C + By_{t-1} + \cdots + B_p y_{t-p} + \varepsilon_t, \tag{1}$$

$$\varepsilon_t \sim N(0, \Sigma), \tag{2}$$

where C is a constant $(n \times 1)$ parameter vector and B_r, $r = 1, \ldots, p$, are constant $(n \times n)$ parameter matrices. Σ denotes the covariance matrix of the error term ε_t, and y_t is a vector of the endogenous variables observed at time t.

We use a prior distribution belonging to the Normal-Inverse-Wishart family:

$$\Sigma \sim IW(\Psi, d) \tag{3}$$

$$\beta \sim N(b, \Sigma \otimes \Omega), \tag{4}$$

where b is the prior mode of $\beta = vec([C, B_1, \ldots, B_p]')$, the vector of coefficients of the VAR model, Ω is a weight matrix, d is the degrees-of-freedom parameter of the Inverse-Wishart distribution, and Ψ is its scale matrix. This choice is well justified given its popularity in previous studies and the fact that this family of distributions admits a closed-form solution for the marginal likelihood (ML) function of our model. The prior densities and their corresponding parameters are shown in Table 3 in the Appendix.

For the (conditional) Gaussian prior distribution for β, we use the Minnesota prior as the baseline for our analysis. This prior assumes a univariate random walk process for each endogenous variable. The informativeness of this prior is controlled by a hyperparameter vector $\gamma = (\psi, \lambda)$, where ψ is a $n \times 1$ vector such that $\Psi = diag(\psi)$.[2] The key hyperparameter λ controls the scale of the covariances in $\Sigma \otimes \Omega$, or the overall tightness of the prior. The vector ψ defines the prior mean of the variance of the VAR coefficients B_r, *i.e.* the prior mean of the main diagonal in Σ. These hyperparameters together

[2]Notice that this is the main difference compared to the standarard Bayesian approach, where the hyperparameter γ would be assumed a constant hyperparameter.

control the informativeness of the prior on β. For further details on specifying the prior in this form, we refer the reader to Giannone et al. (2015). We specify relatively loose priors on the hyperparameters, following Giannone et al. (2015).

In much of the literature, it has been customary to use rules-of-thumb and judgment calls in adjusting the informativeness of these priors. Here, however, we follow Giannone et al. (2015) and study a hierarchical model, treating the vector γ as a hyperparameter and its prior density as a hyperprior.

Giannone et al. (2015) show that the posterior for this hyperparameter can be written as $p(\gamma|y) \propto p(y|\gamma)p(\gamma)$. Here,

$$p(y|\gamma) = \prod_{t=p+1}^{T} p(y_t|y^{t-1}, \gamma), \tag{5}$$

where y^t is a vector of observations prior time t. This function is the marginal likelihood (ML) function, *i.e.* the density of the data as a function of the hyperparameters γ, after intergrating out uncertainty about model parameters. It follows that for an uninformative hyperprior distribution $p(\gamma)$, posterior maximization of the hyperparameter distribution coincides with minimizing the forecast error of the one-step-ahead forecast of the model.

3.2 Conditional forecasts

Conditional forecasts are computed by applying the Kalman filter to the state-space presentation of the VAR model (Lütkepohl, 2005, Chapter 18). The conditioning variables are listed in the center panel of Table 2 in the Appendix. They are treated as observed variables over each forecast horizon. All other model variables, *i.e.* all variables for which no conditioning information is available, are treated as latent variables. Forecasts for these variables are obtained by running the forecasting steps of the Kalman filter recursion using the estimated values of the parameters in the transition and the measurement equations.

4 Conditional forecasts for household loan stocks

This section starts with a description of the data that is used in the model estimation and evaluation. We also describe different model specifications that we use in order to test the forecasting performance of our BVAR model.

Robustness of the results is tested against different model specifications and estimation periods, in terms of *root mean squared forecasting errors* (RMSFE). We also asses the usefulness of using conditioning information in producing the forecasts.

4.1 Data

We use quarterly data from 2003Q1 to 2021Q4.[3] The data are sourced from the Bank of Finland, Statistics Finland, and the ECB. Table 2 in the Appendix summarizes the variables and indicates how they have been modified. The model is estimated in (log-)levels. Macroeconomic aggregates and loan stock variables are expressed in log-levels, while interest rates and interest-rate spreads are expressed in levels (percentage points).

We can divide the data in three categories. In the first category, we have the three household loan stocks of interest, for which forecasts are produced. In the second category, we have endogenous variables also included in the Aino 2.0 model, for which forecasts produced using that model are available. As our BVAR model is designed to be used as a satellite model for the Aino 2.0 model, these forecast paths can serve as conditioning information when predicting the household loan stocks. In the next subsection, when we evaluate the model performance, we use the realized values of the variables in the second category as conditioning information.[4]

The third category consists of all other endogenous variables not included in the Aino 2.0 model and for which we lack forecast paths. These variables are treated as unobservables in the conditional forecasting procedure. The inclusion of these variables in alternative specifications of our model is motivated by economic theory. We test whether they are informative for the dynamics of the loan stock variables of interest and whether they improve the forecasting performance of the model.

Next, we describe the variables included in each of these categories. The first category includes the three types of household loan stocks that the model is built to predict. They are *Loans for house purchase (LHP)*, *Consumer credit (LCC)*, and *Other loans (LOL)*. They contain loans granted by MFIs resident in Finland to households and non-profit institutions serving households. The loan stocks do not contain loans granted to housing corporations as they are classified as corporate loans.

The second category includes six endogenous variables from the Aino 2.0 model. The *euro area interest rate (ESHORT)* is the 3-month Euribor rate. *Finnish GDP (FGDP)* is the quarterly volume index of GDP, and *Finnish HICP (FHICP)* is the monthly Finnish Harmonized Index of Consumer Prices aggregated to quarterly frequency. *Disposable income (INDH)* is the quarterly nominal disposable income of Finnish households and non-profit institutions serving households. The *household saving rate (SAVRH)* is the quarterly net saving rate of Finnish households and non-profit institutions serving

[3]This time period comprises both the period of high loan growth that begins in the early 2000s and the period of lower loan growth in the aftermath of the global financial crisis. The choice of the estimation sample is constrained by data availability.

[4]We acknowledge that this procedure likely overestimates the prediction accuracy of the BVAR model compared to real-time forecasting, where predicted values of the variables in the second category serve as conditioning information. In any case, our approach provides a fair comparison between different model specifications.

households.

The fourth loan stock variable we include are *loans to non-financial corporations (LNFC)*, which is the stock of loans granted by MFIs resident in Finland to non-financial corporations, including housing corporations. Last, the *house price index (HP)* is an aggregate index of nominal prices of all dwellings in Finland.[5]

The third category includes all other endogenous variables. The *lending spread (FSPREAD)* is calculated as the spread between the weighted average interest rate on all loans granted to households by MFIs resident in Finland and the short-term interest rate.[6]

The *Bank Lending Survey (BLS)* composite index collects information on non-price credit conditions, which might remain invisible if we only included the lending spread in the model. In the survey, banks are asked whether they have tightened or loosened their credit terms and conditions on various aspects of loan contracts. A positive (negative) number on each sub-index corresponds to the net share of banks that have reported tightening (loosening) on their terms and conditions. Our *BLS* variable is an average value of indices measuring changes in collateral requirements, margins on average loans, margins on riskier loans, maturity, and non-interest rate charges on loans to households. Finally, the *CISS index* is a composite index of 15 financial stress indices for the euro area selected from five market categories: the financial intermediaries sector, money markets, equity markets, bond markets, and foreign exchange markets (Kremer, Lo Duca, and Holló, 2012).

4.2 Model evaluation

This subsection describes our approach for model selection, and discusses various dimensions of model specification: variable selection, lag selection, choice of estimation method, and the role of conditioning information. Our aim is to build a model that is able to forecast the household loan stocks of interest as accurately as possible. We evaluate the forecasting performance of each specification by computing out-of-sample RMSFEs, calculated using an expanding estimation sample. When calculating the RMSFEs, the out-of-sample forecasts generated by the model are compared to the true (ex-post) observed data at each estimation sub-sample and horizon. For illustrative purposes, expanding window forecasts for Model 9, together with observed data from 2003Q1 to 2021Q4, are shown in Figure 8 in the Appendix.

[5]Unlike the rest of the variables in this category, a separate satellite VAR model of the Aino 2.0 model is used to forecast house prices. Like the loan stock variables in this exercise, the forecasts for house prices are consistent with predictions for key macro variables of the Aino 2.0 model.

[6]We also considered the lending spread on all loans granted to the private sector, but found that the lending spread on household loans is more useful in predicting household loan stocks. These results are omitted here, but available on request.

	Model number									
(1) Loan stock variables	1	2	3	4	5	6	7	8	9	10
Loans for housing purchases (LHP)	x	x	x	x	x	x	x	x	x	x
Consumption credit (LCC)	x	x	x	x	x	x	x	x	x	x
Loans for other purposes (LOL)	x	x	x	x	x	x	x	x	x	x
(2) Conditioning variables										
Euro area interest rate, 3 months Euribor (ESHORT)	x	x	x	x	x	x	x	x	x	x
Finnish GDP (FGDP)	x	x	x	x	x	x	x	x	x	x
Finnish HICP (FHICP)	x	x	x	x	x	x	x	x	x	x
Loans to non-financial corporations (LNFC)	x	x	-	x	x	x	x	x	x	x
Disposable income (INDH)	x	x	x	-	x	x	x	x	x	x
Savings rate (SAVHR)	-	x	x	x	-	x	x	-	-	-
Housing prices (HP)	-	-	x	x	x	x	x	-	-	-
(3) Other variables										
Lending spread (FSPREAD)	x	x	x	x	x	-	x	x	x	x
Bank lending survey (BLS)	-	-	-	-	-	-	-	x	-	x
Euro area financial stress index (CISS)	-	-	-	-	-	-	-	-	x	x

Table 1: Description of the model specifications used throughout the analysis.

4.2.1 Variable selection

The sets of endogenous variables we evaluate are summarized in Table 1. Each specification naturally includes all three loan stock variables of interest (top panel in the table). In addition, the specifications include different combinations of conditioning variables (center panel) and other endogenous variables (bottom panel).

We evaluate the forecast performance of these different models over several dimensions. Models 1–7 are combinations of the conditioning variables and the variable *FSPREAD*. Models 8–10 evaluate the impact of two additional variables compared with Model 1, *CISS* and *BLS*, which may provide useful information about developments in loan stocks, but for which we assume that no conditioning information is available.

Model 9 is the specification that we find to provide the best forecasting accuracy overall. This model includes the three loan stock variables of interest (*LHP*, *LCC*, and *LOL*), five endogenous conditioning variables from the Aino 2.0 model (*ESHORT*, *GDP*, *HICP*, *LNFC*, and *INDH*), the lending spread *FSPREAD*, and the CISS financial stress index (*CISS*).

We start by studying the overall forecasting performance for the components of the household loan stock with different specifications. These results are summarized in Figure 3.

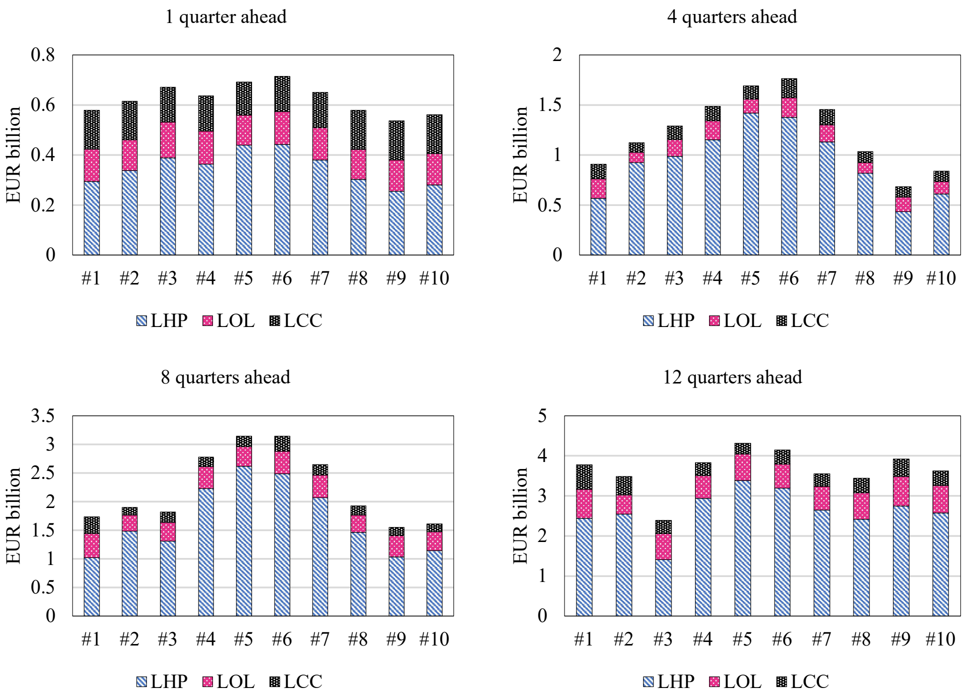

Figure 3: Absolute root means squared forecast errors (RMSFEs) of conditional loan stock forecasts, for different model specifications. RMSFEs are reported for each model at four different forecasting horizons. All models are estimated using lag length of one ($p = 1$) and the prior selection method of Giannone et al. (2015). *LHP*, *LOL*, and *LCC* stand for loans for house purchase, other loans, and consumer credit, respectively.

Each model specification is estimated using an expanding window running from 2003Q1 to 2018Q4, and forecasts are produced for one, four, eight and twelve quarter horizons at each sub-sample. Finally, we compute the out-of-sample root mean squared forecast errors (RMSFE) in absolute terms at each forecast horizon. Estimation is performed using the optimal prior selection method proposed by Giannone et al. (2015).

In terms of RMFSEs, the most parsimonious specification (Model 1) compares favorably against Models 2–7, implying that of the variables included in the set of conditioning variables, *ESHORT, FGDP, FHICP, LNFC*, and *INDH* are the most useful variables in predicting the future evolution of the different components of household loan stocks. *FSPREAD* is also included in this model specification, although it is not included in the set of conditioning variables. It is included, since *FSPREAD* appears as an obvious candidate in controlling for the dynamics of the loan stocks.[7]

[7]We also evaluated model specifications more parsimonious than Model 1, but their forecasting performance was

With Models 8–10, we investigate further whether other variables, besides those for which we assume conditioning information is available, would improve the forecasting accuracy even further. Comparison across all models (Models 1–10) suggests that information that improves forecasting accuracy is revealed by the *CISS* stress index.

It comes as no surprise that this particular variable is useful. It is a composite indicator that captures accurately the current stress level of euro area financial markets, a proxy for credit supply conditions in the euro area. However, inclusion of the domestic *BLS* index (Models 8 and 10), which is a measure of domestic credit supply conditoins, does not further enhance model performance. Model 9 outperforms the rest of the models at horizons from one to eight quarters. For the longest horizon of twelve quarters, the models with *BLS* perform better than Model 9.

The biggest improvements in accuracy are gained from the reduced RMSFEs of the housing loans, which are the largest component in the total loan stock. In relative terms, forecasting accuracy is quite even across all models when it comes to the other components of household loans. This is illustrated in Figure 7 in the Appendix, where the forecast accuracy of different models are compared across different components of the loan stock, relative to Model 1, our most parsimonious model.

4.2.2 Impact of the estimation method

The estimation method may have substantial impact on the forecast accuracy. We illustrate this in Figure 4, which displays the RMSFEs of forecasts for Models 1–10 estimated using different methods and one lag ($p = 1$).[8]

Figure 4 reports the RMSFEs of a particular model estimated using standard Bayesian methods with Minnesota-type priors, under the assumption of constant hyperparameters and using the Giannone et al. (2015) estimation method explained in the previous section. The RMSFEs of these two methods for Models 1–10 are provided at four forecasting horizons.

One-quarter-ahead forecasts are not substantially affected by the estimation method, but the difference in accuracy is more striking for the longer horizons. Especially Models 1 and 9, which were shown to perform well overall, clearly benefit from the Giannone et al. (2015) method at 4- and 8-quarter forecasting horizons. This prior selection method improves the forecasting performance in most models at horizons longer than one quarter. The most notable exception is the 12-quarter-ahead forecast using Model 9, where the optimal prior selection algorithm slightly *decreases* forecast accuracy. Overall,

worse than that of Model 1. These results are not reported here, but available on request.

[8]Estimation results using the OLS method are also available on request. The large dimensionality of the model and the relatively short span of our dataset mean that without any shrinkage in the parameter space, the estimation uncertainty dominates the positive sides of not having to impose any prior judgement on the parameter values. Thus, these relatively large RMSFEs have been omitted here.

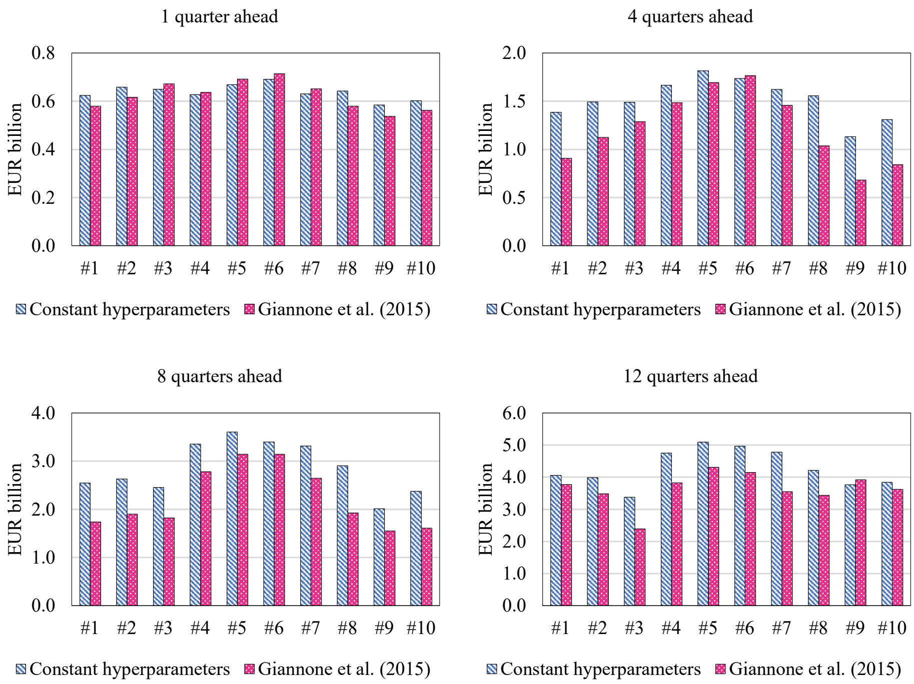

Figure 4: Comparison of different model specifications and estimation methods. RMSFEs of the total stock of household loans are reported for Models 1–10 with a lag length of one ($p = 1$), estimated using two different Bayesian estimation methods, conditional on the observed values of the conditioning variables.

differences between the two methods are not significantly different at this longest horizon.

4.2.3 Lag selection

Figure 5 compares the forecast accuracy of Models 1–10 using different lag lengths p of the AR polynomial, when the models are estimated with the optimal prior selection method. Including one additional lag to our preferred ten-variable model (Model 9) adds 100 additional parameters to be estimated, so it is clear that the priors must be quite informative about these variables.

Models with two lags ($p = 2$) seem to produce the most accurate forecasts of the household loan stocks. The benefit of adding a second lag to the specification is the most pronounced at the one quarter horizon, where this additional lag decreases the RMSFEs of Model 9 by almost one third. The same phenomenon also applies to several other model specifications. Model 9 does not benefit significantly from another lag when forecasts are made for longer horizons. Indeed, adding a larger number of lags

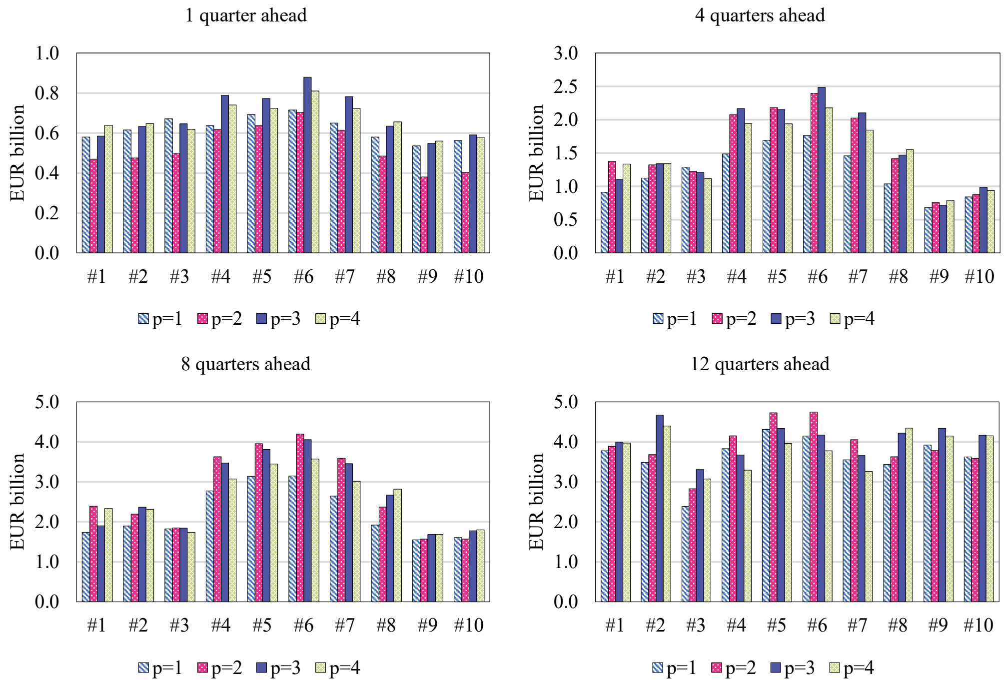

Figure 5: Comparison of different model specifications and for different lag lengths p of the AR polynomial. RMSFEs of the total stock of household loans are reported for each specification and at four different forecasting horizons, conditional on the observed values of the conditioning variables.

seems to reduce forecast accuracy. We conclude that Model 9 with two lags has the best forecasting performance overall among our various model specifications.

4.2.4 The role of conditioning information

The role of conditioning information in forecasting performance is investigated in Figure 6. We condition on the (ex-post) observed paths of the conditioning variables at each sub-sample. Each model has been estimated using lag length of $p = 2$ and the optimal prior selection method of Giannone et al. (2015).

Conditioning information does not necessarily improve forecasts. This can be seen by looking at Models 4–7, where the forecast accuracy is significantly better without conditioning. In some other cases, however, knowledge of the future paths of the conditioning variables does improve the accuracy substantially. This is the case with Models 1 and 9, for example. The RMSFEs of conditional forecasts are almost one half smaller than their unconditional counterparts for Model 9 at the 4- and 8-quarter

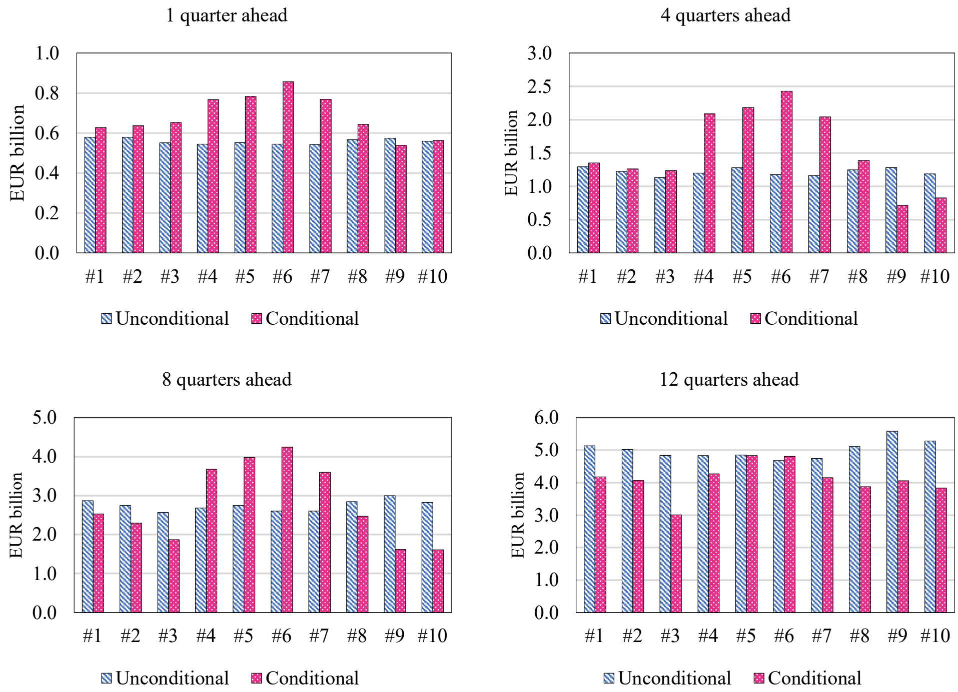

Figure 6: Comparison of different model specifications and the role of conditioning information. RMSFEs are reported for the total stock of household loans. Each model has been estimated using the prior selection method by Giannone et al. (2015) and using the lag length of two ($p = 2$).

horizons. The difference in accuracy is significant also at the longest forecasting horizon of 12 quarters, but less so than at the shorter ones, in relative terms.

The fact that information about the future values of the conditioning variables can be used to improve forecast accuracy over the unconditional forecasts means, first, that there is a significant amount of information in the conditioning variables about the future values of the loan stocks; and second, that the estimation method is capable of capturing the dynamics between the endogenous variables, even in such a multidimensional system.

The Giannone et al. (2015) algorithm seems to work well in selecting the hyperprior parameters in such a way as to capture the dynamics between the endogenous variables. When the priors are too tight, the unit root assumption of the variables implies no linear predictability of the loan stocks, even given the future paths of other variables. When priors are too slack, the overwhelming estimation uncertainty potentially makes the estimation results unstable. This could make information about the future values of the endogenous conditioning variables useless as the dynamics between these variables may well be misspecified.

Thus, our preferred specification (Model 9) is well-suited for producing conditional forecasts in which future paths of key macroeconomic variables are used as conditioning information.

5 Conclusions

We specify a new forecasting model for Finnish household loan stocks. This model complements the Aino 2.0 DSGE model in forecasting exercises as household loans are not incorporated in the core Aino 2.0 model. Our forecasts are conditional on and thus consistent with the macroeconomic forecasts derived from the Aino 2.0 model.

The BVAR model serves as a good candidate for forecasting purposes. Generally speaking, linear systems provide good approximations for general equilibrium models of macroeconomic dynamics, and thus can potentially be quite good in controlling for the dynamics between macroeconomic variables and loan stocks.

We propose the use of Bayesian estimation methods for fitting the model. In doing so, we mitigate the eminent curse of dimensionality that arises from the short data sample and large set of endogenous variables included in our analysis. Bayesian methods can be used to shrink the parameter space toward our prior information. They also allow using information provided by the data whenever it is informative about the variables of interest.

Our priors reflect a common unit root assumption for the data generating process, formulated by Litterman (1979). We then apply the optimal prior selection method by Giannone et al. (2015), which allows us to select the optimal tightness of our priors to reflect the richness of the information in the data and to be more agnostic about the prior when the data provide a clearer picture about the dynamics of the system. Giannone et al. (2015) show that this method works especially well in forecasting as the optimal prior selection algorithm coincides with optimizing the one-step-ahead forecast accuracy in the case of uninformative priors.

We test our model performance in an out-of-sample forecasting exercise, whereby we calculate conditional forecasts of loan stocks conditional on the true observed paths of the endogenous macro variables included in our model. This exercise is repeated for several different specifications of the model. We use the forecasting performance, measured in the terms of RMFSEs, as our criterion for model selection. We conclude that our preferred specification, estimated with two lags in the autoregressive polynomial, provides a robust and well-performing specification for the purpose of forecasting household loan stocks.

References

Borio, C. E. and M. Drehmann (2009). Assessing the risk of banking crises – revisited. *BIS Quarterly Review, March 2009*.

Giannone, D., M. Lenza, and G. E. Primiceri (2015). Prior selection for vector autoregressions. *Review of Economics and Statistics 97*, 436–451.

Jordà, O., M. Schularick, and A. M. Taylor (2013). When credit bites back. *Journal of Money, Credit and Banking 45*, 3–28.

Kilponen, J., S. Orjasniemi, A. Ripatti, and F. Verona (2016). The Aino 2.0 model. *Bank of Finland Research Discussion Paper 16/2016.*

Kremer, M., M. Lo Duca, and D. Holló (2012, March). CISS – a composite indicator of systemic stress in the financial system. *European Central Bank Working Paper Series 1426.*

Lainà, P., J. Nyholm, and P. Sarlin (2015). Leading indicators of systemic banking crises: Finland in a panel of EU countries. *Review of Financial Economics 24*, 18–35.

Litterman, R. B. (1979). Techniques of forecasting using vector autoregressions. *Federal Reserve Bank of Minneapolis Working Paper 115.*

Lütkepohl, H. (2005). *New introduction to multiple time series analysis.* Springer.

Mian, A. and A. Sufi (2010). Household leverage and the recession of 2007–09. *IMF Economic Review 58*, 74–117.

Mian, A., A. Sufi, and E. Verner (2017). Household debt and business cycles worldwide. *Quarterly Journal of Economics 132*, 1755–1817.

Nyholm, J. and V. Voutilainen (2021). Quantiles of growth: Household debt and growth vulnerabilities in Finland. *Bank of Finland Economics Review 2/2021.*

Schularick, M. and A. M. Taylor (2012). Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008. *American Economic Review 102*, 1029–61.

Verner, E. and G. Gyöngyösi (2020). Household debt revaluation and the real economy: Evidence from a foreign currency debt crisis. *American Economic Review 110*, 2667–2702.

Appendix

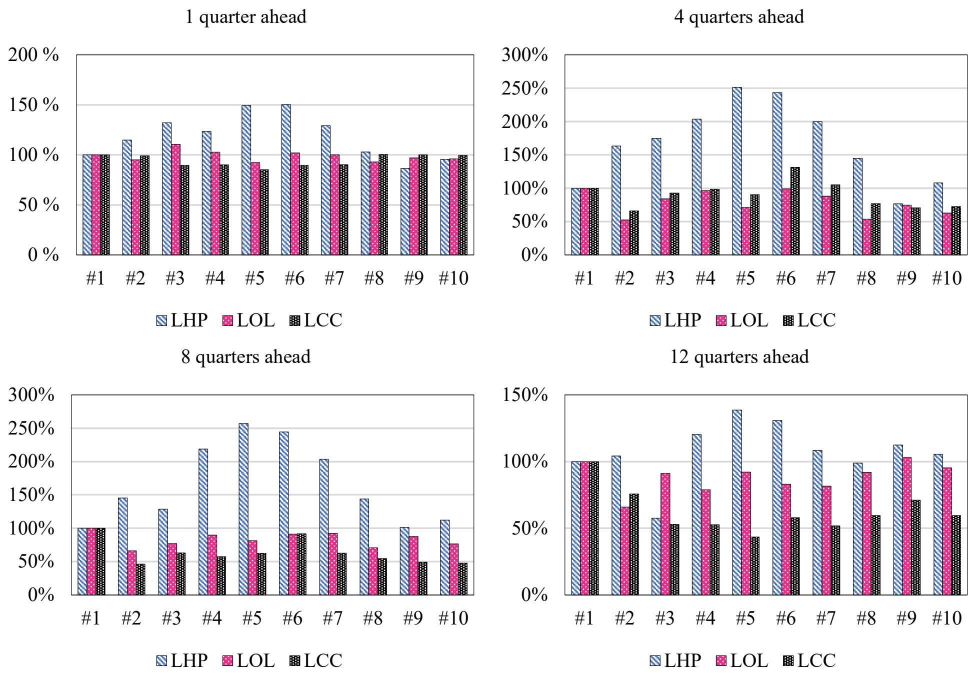

Figure 7: Comparison of different model specifications relative to Model 1. RMSFEs of each component of the household loan stock are reported relative to the RMSFEs provided by the Model 1, using conditional forecasts, for four forecasting horizons.

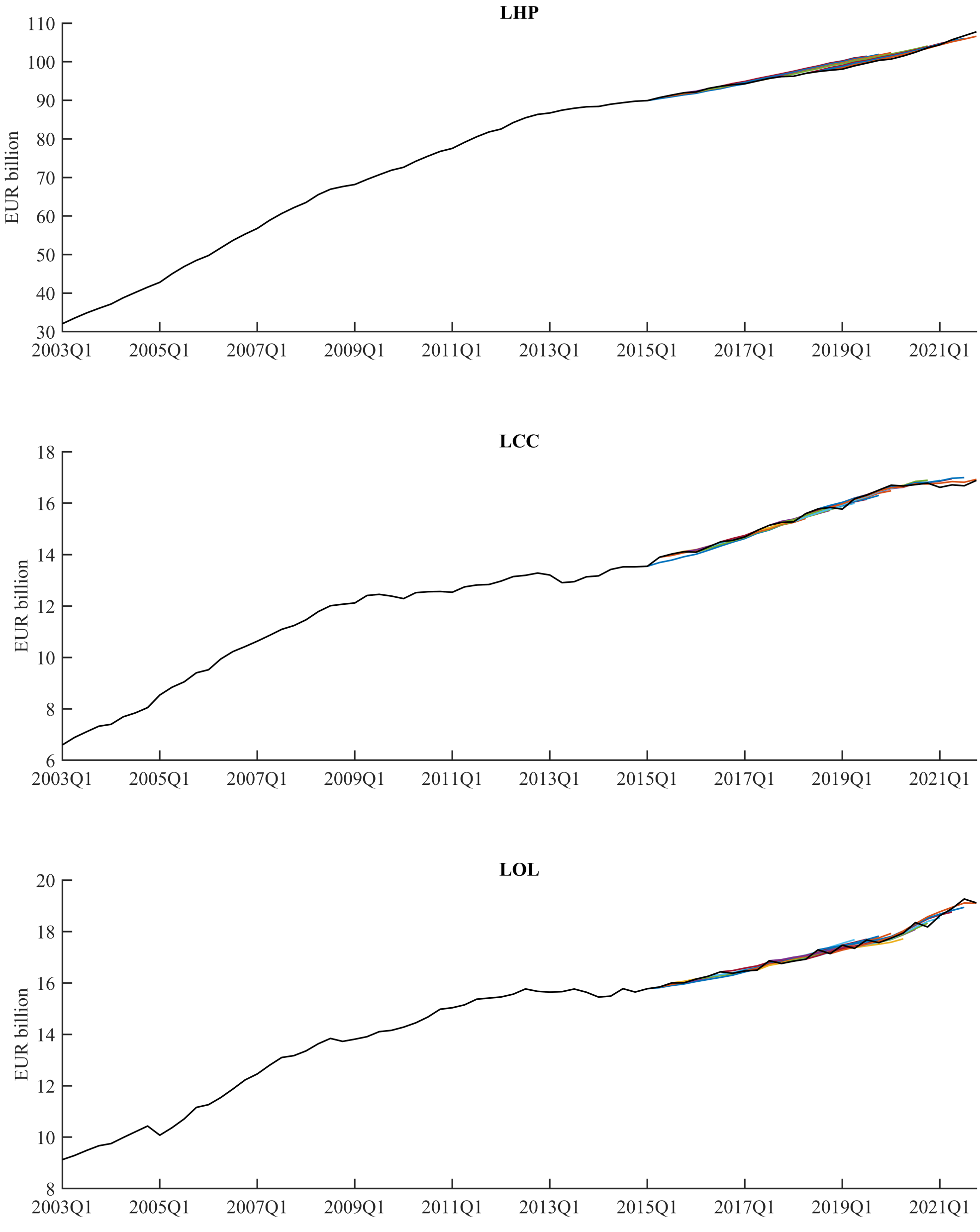

Figure 8: Expanding window forecasts for household loan stocks, 12 quarters ahead. Forecasts are produced with Model 9 with lag length $p = 2$, using the algorithm proposed by Giannone et al. (2015). *LHP*, *LOL*, and *LCC* stand for loans for house purchase, other loans, and consumer credit, respectively.

Abbreviation	Definition	Modifications	Source
LHP	Loans to households and non-profit institutions serving households, issued by MFIs resident in Finland, loans for house purchase, EUR million	Log-levels	BoF
LCC	Loans to households and non-profit institutions serving households, issued by MFIs resident in Finland, consumer credit, EUR million	Log-levels	BoF
LOL	Loans to households and non-profit institutions serving households, issued by MFIs resident in Finland, other loans, EUR million	Log-levels	BoF
ESHORT	Euro area interest rate, 3-month Euribor	Percentage points	ECB
FGDP	Finnish GDP volume, in 2015 prices, EUR million	Log-levels	Statistics Finland
FSPREAD	Lending spread: average interest rate on household loans issued by Finnish MFIs, minus euro area short term interest rate	Percentage points	ECB, BoF
FHICP	Finnish harmonized index of consumer prices, 2005=100	Log-levels	Statistics Finland
LNFC	Loans to non-financial corporations (including housing corporations), issued by MFIs resident in Finland, EUR million	Log-levels	
INDH	Disposable income, households and non-profit institutions serving households, EUR million	Log-levels	Statistics Finland
HP	House price index, all dwellings, entire country, 2010=100	Log-levels	Statistics Finland
SAVRH	Net saving rate, households and non-profit institutions serving households	Percentage points	Statistics Finland
BLS	Bank lending survey, credit terms and conditions, average of collateral requirements, margin on average loans, margin on riskier loans, maturity, and non-interest charges	Levels	ECB, BoF
CISS	Composite index for systemic stress in the financial markets in the euro area (Kremer et al., 2012)	Levels	ECB

Table 2: Description of the data used in the analysis. BoF: Bank of Finland. ECB: European Central Bank.

	Parameter	Distribution	Mode / mean	s.d. / cov
Prior distributions for VAR coefficients	Σ	$IW(\Psi, d)$	$\Psi(d-n-1)^{-1}$, where $\Psi = diag(\psi)$	undefined
	$\beta \vert \Sigma$	$N(b, \Sigma \otimes \Omega)$	$E[(B_r)_{ij} \vert \Sigma] = \begin{cases} 1 \text{ if } i=j \text{ and } r=1 \\ 0 \text{ otherwise} \end{cases}$	$\text{cov}((B_s)_{ij}, (B_r)_{kl} \vert \Sigma) = \begin{cases} \lambda^2 s^{-2} \frac{\Sigma_{ik}}{\psi_j/(d-n-1)} \text{ if } j=l \text{ and } s=r \\ 0 \text{ otherwise} \end{cases}$
Hyperprior distributions	λ	$\Gamma(k_\lambda, \theta_\lambda)$	0.2	0.4
	$\psi_i(d-n-1)^{-1}$	$Inv.\Gamma(\xi_\psi, \zeta_\psi)$	0.02^2	undefined
	d	constant $(n+2)$		

Table 3: Summary of the prior and hyperprior distributions.

BoF Economics Review

2020	No 1	Crowley, Patrick M.; Hudgins, David: How Effective is the Taylor rule? : Some Insights from the Time-Frequency Domain
	No 2	Ambrocio, Gene; Juselius, Mikael: Dealing with the costs of the COVID-19 pandemic – what are the fiscal options?
	No 3	Kilponen, Juha: Koronaviruskriisi leikkaa syvän loven Suomen talouteen
	No 4	Laine, Olli-Matti; Lindblad, Annika: Nowcasting Finnish GDP growth using financial variables : a MIDAS approach
	No 5	Kortelainen, Mika: Yield Curve Control
	No 6	Ambrocio, Gene: European household and business expectations during COVID-19 : Towards a v-shaped recovery in confidence?
	No 7	Nissilä, Wilma: Probit based time series models in recession forecasting : A survey with an empirical illustration for Finland
	No 8	Grym, Aleksi: Lessons learned from the world's first CBDC
2021	No 1	Kärkkäinen, Samu; Nyholm, Juho: Economic effects of a debt-to-income constraint in Finland : Evidence from Aino 3.0 model
	No 2	Nyholm, Juho; Voutilainen, Ville: Quantiles of growth : household debt and growth vulnerabilities in Finland
	No 3	Juselius, Mikael; Tarashev, Nikola: Could corporate credit losses turn out higher than expected?
	No 4	Nelimarkka, Jaakko; Laine, Olli-Matti: The effects of the ECB's pandemic-related monetary policy measures
	No 5	Oinonen, Sami; Vilmi, Lauri: Analysing euro area inflation outlook with the Phillips curve
	No 6	Pönkä, Harri; Sariola, Mikko: Output gaps and cyclical indicators : Finnish evidence Analysing euro area inflation outlook with the Phillips curve
	No 7	Hellqvist, Matti; Korpinen, Kasperi: Instant payments as a new normal : Case study of liquidity impacts for the Finnish market
	No 8	Markkula, Tuomas; Takalo, Tuomas: Competition and regulation in the Finnish ATM industry
	No 9	Laine, Tatu; Korpinen, Kasperi: Measuring counterparty risk in FMIs
	No 10	Kokkinen, Arto; Obstbaum, Meri; Mäki-Fränti, Petri: Bank of Finland's long-run forecast framework with human capital
2022	No 1	Norring, Anni: Taming the tides of capital – Review of capital controls and macroprudential policy in emerging economies
	No 2	Gulan, Adam; Jokivuolle, Esa; Verona, Fabio: Optimal bank capital requirements: What do the macroeconomic models say?
	No 3	Oinonen, Sami; Virén, Matti: Has there been a change in household saving behavior in the low inflation and interest rate environment?
	No 4	Nyholm, Juho; Silvo, Aino: A model for predicting Finnish household loan stocks

https://www.suomenpankki.fi/en/media-and-publications/publications/studies-and-reports/bof-economics-review/