



**LIETUVOS BANKAS**  
EUROSISTEMA

# Lithuanian house price index: modelling and forecasting

Occasional Paper Series

No 28 / 2019

Laurynas Naruševičius<sup>\*</sup>

Tomas Ramanauskas

Laura Gudauskaitė

Tomas Reichenbachas

November 2019

---

<sup>\*</sup>Corresponding author: Financial Stability Department, Bank of Lithuania. Email: LNarusevicius@lb.lt.

© Lietuvos bankas, 2019

Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

Gedimino pr. 6, LT-01103 Vilnius

[www.lb.lt](http://www.lb.lt)

The series is managed by Applied Macroeconomic Research Division of Economics Department and the Center for Excellence in Finance and Economic Research.

The views expressed are those of the author(s) and do not necessarily represent those of the Bank of Lithuania.

## CONTENTS

Abstract.....	5
1. Introduction .....	6
2. Discussion about the fundamentals of house price development .....	6
3. Discussion about house price modeling and forecasting .....	8
4. Dynamics of Lithuanian house price index .....	9
5. Fundamentals of the Lithuanian house price index: VECM analysis .....	10
6. Determinants of the house price growth and short term forecasting: ARDL model.....	16
7. Short and medium term forecasting: forecast combination .....	19
References.....	26
Appendix A .....	29
Appendix B .....	30
Appendix C .....	31

## **ABSTRACT**

Timely monitoring of the housing market developments in Lithuania is one of the key elements in the analysis framework of the macroprudential authority aiming to contribute to financial stability in Lithuania. In this paper, we addressed three important questions related to Lithuanian house prices, namely, whether house prices are under- or over valued, which explanatory variables have the biggest impact on the growth of house prices and what the future development of the Lithuanian house price index could be. Three separate modelling and forecasting exercises were performed in order to tackle these questions. The first exercise employs the vector error correction modelling (VECM) approach to assess under- or overvaluation of the house prices. We then use an autoregressive distributed lag model (ARDL) to evaluate which explanatory variables have the biggest impact on house price growth. As the last exercise, we develop a suite of models that are used to forecast future development of the house price index. The analysis presented in this paper may be viewed as a further step towards more formalised modelling and forecasting of the Lithuanian house price index.

**JEL:** C22, C32, C53, E37, R30

**Keywords:** House price index, fundamental value, time series models, forecasting, forecast combination

## 1. INTRODUCTION

The Global Financial Crisis showcased how increase in house price volatility may affect real economic activity and induce stress in the financial system. The Lithuanian housing market was very buoyant during the boom years and then declined sharply in 2009. This had huge negative consequences for households with mortgage loans, banks that experienced deterioration of their balance sheets, the construction and real estate sector, and the overall economy. Therefore, timely monitoring of housing market developments in Lithuania is one of the key elements in the analysis framework of the macroprudential authority contributing to financial stability in Lithuania.

From the macroprudential perspective, there are three main questions regarding house price developments that are important for warranted and timely policy actions. First, policymakers need to assess whether the house prices are in line with their fundamentals and identify periods of under- or overvaluation of house prices. The second crucial question relates to identification of those fundamentals and other major drivers of house price changes. And lastly, policymakers need to form forecasts of future developments of house prices. In this paper, we address all these questions by undertaking three separate modelling and forecasting exercises. The first exercise employs the vector error correction modelling (VECM) approach to assess under- or overvaluation of house prices. We then use an autoregressive distributed lag model (ARDL) to evaluate which explanatory variables have the biggest impact on house price growth. As the last exercise, we develop a suite of models that are used to forecast the future development of the house price index.

We develop a VECM model, which is used primarily to assess the long-term econometric equilibrium relationship between house prices and other macroeconomic variables (namely, real interest rates, construction costs, credit and investment imbalance<sup>1</sup> and household indebtedness indicator), as well as analyse the impact of shocks to fundamental and non-fundamental variables on the house price dynamics. The model could also be used for house price forecast error decomposition, historical shock decomposition, unconditional and conditional forecasting and other analytical purposes. An alternative, ARDL modelling approach allows us to examine a broader array of indicators that can be useful in explaining house price dynamics. The results reveal that interest rates on new loans to households and building permits are among the key factors for changes in house prices, whereas other variables, such as number of housing transactions, house price to income ratio, etc., have a smaller impact on the dependent variable. In the forecasting exercise, various univariate and multivariate time series models are employed to build a suite of models. For robustness checks and in order to reduce model uncertainty we consider several forecast combination approaches that improve forecasting accuracy compared to individual models. We suggest that the combination of forecasts could be used to predict future changes in the Lithuanian housing price index.

The outline of this paper is as follows: Section 2 provides some theoretical background for housing market modelling. Section 3 contains a brief presentation of relevant empirical literature on house price modelling and forecasting. Section 4 outlines the historical developments of the Lithuanian house price index. Section 5 describes estimation of the fundamental house price. Section 6 provides ARDL modelling results. In Section 7 we describe a suite of models used to forecast the house price index. The last section concludes.

## 2. DISCUSSION ABOUT THE FUNDAMENTALS OF HOUSE PRICE DEVELOPMENT

We begin our analysis by presenting a theoretical framework for housing market modelling. A commonly used theory for the drivers of house prices in the analysis of house prices has been the life-cycle model (see, e.g., Poterba, 1984; Muellbauer & Murphy, 1997; Meen, 2002; Anundsen, 2016). This approach centres on a representative agent that maximises lifetime utility with respect to consumption of housing and composite

---

<sup>1</sup> Ratio of new housing loans to nominal housing investment.

consumption good. The marginal rate of substitution between these two goods satisfies the following equilibrium condition:

$$\frac{U_H}{U_C} = PH_t \left[ (1 - \theta_t)i_t - \pi_t + \delta - \mathbb{E} \left[ \frac{\Delta PH_{t+1}}{PH_t} \right] \right] \quad (1)$$

where  $U_H$  is the marginal utility of housing goods ( $H$ ),  $U_C$  is the marginal utility of composite consumption goods ( $C$ ),  $PH_t$  represents house prices. Time-varying tax rates, mortgage interest rates, and inflation rate are denoted by  $\theta_t$ ,  $i_t$ , and  $\pi_t$ , respectively.  $\delta$  is the housing depreciation rate and  $\mathbb{E}[\Delta PH_{t+1}/PH_t]$  is the expected real capital gain. Equation (1) states that the marginal rate of substitution between housing and other goods is equal to what it costs to own one more unit of a property, measured in terms of forgone consumption of other goods.

Market efficiency also requires the following arbitrage relationship to be satisfied in equilibrium:

$$Q_t = PH_t \left[ (1 - \theta_t)i_t - \pi_t + \delta - \mathbb{E} \left[ \frac{\Delta PH_{t+1}}{PH_t} \right] \right] \quad (2)$$

where  $Q_t$  is the real imputed rent on housing services. This implies that price-to-rent ratio is proportional to the inverse of the user cost:

$$\frac{PH_t}{Q_t} = \frac{1}{UC_t} \quad (3)$$

where  $UC_t$  is the real user cost defined by:  $UC_t = (1 - \theta_t)i_t - \pi_t + \delta - \mathbb{E}[\Delta PH_{t+1}/PH_t]$ . Poterba (1984) interprets relationship (3) as an inverted housing stock demand function. In the literature, unobservable real imputed rent is approximated by observable rent  $R_t$  or it is assumed that it is proportional to income and the housing stock. The first assumption leads to:

$$\frac{PH_t}{R_t} = \frac{1}{UC_t} \quad (4)$$

Instead, it can be assumed that imputed rent is determined by the following expression:

$$R_t = Y_t^{\beta_y} H_t^{\beta_h}, \beta_y > 0, \beta_h < 0 \quad (5)$$

where  $Y$  is households' income and  $H$  is housing stock. This leads to the following equation:

$$\frac{PH_t}{Y_t^{\beta_y} H_t^{\beta_h}} = \frac{1}{UC_t} \quad (6)$$

Equations (4) and (6) are then used as a starting point from which econometric models of house prices are built. In empirical analysis many authors use the semi-logarithmic model to evaluate house price developments:

$$ph_t = \beta_y y_t + \beta_h h_t + \beta_{UC} UC_t \quad (7)$$

where expected signs of the coefficients are:  $\beta_y > 0$  and  $\beta_h, \beta_{UC} < 0$ .

Equation (7) can also include additional determinants of house prices. For example, Anundsen & Jansen (2013) argue that credit is also an important variable, which can determine house price developments. In their paper, the authors use the following expression:

$$ph_t = \beta_y y_t + \beta_h h_t + \beta_{UC} UC_t + \beta_d D_t \quad (8)$$

where  $D$  is households' debt and the expected sign of the coefficient is:  $\beta_d > 0$ .

Equations (7) and (8) can be estimated empirically, which would result in the so-called fundamental house price. Large positive and systematic deviations of actual house prices from the estimated fundamental value can be considered as an indication of unsustainable developments (overvaluation) in house prices.

### 3. DISCUSSION ABOUT HOUSE PRICE MODELLING AND FORECASTING

The empirical literature on house price modelling and forecasting consists of two broad streams. One part of the literature tries to estimate the relationship between house price development and other macroeconomic variables or tries to estimate fundamental house prices. The other part of the literature tries to forecast house price development in the future. Even though authors may use similar models, for example, the vector autoregressive model (VAR), the difference is that they put emphasis either on in-sample fit or on out-of-sample forecasting performance.

Studies examining the fundamental house price usually involve variables from the demand equation, because there is an assumption that the supply of housing is relatively inelastic in the short or medium term. The estimation of the fundamental house price allows one to determine periods of booms and busts in the housing development. In such studies, authors use an error correction model (ECM) or vector error correction model where the level of house prices is often linked to household income, interest rate, mortgage lending, demographic and labour market variables. For example, de Haas & de Greef (2000) use an ECM model for the Dutch housing market and find that house prices depend on bank lending, even controlling for household income, mortgage interest rate, and other important variables. The authors also find that house prices can deviate significantly from their long term fundamental value, and the adjustment process can take a few years. Oikarinen (2008) uses VECM models for Finland's data and also finds that two-way interaction between house prices and credit is very important because this relationship can amplify the boom-bust cycle in the economy and increase instability in the financial sector. Similarly, Greiber & Setzer (2007) observe that excess liquidity can fuel house prices in the US.

Other approaches are also used to estimate boom and bust periods or the deviation from the fundamental house prices. For example, Kajuth et al. (2013) apply panel regression for Germany's administrative districts and take the regression residuals as a measure for deviations of actual house prices from their fundamental equilibrium level. Gerdesmeier et al. (2012) take a less common approach to detect boom and bust episodes in the euro area housing market. The authors apply quantile regression and find that beside variables such as disposable income, user cost rate and unemployment rate, it is useful to include credit variables in the specification as they help explain the boom/bust cycle of house prices. Anundsen (2016) considers four alternative econometric methods (for example, tests based on ADF-test statistic) to construct indicators of housing market imbalances for the US, Finland and Norway. He estimates that only one of the measures indicates imbalances in the Finnish housing market, whereas none of the measures suggest a bubble in Norway.

Some authors in the literature analyse the relationship between house prices and other determinants in order to estimate the response of the house prices to changes in those variables. Katrakilidis & Trachanas (2012) study dependence of house prices on macroeconomic variables in Greece using asymmetric autoregressive distributed lag methodology. Their results reveal significant differences in the response of house prices to positive or negative changes of the explanatory variables in both the short and long term. Many authors apply some kind of VAR type model to assess the impact of monetary policy shock to house prices. Carstensen et al. (2009) analyse 12 European Union countries using a panel VAR model, Eickmeier & Hofmann (2010) and Gupta et al. (2010) employ a factor-augmented VAR (FAVAR) model to the US and South Africa data, respectively, Jarociński & Smets (2008) use Bayesian VAR models for the US data and Robstad (2014) apply such models for the Norwegian data. Most of the authors conclude that house price response to a monetary policy shock is negative and persistent. In addition, Carstensen et al. (2009) show that in countries with a



more pronounced reaction of house prices the propagation of monetary policy shocks to macroeconomic variables is amplified. Meanwhile, Gupta et al. (2010) find that responses are heterogeneous across the affordable, middle and luxury segments of the housing market.

Another broad stream in the literature focuses on the out-of-sample forecasting performance of the models. Authors explore a wide range of models, from univariate models, like the autoregressive moving average model (ARMA) (Barari et al., 2014) or the ARDL model (Rapach & Strauss, 2007), to multivariate models like VAR, BVAR (Das et al., 2011; Gupta & Miller, 2009) or large scale models like dynamic factor models (DFM) (Emiris, 2016), FAVAR models (Gupta et al., 2009) and various other model classes. The choice of variables as predictors for house prices is similar as in the modelling literature, i.e. macroeconomic, monetary, and demographic fundamentals are used in the models. In general, several lessons can be learned from these studies.

Firstly, taking into account additional variables improves the forecasting accuracy of an autoregressive model. For example, an de Meulen et al. (2011) show that ARDL and VAR models with additional macroeconomic variables outperform autoregressive model. Secondly, Gupta et al. (2009) find that a small scale BVAR model is more accurate than a large scale BVAR model. They conclude that only a few fundamental variables are important for house price developments.

Furthermore, allowing estimated model parameters to vary over time gives better forecasting performance. Bork and Møller (2015) and Risse and Kern (2016) use dynamic model averaging (DMA) and dynamic model selection (DMS) and conclude that forecasting accuracy improves substantially when the entire forecasting model is allowed to change over time. Similarly, Barcilar et al. (2015), with a smooth transition autoregressive model (STAR), and Kouwenberg & Zwinkels (2014), with a smooth transition VECM model, show that smooth transition models in an out-of-sample forecasting assessment perform better than competing standard static AR or VECM models.

Lastly, forecast combination can address the issues of unknown data generating process and model uncertainty. Forecast combination is gaining popularity in the empirical studies, where this approach has been found to frequently outperform forecasts from the best-performing model in real time. Such results were found by Drought & McDonald (2011) for New Zealand's, an de Meulen et al. (2011) for Germany's and Rapach & Strauss (2007) for US house prices forecasting exercises.

To summarise, many modern time series modelling and forecasting techniques have been used in the literature to analyse house price development in the past and to predict it in the future. None of the model class appears to be clearly superior to others. In general, the choice of modelling approach and its empirical estimation depends on the research question and on the quality and availability of data.

#### **4. DYNAMICS OF LITHUANIAN HOUSE PRICE INDEX**

The Lithuanian housing market developments and price index dynamics from the 2000s comprise several characteristic phases. A modern housing market, where banks more actively issued mortgage loans and investors began to participate beside individual buyers, emerged around 2000. The initial phase can be considered to span the period until May 2004 when Lithuania joined the European Union. Following this structural change, consumers, businesses and banks began to form very optimistic expectations with regard to economic progress, improvement in living standards and their rapid convergence to the level of incumbent EU member states. High economic growth bolstered by rising household incomes and strong consumer spending, as well as negative real lending rates and large credit flows directed towards the housing market – all of this contributed to a rapid increase in house prices, which lasted until its peak in Q2 2008 (see Fig. 1). During that period, the annual house price growth averaged 35.3%. The highest quarterly and annual growth figures (21.5% and 64.2%, respectively) were recorded in Q4 2005. The bust of the house prices coincided with the global financial crisis, which severely affected credit supply. Concurrently, general economic

conditions worsened significantly, denting affordability of housing and leading to a plummeting demand for housing. The house price trough was reached in Q1 2010, and peak-to-trough decline in house prices amounted to 38.5%. The largest negative quarterly change in house price index was recorded in Q1 2009 and equalled 20.0%.

The post-crisis stagnation phase was between Q1 2011 and Q4 2013, when the house price index increased only by 3%. In Q3 2011, the Bank of Lithuania introduced the Responsible Lending Regulations (RLR), which limited borrowers' loan-to-value (LTV) and debt service-to-income (DSTI) ratios, as well as the maximum loan duration. RLR restricted unsustainable borrowing which could then translate into house price growth. The most recent phase of development of the Lithuanian house price index, or a more pronounced price recovery phase, spanned the period from Q4 2013 until the end of the sample in Q2 2018. During this period, house prices continually exhibited positive annual growth rates.

Fig. 1. Development of the Lithuanian house price index: level (left), quarterly growth (centre), annual growth (right)



Due to the strongly expressed boom/bust cycle of the Lithuanian house price index, quarterly and annual growth rates are very volatile which makes modelling and forecasting quite challenging. During the sample period, i.e. from Q1 2002 to Q2 2018, the average quarterly growth rate was equal to 2.2% with a standard deviation of 6.3 p.p. At the same time, the average annual growth rate was equal to 10.4% with a standard deviation of 18.9 p.p. However, after the crisis volatility in the data has considerably decreased. Calculating quarterly and annual growth rates from Q1 2011 to the end of the sample we get the average quarterly growth of 1.2% (with standard deviation of 1.7 p.p.) and the average annual growth of 4.8% (with standard deviation of 3.4 p.p.). High volatility and a change in it could have an impact on confidence intervals of the estimated model coefficients and forecasts.

## 5. FUNDAMENTALS OF THE LITHUANIAN HOUSE PRICE INDEX: VECM ANALYSIS

In this section we develop a VECM model to assess the cointegrating relationship (or a long-term econometric equilibrium relationship) between house prices and other macroeconomic variables, as well as analyse the impact of shocks to fundamental and non-fundamental variables on house price dynamics. The model could also be used for house price forecast error decomposition, historical shock decomposition, unconditional and conditional forecasting and other analytical purposes.

## 5.1. ECONOMETRIC MODEL

To estimate the VECM model we employ the standard Engle-Granger procedure (Engle & Granger, 1987). First, the cointegrating relationship among a number of  $I(1)$ , i.e. integrated of order one, variables is estimated by ordinary least squares (OLS):

$$y_{1t} = \beta_1 + \beta_2 y_{2t} + \dots + \beta_n y_{nt} + \varepsilon_t \quad (9)$$

Here  $y_{1t}$  is the housing price variable, other  $y$ 's denote other macroeconomic variables and  $\varepsilon_t$  is a stationary error term. Using residuals  $\hat{\varepsilon}_t$  from the estimated cointegrating relationship in dynamic equations of individual variables, one can formulate the error correction model. In matrix notation, it can be written as follows:

$$\Delta y_t = \alpha \hat{\varepsilon}_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \epsilon_t \quad (10)$$

where  $y_t = (y_{1,t}, y_{2,t}, \dots, y_{n,t})$  is a  $n \times 1$  vector of endogenous variables and  $\alpha$  denotes a commensurately sized vector of adjustment parameters measuring how individual variables respond to last period's deviations from the equilibrium.  $\Gamma_i$  is an  $n \times n$  coefficient matrix associated with  $i$ -th lag. Parameter  $p$  denotes the number of lags in the model's VAR form. Model variables and the number of lags are chosen so as to ensure that model disturbances  $\epsilon_t$  follow a multivariate normal distribution. Finally,  $\Delta$  denotes a difference operator.

## 5.2. MODEL VARIABLES AND DATA

Taking the inverse housing demand function as a starting point for our empirical analysis of house price determination, we expect house prices to be positively linked to rent prices, inflation and expectations of future growth of house prices and negatively linked to interest rates, as well as maintenance and depreciation rates.

Though this inverse housing demand function, or housing user cost approach, forms the basis of much of the empirical research on house price fundamentals, there are some caveats associated with this approach and with interpretation of this relationship as fundamentals. First, "explaining" high house prices by expensive rents may be misleading because rents can also be procyclical and can strongly decline during a downturn in the housing market. Notably, this has been the case in Lithuania.

Also, the formula suggests that high house prices can be justified by expected future house price growth. Given the element of reflexivity in this formulation, one has to be careful and distinguish whether these expectations are rational and reflect underlying structural developments or they are not fully rational and reflect subjective market sentiment. In fact, one of the most popular definitions of a house price bubble, formulated by Stiglitz (1990), suggests that a bubble exists if the reason that the price is high today is only because investors believe that the selling price will be high tomorrow – when "fundamental" factors do not seem to justify such a price. In this light it is logical in the empirical analysis to replace the unobserved and somewhat vague expectations variable with structural imbalances and other factors that are likely to drive house price growth. At the same time, house price inertia, adaptive expectations and bubble formation tendencies can be modelled by including house price lags in the model.

Finally, in its pure theoretical form, the housing demand function leaves out credit. In reality, however, house prices and credit are strongly intertwined and it can be difficult to reasonably explain house price dynamics without taking into consideration developments in the credit market. So, if the aim is to explain actual house price dynamics using an empirical model, one should try to incorporate some credit-related variables (even though they are usually regarded as non-fundamental) in the model in addition to those that can be thought of as fundamentals.

Because of the above-mentioned caveats and data issues, in our empirical analysis we had to depart quite far from the initial theoretical formulation (equation 7). The analysis was affected by the usual problem of short data series. Quarterly data spanned the period from Q4 2004 to Q2 2018 (a total of 55 observations), which

essentially included one very significant boom-bust cycle. Against the backdrop of huge cyclical variation, there were cases of very strong covariation between house prices and some other variables (e.g. rent prices), rendering a number of other reasonable candidate variables insignificant. With this in mind and given the lack of sufficient quality data on rent prices in Lithuania, we tried to proxy the rent price variable by demographic, household income, housing stock variables and other indicators that could potentially affect conditions in the rental market. However, none of them appeared significant and were therefore left out of the analysis. We tried to proxy unobserved expectations of house price dynamics by consumer and business confidence variables, construction price index, household indebtedness, variables reflecting imbalances between financial capital (housing loans) and physical investment, etc.

The working version of the VECM model includes five variables:

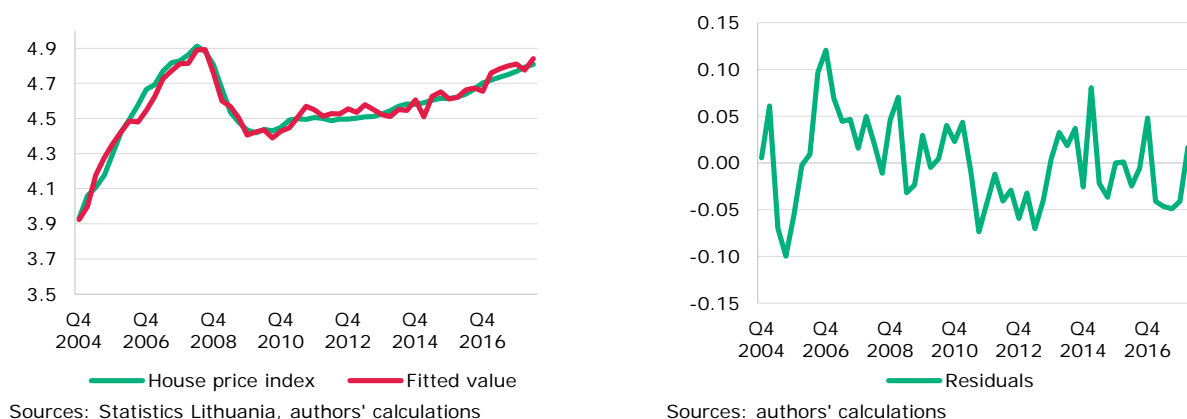
- House price index (denoted as hpi)
- Real (HICP inflation-adjusted) lending rate on housing loans ( $r$ )
- Construction cost price index (ccpi)
- New housing loans to nominal housing investment ratio (credInv)
- Existing housing loans to nominal GDP ratio (debtRatio)

Data series are seasonally adjusted using Census X-13 seasonal adjustment procedure. All variables, except for the real loan rate, enter the model in logs. Application of the augmented Dickey-Fuller test confirms that all variables are nonstationary and integrated of order one.

### 5.3. MODEL RESULTS

As the first stage of the Engle-Granger procedure, we estimate the cointegrating relationship among selected variables (see Fig. 2). Applying the Engle-Granger test, the null hypothesis of no cointegration can be rejected at 10% significance level. Coefficient signs are as expected: house prices are negatively linked to real lending rates and positively related to all other variables in the cointegrating relationship, indicating that higher house prices are positively associated with higher construction costs, stronger credit flows and larger household indebtedness (see Table A1, Appendix A). The long-run sensitivity of house prices to real lending rates is moderate: a 1 p. p. increase is associated with a 1.1% decline in house price over the long run. Notably, house prices are found to be very sensitive to increases in construction costs, with elasticity of 1.54.

Fig. 2. Cointegrating relationship: house price index and the fit from the estimated long-run relation (left) residuals from cointegrating relationship (right)



Various model selection criteria (Akaike information criterion, Bayesian information criterion and the likelihood ratio test) suggest that the optimal number of lags to include in the VAR form of the model is two, which is equivalent to one lag in the VECM form. However, the Ljung-Box test indicates that residuals of the

house price equation in the resulting VECM model are autocorrelated. Attempting to tackle the problem but at the same time avoiding the need to estimate too many parameters associated with inclusion of additional lags, we tried adding individual lagged (exogenous or endogenous) variables to the model. The residual autocorrelation problem was corrected when we included the fourth lag of the household debt ratio. Thus, our working model is:

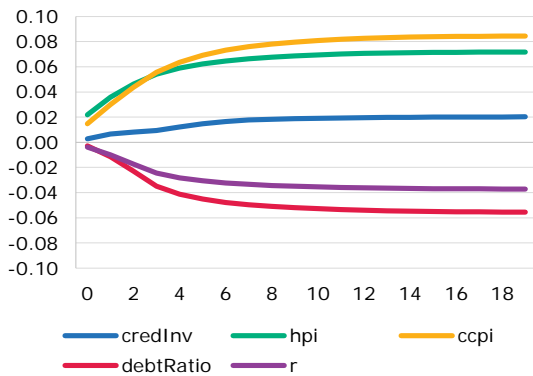
$$\Delta \mathbf{y}_t = \alpha \hat{\boldsymbol{\epsilon}}_{t-1} + \Gamma_1 \Delta \mathbf{y}_{t-1} + \boldsymbol{\gamma} \Delta \text{debtRatio}_{t-4} + \boldsymbol{\epsilon}_t \quad (11)$$

where  $\mathbf{y}_t = (hpi_t, credInv_t, debtRatio_t, ccpi_t, r_t)$  and  $\boldsymbol{\gamma}$  is the coefficient vector associated with the fourth lag of *debtRatio* variable. Since this variable is endogenous, we actually have a restricted VECM model – term  $\boldsymbol{\gamma} \Delta \text{debtRatio}_{t-4}$  can be equivalently expressed as  $\Gamma_4 \Delta \mathbf{y}_{t-4}$ , where the third column of parameter matrix  $\Gamma_4$  equals  $\boldsymbol{\gamma}$  and all other matrix elements are zero.

Estimation results show that house prices are modestly pulled to the long-term equilibrium as they tend to correct around 9% of the deviation from the long-term equilibrium in the preceding quarter. The value of the error correction parameter is in line with typical values reported in the literature on house price fundamentals (where they range from -0.07 to -0.25), though in our case the parameter is not statistically significant. In contrast, the credit-investment imbalance variable (*credInv*) is strongly pushed away from the equilibrium values, suggesting that this imbalance is a significant source of instability in the system. The result is logical, because historically, as Lithuania's house price bubble inflated, housing loans rose faster than investment and during the bubble deflation phase a drop in new housing loans was stronger than a decrease in housing investment. One more thing to note is that in the house price equation there is a strong autoregressive element – the positive coefficient associated with the lagged change in house prices, sometimes in the literature referred to as the “bubble generator” parameter, is large and significant, indicating inertia of house price movements and arguably a strong role of adaptive market expectations.

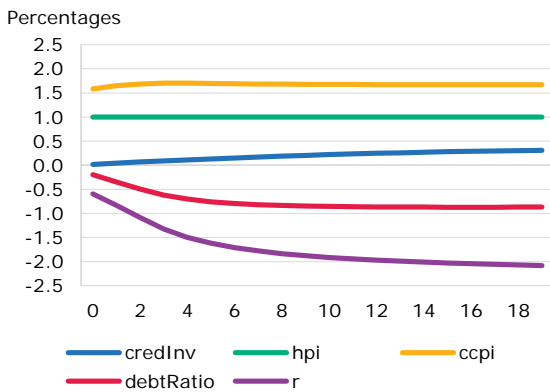
The Granger representation theorem implies that our VECM model can be expressed as a nonstationary reduced-form VAR. This enables us to calculate impulse response functions. Generalised impulse responses of the house price variable to one standard deviation shocks to each variable are shown in Fig. 3. The house price variable is strongly positively affected by shocks to construction costs and the house price itself. The house price response to a change in credit-investment imbalance in the housing market is moderately positive. House prices negatively react to real lending rate and indebtedness shocks. The impulse responses can also be transformed into various multipliers. For example, the cumulative multiplier can be calculated as a cumulative impulse response of the response variable divided by a cumulative response of the shock variable to a shock to itself. As can be seen from Fig. 4, a 1% increase in construction costs is associated with a 1.6% increase in house prices, whereas a 1 p.p. rise in real lending rates leads to a 2% decline in house prices over the long run.

Fig. 3. Generalised impulse responses of house prices to one standard deviation shocks



Sources: authors' calculations

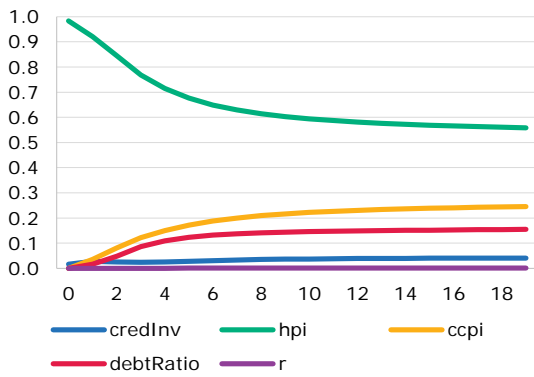
Fig. 4. Cumulative multipliers of house prices



Sources: authors' calculations

In order to identify structural (orthogonalised) shocks, we apply Cholesky decomposition procedure, which requires ordering variables from most endogenous to most exogenous. Based on the Granger causality analysis and economic considerations, we put the system's variables in the following order: *credInv*, *hpi*, *ccpi*, *debtRatio*, and *r*. Structural shock identification enables us to perform shock decomposition analysis. The forecast error variance decomposition reveals the relative importance of various structural shocks in determining stochastic variation in house prices. Fig. 5 shows that in the very short term the variation in house prices is driven predominantly by structural shocks to the house price variable itself, whereas in the long run a significant fraction of house price variation is explained by structural shocks to construction costs and household indebtedness.

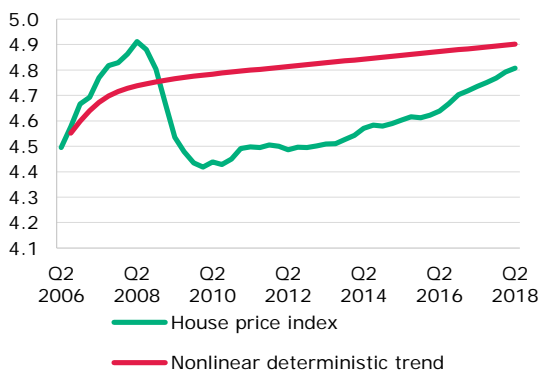
Fig. 5. House price forecast error variance decomposition



Sources: authors' calculations

Historical shock decomposition allows us to decompose the dynamics (more specifically, the change from the values observed at the start of the sample) of a given variable into a deterministic (nonlinear trend) component and the cumulative impact of individual structural shocks. The deterministic components of each variable represent the system's dynamics in the absence of structural shocks, so they approximate the model's long-term econometric equilibrium path. It is not the equilibrium or sustainable growth path in a strict economic sense because the trend depends on initial conditions and also on the "nonfundamental" variables included in the system. Nevertheless, the estimates of such trends reflect the variables' "balanced" paths with regard to the rest of the system – the variables are beaten off of these paths by structural shocks to the system. As can be seen from Fig. 6, in 2006–2008 house prices rose above their balanced path, then in 2008 and 2009 abruptly fell well below it and starting from 2012 started gradually closing the gap from below. According to the model, house prices were still slightly below their econometric balanced path at the end of the sample, in 2018.

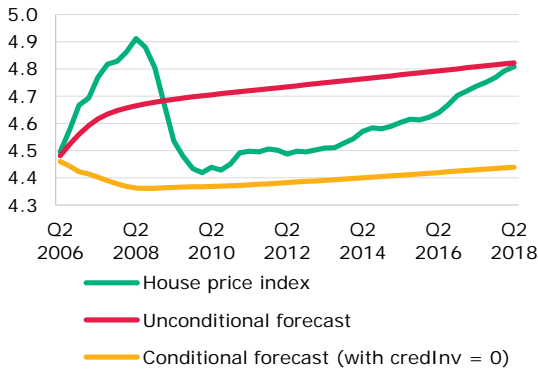
Fig. 6. Deterministic component (nonlinear trend) of house price dynamics



Sources: Statistics Lithuania, authors' calculations

The model can also be directly applied for forecasting house prices (and other model variables), adding to the suite of specialised forecasting models described in the sections below. Notably, the model can be used for conditional forecasts and counterfactual analysis, which allows us to analyse the dynamics of the system subject to externally imposed restrictions on paths of selected variables. In this regard, we are specifically interested in assessing the impact of unbalanced credit and investment dynamics on house prices. If we assume a balanced credit and investment dynamics (set *creditInv* equal to zero), we get the result that excess credit flows may have raised house prices over the analysed period by about 40% (see Fig. 7).

Fig. 7. Unconditional and counterfactual house price trends



Sources: Statistics Lithuania, authors' calculations

## 6. DETERMINANTS OF THE HOUSE PRICE GROWTH AND SHORT TERM FORECASTING: ARDL MODEL

In this section, we analyse the potential factors that determine the dynamics of Lithuanian house prices. We employ an ARDL model to evaluate house price changes and to make short term forecasts (up to four quarters). In this section, we follow the work of Rapach & Strauss (2007), who model and forecast the US house price dynamics.

### 6.1. ECONOMETRIC MODEL

Let  $\Delta y_t = y_t - y_{t-1}$ , where  $y_t$  is the log level of Lithuanian house price index at time  $t$ . Then the growth rate of house price index from time  $t$  to  $t + h$  ( $h$  is the forecast horizon, which is four quarters in this exercise)  $y_{t+h}^h$  is defined as follows:

$$y_{t+h}^h = \frac{1}{h} \sum_{j=1}^h \Delta y_{t+j} \quad (12)$$

Furthermore, let  $x_{i,t}$  ( $i = 1, \dots, n$ ) represent potential determinants of house price growth rate. An ARDL model can be written as:

$$y_{t+h}^h = \alpha + \sum_{j=0}^p \beta_j \Delta y_{t-j} + \sum_{j=0}^q \gamma_j x_{i,t-j} + \varepsilon_{t+h}^h \quad (13)$$

where  $\varepsilon_{t+h}^h$  is an error term. In this exercise, the ARDL model parameters were estimated using OLS and the model is used in two ways. First, we estimate the model with a full data sample in order to determine which variables made the biggest impact on house price growth. Second, equation 13 is used to construct a set of recursive (expanding estimation window) out-of-sample forecasts of  $y_{t+h}^h$  using information available at time  $t$ .

Due to short time series, the lag length  $p$  is set to 0, i.e. only  $\Delta y_t$  is included in the equation. The lag length  $q$  can vary between 0 and 2. Furthermore, the number of determinants  $x_{i,t}$  can also vary between one to four predictors per model. In this way, we estimated a few thousand different ARDL models.

### 6.2. EXPLANATORY VARIABLES

Data availability in this modelling exercise determined the data sample which covers the period from Q2 2006 to Q2 2018. We include 12 potential explanatory variables of house price growth rate which could be attributed to three main groups. Demand-related variables provide measures of the ability of households to purchase housing:



- Population
- Unemployment rate
- Wages
- Consumer price index
- Consumer confidence index
- House price to income ratio

Supply-related variables may reflect supply conditions in the housing market with a potential impact on house prices:

- Building permits
- Investment in real estate
- Construction cost index
- Number of housing transactions

Meanwhile, bank credit related variables show households' access to credit for housing purchases:

- Interest rate on new loans for households
- New loans for households

In this modelling exercise the house price index and some of the explanatory variables were seasonally adjusted: wages, building permits, investment in real estate, construction cost index, number of housing transactions. Furthermore, all of the variables were transformed by taking first the differences of log levels to make them stationary, with the exceptions to unemployment rate, house price to income ratio and interest rate (which were taken in levels) and building permits, number of housing transactions and new loans for households (which were taken in log levels).

In order to reduce the number of ARDL models, sign restrictions were applied on the estimated coefficients. A positive sign restriction was imposed on population, wages, consumer price index, consumer confidence index, construction cost index and new loans for households. A negative sign restriction was imposed on unemployment rate, building permits, investment in real estate and interest rate on new loans to households. The number of housing transactions and house price to income ratio were left unrestricted.

### 6.3. MODEL EVALUATION AND COMBINATION

All ARDL models were evaluated by their performance in various model accuracy testing procedures. First of all, we estimated all models using the full data sample and kept only those models that have coefficients in accordance with the sign restrictions mentioned above. Then the in-sample fit was examined by the mean absolute scaled error (MASE), sign accuracy, and p-value of the coefficients. Introduced by Hyndman & Koehler (2006), MASE is the mean absolute error of the fitted values, divided by the mean absolute error of the one-step naive forecast:

$$MASE = \frac{1}{T} \sum_{t=1}^T \left( \frac{|y_t - f_t|}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|} \right) \quad (14)$$

where  $y_t$  is actual data and  $f_t$  is fitted values. In our further analysis we kept models where  $MASE < 1$ , which indicates a better accuracy compared to a simple one-step naive forecast. Next, we compared the sign of the fitted value and actual data, i.e. we compared whether the growth of house prices or decline of it was captured. We put the restriction that sign accuracy is greater or equal to 0.75, which means that at least three times out of four, models correctly predicted the direction of the development. Lastly, we kept only those models for which the p-value of the least significant coefficient was lower than 0.5.

Another model evaluation criterion was based on model performance during a crisis period. We re-estimated models from the start of the sample until Q3 2008 and forecasted four quarters ahead. Then we examined coefficient sign accuracy and retained only those models, which predicted decline in house prices at least in one quarter.

The last model evaluation criterion was based on model performance in the out-of-sample forecasting exercise. Initially, models were estimated using the data from the training sample, i.e. from Q2 2006 until Q4 2011. Then the short-term (i.e. four quarters ahead) forecasts were calculated. All models were then re-estimated on an incrementally expanding (by one quarter) data window to produce a new set of forecasts. The expanding window estimation was repeated until Q2 2017. Once again, we kept only those models that had the average forecast sign accuracy greater than 0.5, i.e. on average at least two times out of four correctly predicted the direction of the development.

Only models that passed all accuracy testing procedures were used in the further analysis. Since it is not obvious which individual model is best suited for the analysis and short term forecasting, we applied model combination, which provided a way to incorporate information from several models. The model combination allows having more variables as explanatory variables compared to a single model and as noted by Rapach & Strauss (2007) helps improve forecasting accuracy.

Individual ARDL models were combined into one model based on the out-of-sample average root mean squared errors (RMSE). More specifically, the weight of the model is calculated as follows:

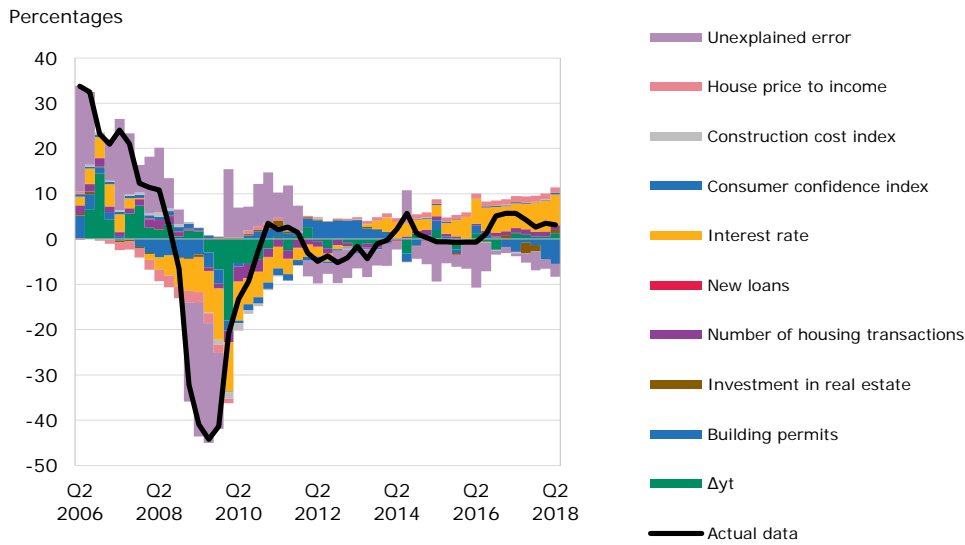
$$w_i = \frac{(1 - RMSE_i)}{\sum_{i=1}^M (1 - RMSE_i)} \quad (15)$$

where  $w_i$  is the weight of the individual model,  $RMSE_i$  is the average out-of-sample RMSE of the model, and  $M$  is the number of models. In this way, the model with the lowest forecast errors receives the highest weight. A combination of the in-sample fitted values gives insights to which variables had the biggest impact on house price growth in the past. Meanwhile, a combination of out-of-sample forecasts helps to explain which explanatory variables will affect future development.

#### 6.4. MODELING RESULTS

Empirical estimation provided 18 individual ARDL models that passed all model evaluation criteria. The results from the estimated models were then aggregated into one weighted ARDL model. Fig. 8 shows which explanatory variables had the biggest impact on the deviation from the long term growth of house prices. Four variables were not included in any of the individual models: consumer price index, wages, population and unemployment rate. The figure shows that interest rate is one of the main drivers of house price developments. The low interest rate environment mainly caused the recent increase in house price growth. From the supply side variables, building permits had the biggest impact in the past. All other variables played a lesser role in affecting the dependent variable. The figure also reveals that quite a big part of the house price changes, especially during the crisis, is still left unexplained by the model. This may suggest that the actual decrease in growth could be attributed to changes in perception of the house value which could not be captured by any explanatory variables. Another possible explanation could be that the relationship between house price growth and other variables is more contemporaneous. However, given the forecasting purpose of the model, we could not include higher lags of the explanatory variables in the model.

Fig.8. Determinants of the house price growth in the past



Sources: Statistics Lithuania, authors' calculations

The combined ARDL model was also used to produce short term out-of-sample forecast. Fig. B1 (in Appendix B) shows the forecasts at different forecasting horizons. The results show that the model tends to overpredict the growth of house prices at the beginning of the forecasting sample. This result may come from the fact that data showed high volatility before and during the crisis, which may affect the size of the coefficients in the individual ARDL models. Nevertheless, at the end of the sample, combined ARDL forecasts are more accurate. Moving forward, we can expect better forecasting performance as a longer data sample will give better estimates of the model coefficients.

## 7. SHORT AND MEDIUM TERM FORECASTING: FORECAST COMBINATION

In this section, we present a suite of models aimed at forecasting the Lithuanian house price index in the short and medium term (up to ten quarters). We explore univariate and multivariate time series models and compare their performance by out-of-sample forecasting accuracy. Thus, this modelling exercise follows a stream of literature that focuses on the forecasting performance of the models.

### 7.1. ECONOMETRIC MODELS

The starting point of our forecasting exercise is a simple naïve forecast of the log level of the house price index. Thus, our benchmark model assumes that during the forecasting period the level of house price index will remain the same as the last observed value. All other models will be compared to this benchmark model.

#### ARMA

In a univariate case, the ARMA model is used, which is the most common empirical approach in modelling and forecasting time series data. The classical  $ARMA(p, q)$  model, where  $p$  and  $q$  denote the AR and MA orders, can be written as:

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + e_t + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} \quad (16)$$

where  $y_t$  is a variable of interest,  $c$  is a constant,  $e_t$  is an error term,  $\phi_1, \dots, \phi_p$  and  $\theta_1, \dots, \theta_q$  are parameters of the model. The model is fitted using the maximum likelihood estimation method.

The house price index is forecasted taking the AR order from 1 to 4 ( $p = 1, 2, 3, 4$ ), and the MA order from 0 to 2 ( $q = 0, 1, 2$ ). Furthermore, each ARMA model specification is estimated taking house price index in log level,

quarterly growth rate, and annual growth rate. Overall, 36 different ARMA models were estimated and included in the suite of models.

## VAR

Gupta et al. (2009) and an de Meulen et al. (2011) showed that including additional macroeconomic variables improves forecasting accuracy compared to autoregressive models. Therefore, VAR-type models are included in the suite of models for forecasting house price index.

The classical VAR model with  $n$  endogenous variables and  $p$  lags was suggested by Sims (1980) and can be written as:

$$Y_t = B_0 + B_1 Y_{t-1} + B_2 Y_{t-2} + \dots + B_p Y_{t-p} + u_t \quad (17)$$

where  $Y_t = (y_{1,t}, y_{2,t}, \dots, y_{n,t})$  is a  $n \times 1$  vector of endogenous variables,  $B_0$  is a  $n \times 1$  vector of constant terms,  $B_1, B_2, \dots, B_p$  are  $n \times n$  matrices of coefficients and  $u_t$  is a  $n \times 1$  vector of residuals, which follows a multivariate normal distribution, i.e.  $u_t \sim \mathcal{N}(0, \Sigma)$ . Estimation of the model is done by the ordinary least squares (OLS) and forecasting is straightforward (see, e.g., Lütkepohl, 2005).

The house price index is forecasted using seven alternative VAR model specifications (see Table 1). Due to data limitations and following Gupta et al. (2009), who found that small scale VAR-type models are more accurate than large scale models, we explore different specifications which include from four to six variables in the model. All specifications follow the same structure: the house price index is modelled with interest rate, lending to households, macroeconomic and labour market variables.

Table 1. Estimated VAR model specifications

Specification	Variables					
Spec. 1	Real GDP	House price index	Credit to households	Interest rate on loans to households		
Spec. 2	Real GDP	House price index	Credit to households	Interest rate on loans to households	HICP	
Spec. 3	Real GDP	House price index	Credit to households	Interest rate on loans to households	HICP	Unemployment rate
Spec. 4	Disposable income	House price index	Credit to households	Interest rate on loans to households	HICP	
Spec. 5	Disposable income	House price index	Credit to households	Interest rate on loans to households	HICP	Unemployment rate
Spec. 6	Real private consumption	House price index	Credit to households	Interest rate on loans to households	HICP	
Spec. 7	Real private consumption	House price index	Credit to households	Interest rate on loans to households	HICP	Unemployment rate

As in the case of autoregressive models, each VAR model specification is estimated taking endogenous variables in log levels (except interest rate and unemployment rate which are taken in percentages), quarterly growth rates (interest rate and unemployment rate are taken in quarterly changes) or annual growth rates (interest rate and unemployment rate are taken in annual changes). Furthermore, each specification is estimated taking the number of lags from 1 to 4, i.e.  $p = 1, \dots, 4$ . Overall, 84 different VAR models were estimated and included in the suite of models.

## BVAR

Even though a VAR model is one of the main tools for macroeconomic modelling and forecasting, the estimation of these models poses certain challenges. The VAR model requires many parameters to be estimated because typically equal lag lengths for all variables are used, including many that are statistically insignificant. The rich parameterisation of VAR models (over-parameterisation problem) brings a risk of overfitting the data, inefficient estimates and possibly large out-of-sample forecasting errors. Consequently, this limits the endogenous variables and the number of lags, which can be included in the simple VAR model to avoid those problems. Litterman (1986) and many other authors used the Bayesian estimation of the VAR model (BVAR) to overcome the over-parameterisation problem. See, for example, Karlsson (2013) for extensive literature on forecasting using Bayesian VAR models. The Bayesian approach suggests a solution to this curse of dimensionality by introducing prior distributions on parameters of interest. According to Karlsson (2013), BVAR models usually show better forecasting performance compared to VAR models estimated with frequentist techniques.

In Bayesian econometrics, every parameter of interest is treated as a random variable, characterised by some underlying probability distribution. In this study, we employ widely used Minnesota prior distributions (Litterman, 1986). Instead of eliminating lags, the Bayesian estimation imposes restrictions on the coefficients across different lag lengths. Minnesota priors take the fact that more recent values of a data series usually contain more information about the current value of the series than past values. Furthermore, it also takes the fact that past values of a given variable contain more information about its current state than past values of other variables. In this framework, it is assumed that the VAR residual variance-covariance matrix  $\Sigma$  is known and the only object left to estimate is the vector of coefficients  $\beta$ . It is assumed that  $\beta$  follows a multivariate normal distribution:

$$\beta_{ii} \sim \mathcal{N}(1, \sigma_{\beta_{ii}}^2) \text{ and } \beta_{ij} \sim \mathcal{N}(0, \sigma_{\beta_{ij}}^2) \quad (18)$$

where  $\beta_{ii}$  denotes the coefficients associated with the lagged dependent variable in each equation of the VAR model,  $\beta_{ij}$  represents any other coefficient. The prior variances  $\sigma_{\beta_{ii}}^2$  and  $\sigma_{\beta_{ij}}^2$  specify uncertainty about the prior means. Based on Litterman (1986), for coefficients relating endogenous variables to their own lags, the variance is given by:

$$\sigma_{\beta_{ii}}^2 = \left( \frac{\lambda_1}{l^{\lambda_3}} \right)^2 \quad (19)$$

where  $\lambda_1$  is an overall tightness parameter,  $l$  is the lag considered by the coefficient, and  $\lambda_3$  is a scaling parameter controlling the speed at which coefficients for lags greater than 1 converge to 0 with greater certainty.

For parameters related to cross-variable lag coefficients, the variance is given by:

$$\sigma_{\beta_{ij}}^2 = \left( \frac{\sigma_i^2}{\sigma_j^2} \right) \left( \frac{\lambda_1 \lambda_2}{l^{\lambda_3}} \right)^2 \quad (20)$$

where  $\sigma_i^2$  and  $\sigma_j^2$  denotes the OLS residual variance of the autoregressive models estimated for variables  $i$  and  $j$ , and  $\lambda_2$  represents a cross-variable specific variance parameter.

For exogenous variables (including constant terms), the variance is given by:

$$\sigma_{c_i}^2 = \sigma_i^2 (\lambda_1 \lambda_4)^2 \quad (21)$$

where  $\lambda_4$  is a large variance parameter. In this study, typical values found in the literature for parameters  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$  and  $\lambda_4$  are used, i.e.  $\lambda_1 = 0.1$ ,  $\lambda_2 = 0.5$ ,  $\lambda_3 = 1$  and  $\lambda_4 = 10^2$ .

Empirical BVAR model estimation follows the same strategy as VAR model estimation. Thus, we take the same seven model specifications and estimate them taking the number of lags from one to four. Overall, 84 different BVAR models were estimated and included in the suite of models.

### BVAR WITH STOCHASTIC VOLATILITY

Several authors noted (e.g. Kouwenberg & Zwickels, 2014) that allowing time-variation in parameters may improve house price forecasting performance. Furthermore, Chan & Eisenstat (2018) analysed US macroeconomic data and found strong support for the time-varying parameter VAR with stochastic volatility (TVP-VAR-SV) compared to a constant coefficient VAR with homoscedastic innovations. However, the authors argue that most of the gains come from allowing for stochastic volatility rather than time variation in the VAR coefficients. Following these results in our analysis, we apply the BVAR model with stochastic volatility (BVAR-SV) which was initially proposed by Cogley & Sargent (2005) and Primiceri (2005). Thus, we allow for stochastic volatility but BVAR coefficients are left constant.

In a traditional VAR model it is assumed that the residuals are distributed according to multivariate normal distribution, i.e.  $u_t \sim \mathcal{N}(0, \Sigma)$ . Meanwhile, in the BVAR-SV model, it is assumed that the residuals are independently but not identically distributed across time. Their variance-covariance matrix  $\Sigma$  is allowed to be time varying, hence providing stochastic volatility and introducing heteroscedasticity.

It is assumed that  $\Sigma_t$  can be decomposed into:

$$\Sigma_t = F^{-1} \Lambda_t F^{-1'} \quad (22)$$

where  $F^{-1}$  is a lower triangular matrix with ones on the diagonal and non-zero coefficients below the diagonal. Meanwhile,  $\Lambda_t$  is a period-specific diagonal matrix of variances, i.e.

$\text{diag}(\Lambda_t) = (\bar{s}_1 \exp(\lambda_{1,t}), \bar{s}_2 \exp(\lambda_{2,t}), \dots, \bar{s}_n \exp(\lambda_{n,t}))$ .  $\bar{s}_1, \bar{s}_2, \dots, \bar{s}_n$  are known scaling terms and  $\lambda_{1,t}, \lambda_{2,t}, \dots, \lambda_{n,t}$  are dynamic processes generating the heteroscedasticity of the model. It is assumed that they are characterised by the autoregressive process:

$$\lambda_{i,t} = \gamma_i \lambda_{i,t-1} + \varepsilon_{i,t}, \varepsilon_{i,t} \sim \mathcal{N}(0, \varphi_i) \quad (23)$$

The parameters of interest to be estimated are: the VAR coefficients  $\beta$ , the elements  $f^{-1} = \{f_i^{-1}; i = 2, \dots, n\}$  related to the  $F^{-1}$  matrix, the set of dynamic coefficients  $\lambda = \{\lambda_{i,t}; i = 1, \dots, n; t = 1, \dots, T\}$ , the set of autoregressive coefficients  $\gamma = \{\gamma_i; i = 1, \dots, n\}$  and the heteroscedasticity parameters  $\varphi = \{\varphi_i; i = 1, \dots, n\}$ .

Once again, the empirical BVAR-SV model estimation follows the same strategy as VAR model estimation. Thus, we take the same seven model specifications and estimate them taking the number of lags from one to four. Overall, 84 different BVAR-SV models were estimated and included in the suite of models.

### 7.2. ESTIMATION OF FORECAST ACCURACY

In the forecasting exercise, we used the house price index and other Lithuanian data spanning from Q1 2002 to Q2 2018 ( $T = 66$ ). To compute out-of-sample forecasts, the data sample was split into an in-sample period ( $I = 40$ ) and an out-of-sample period ( $O = 26$ ). With a forecast horizon of ten quarters, this allowed evaluating 17 forecasts ( $N = 17$ ). At the start, all models were estimated using data until Q4 2011 and up to ten quarters ahead forecasts of log level of the house price index were made. All models were then re-estimated by expanding the training data set by one quarter to produce a new set of forecasts. The expanding window estimation was repeated until Q4 2015, which was the final estimation period.

The forecasting performance was assessed using several standard forecast evaluation measures. The main measure of forecasting accuracy in this exercise was the root mean squared error (RMSE) of forecasts at particular period ahead, which is defined as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (y_{n,h} - f_{n,h})^2} \quad (24)$$

where  $f_{n,h}$  is the h-step ahead forecast,  $y_{n,h}$  is the actual value at that step, and  $N$  is the number of forecasts. The forecast accuracy was compared at one-, four-, eight- and ten-quarter-ahead RMSEs across all models. In general, the lower RMSE indicates the better forecasting accuracy of the model.

Another measure for forecasting accuracy is the mean absolute forecast error (MAE) which is defined as:

$$MAE = \frac{1}{N} \sum_{n=1}^N |y_{n,h} - f_{n,h}| \quad (25)$$

In this exercise we calculate the relative absolute error (RMAE):  $RMAE = MAE / MAE^{Naive}$ , where  $MAE^{Naive} = \frac{1}{N} \sum_{n=1}^N |y_{n,h} - f_{n,h}^{Naive}|$  is the MAE of the our benchmark naïve forecast. RMAE smaller than one indicates a better forecasting performance compared to a simple naïve forecast.

The forecasting accuracy was calculated for the log level of the house price index. If the model used quarterly growth rates or annual growth rates of the house price index then forecasts of those models were recalculated to the log level of the index and only then forecasting accuracy was evaluated.

### 7.3. FORECASTING PERFORMANCE

Estimated forecasting accuracy is presented in Appendix C, where tables report RMSE of the individual models at different forecasting horizons. Results in the tables also indicate which models had RMAE lower than one, i.e. which models performed better than a benchmark naïve forecast. Several observations can be made from this forecasting exercise.

First of all, model accuracy depends on the forecasting horizon. If we consider one-quarter-ahead forecasts, only a few BVAR and BVAR-SV models estimated in log levels outperformed a simple naïve forecast. In longer forecasting horizons, many more models were able to give better accuracy.

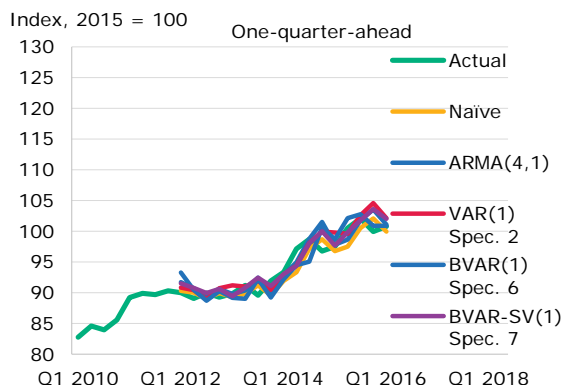
The results also show that univariate models, in general, give better results than VAR models. Especially models estimated in quarterly growth rates, which constitute one of the best model classes in this exercise. Only some of the BVAR and BVAR-SV models outperformed univariate models. Thus, additional variables may improve forecasting performance, but we need Bayesian estimation to exploit that gain.

Relatively short time series may be one of the reasons why VAR models (especially with a higher number of lags) provided poorer forecasting accuracy. A large number of parameters that need to be estimated caused large out-of-sample forecasting errors. Our results are in line with Karlsson (2013), who showed that BVAR models usually show better forecasting accuracy compared to VAR models. Furthermore, the results show that allowing stochastic volatility enhances forecasting performance of the models estimated in log levels. But it seems that there are no gains if we model house price index by taking the quarterly or annual growth rate.

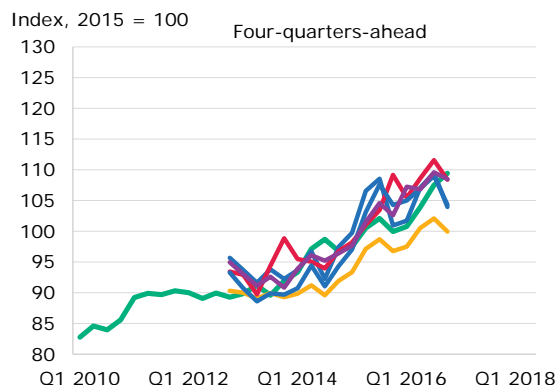
If we consider different multivariate model specifications, then specifications 2 and 6 estimated in log levels provided better forecasting accuracy than other specifications. Thus, it seems that the unemployment rate did not provide additional information. The results also show that multivariate models with one lag outperform models with a higher number of lags. It seems that additional lags are associated with additional parameters to be estimated and this leads to larger out-of-sample forecasting errors.

Fig. 9 shows forecasts of the individual models that performed best (in terms of RMSE) in different model classes. In the univariate case, the best performing model was ARMA(4,1) estimated on quarterly growth rates. The lowest RMSEs between VAR-type models were achieved by VAR(1) Spec. 2 (taking log levels), BVAR(1) Spec. 6 (taking annual growth rates) and BVAR-SV(1) Spec. 7 (taking log levels). Thus, we see that depending on the model class, different data transformation or multivariate model specification would be preferred for the purposes of forecasting the house price index. Also, moving further ahead in the forecasting horizon, the dispersion between models' forecasts increases. One-quarter-ahead forecasts are close, but a range of forecasts at eight quarters ahead is wide. Furthermore, some models, for example, ARMA(4,1) or VAR(1) Spec. 2, showed significant adjustments in their forecasts as new data are taken into account.

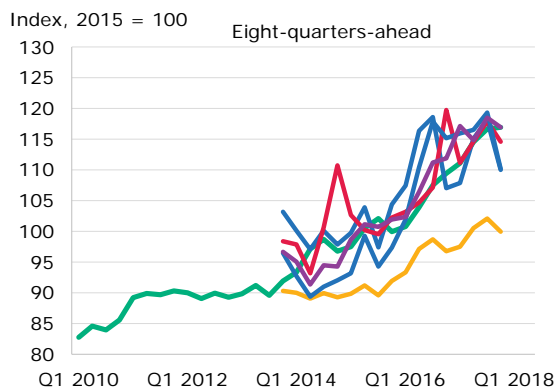
Fig. 9. Forecasts of individual models at different forecasting horizons



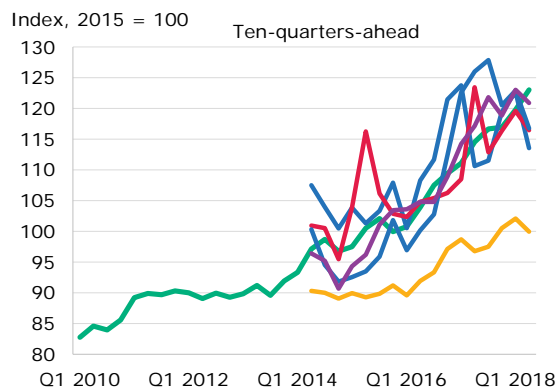
Sources: Statistics Lithuania, authors' calculations



Sources: Statistics Lithuania, authors' calculations



Sources: Statistics Lithuania, authors' calculations



Sources: Statistics Lithuania, authors' calculations

To summarise, BVAR-SV(1) Spec. 7, estimated taking data in log levels, provided the best forecasting accuracy across all estimated models. However, we recognise that the results are driven by the current data sample, which is relatively short and volatile. Thus, the results may change as we will have more data in our analysis.

#### 7.4. FORECAST COMBINATION

Even though BVAR-SV(1) Spec. 7 showed the best out-of-sample forecasting accuracy in our analysis, the data generating process is unknown and model uncertainty risk is still an issue. Furthermore, if we consider many forecasting models, it is unclear which forecast to focus on. Those issues can be addressed by combining the forecasts. Various forecast combination methods have gained ground in the forecasting literature because it is found that in empirical exercises these methods usually outperform forecasts from a single model. A detailed discussion on forecast combinations can be found in Timmermann (2006). The author distinguishes several reasons why forecast combination may be useful. First, forecast combination may be more robust to unknown instabilities (e.g. structural breaks) than an individual model. Second, individual models may be subject to misspecification bias and these biases may be averaged out in the combination forecast.

In our exercise we consider four combination approaches to produce point forecasts at each horizon. The first two approaches are a simple mean and median of the individual forecasts at each forecasting horizon. Timmermann (2006) argues that simple combination methods are hard to beat. The third combination approach is based on individual eight quarters ahead forecasting accuracy (in terms of RMSE) of the model:



$$w_i = \frac{(1 - RMSE_i)}{\sum_{i=1}^M (1 - RMSE_i)} \quad (26)$$

where  $w_i$  is the weights of the individual model,  $RMSE_i$  is the forecast accuracy at eight quarters ahead forecasting horizon, and  $M$  is the number of models. Thus, the model with the lowest forecast error receives the highest weight. The fourth approach is similar to the third, but we take only those models in which RMAE is smaller than one at eight quarters ahead forecasting horizon. Thus, we combine only those models, which were better than the benchmark forecast.

Table 2. Forecasting accuracy of the forecast combination approaches

Combination approach	H1	H4	H8	H10
Mean	0.020	<b>0.035</b>	<b>0.045</b>	<b>0.043</b>
Median	0.019	<b>0.033</b>	<b>0.047</b>	<b>0.046</b>
Weighted mean (RMSE)	0.020	<b>0.034</b>	<b>0.044</b>	<b>0.042</b>
Weighted mean (RMSE, where RMAE < 1)	<b>0.019</b>	<b>0.033</b>	<b>0.042</b>	<b>0.038</b>

Notes: the table provides RMSE at a different forecasting horizon. The bold numbers show models where RMAE < 1.

Table 2 shows the forecasting accuracy of the different forecast combination approaches. The results indicate that even simple combination methods increased forecasting accuracy considerably. Applying weights based on RMSEs further enhances forecasting power. Only three BVAR-SV type models showed a bit better forecasting performance than the first three forecast combination approaches. And only one model, i.e. BVAR-SV(1) Spec. 7, showed better forecasting accuracy than the fourth combination approach. Nevertheless, the forecast combination helps to address model uncertainty issue and is still reasonably accurate over various forecasting horizons.

## CONCLUDING REMARKS

In this paper, we addressed three important questions related to Lithuanian house prices, namely, whether house prices are under or over valued, which explanatory variables have the biggest impact on the growth of house prices and what the future development of the Lithuanian house price index could be. Three separate modelling and forecasting exercises were performed in order to try answering these questions.

The results of the VECM model show that house prices are negatively associated with real lending rates and have a strong positive association with construction costs and credit flows (in excess of housing investment). House prices also exhibit very strong inertia and the pull towards the long-term equilibrium is weak and statistically insignificant. The model helps identify a pronounced boom-bust cycle of house prices and shows that house prices were still slightly below their econometric balanced path at the end of the analysed sample, in 2018.

We employed an ARDL model to assess which explanatory variables had the biggest impact on house price growth. The results revealed that interest rates on new loans to households and building permits are among the key drivers behind house price dynamics in Lithuania. The unexplained error may show that the actual decrease in growth could be attributed to changes in perception of the house value or that the relationship between variables is more contemporaneous. In both cases, the current modelling setup does not allow us to address these issues. The short term out-of-sample forecasting showed that we need a longer time series to make more accurate predictions with this type of model.

Furthermore, a suite of models was built to make short and medium term forecasts of the Lithuanian house price index. We employed various univariate and multivariate time series forecasting models and compared their forecasting accuracy. The results show that the multivariate models outperformed univariate models, but we need Bayesian estimation to exploit the gain of the additional variables. Relatively short time series may be one of the reasons why the traditional VAR models provided relatively poor forecasting accuracy. To

safeguard against model uncertainty we considered several forecast combination approaches that showed improved forecasting accuracy compared to individual models. Thus, the combination of forecasts will be used to predict future changes in the Lithuanian housing price index.

Analysis presented in this paper may be viewed as a further step towards more formalised modelling and forecasting of the Lithuanian house price index at the Bank of Lithuania. Each of the empirical exercises could be improved in various directions. However, we are analysing relatively short and volatile data and, therefore, as more observations become available, a longer data sample will help to model and forecast the Lithuanian house price index more accurately.

## REFERENCES

- Anundsen, A.K. (2016). *Detecting imbalances in house prices: What goes up must go down?* Norges Bank Working Paper 11.
- Anundsen, A.K., & E.S. Jansen (2013). *Self-reinforcing effects between housing prices and credit*. Journal of Housing Economics, Vol. 22(3), pp. 192-212.
- Barari, M., N. Sarkar, S. Kundu, & K.B. Chowdhury (2014). *Forecasting house prices in the United States with multiple structural breaks*. International Economic Review, Vol. 6(1), pp. 1-23.
- Barcilar, M., R. Gupta, & S.M. Miller (2015). *The out-of-sample forecasting performance of non-linear models of regional housing prices in the US*. Applied Economics, Vol. 47(22), pp. 2259-2277.
- Bork, L., and S.V. Møller (2015). *Forecasting house prices in the 50 states using dynamic model averaging and dynamic model selection*. International Journal of Forecasting, Vol. 31(1), pp. 63-78.
- Carstensen, K., O. Hülsewig, & T. Wollmershäuser (2009). *Monetary policy transmission and house prices: European cross-country evidence*. CESifo, Working Paper, No 2750.
- Chan, J.C.C., & E. Eisenstat (2018). *Bayesian model comparison for time-varying parameter VARs with stochastic volatility*. Journal of Applied Econometrics, Vol. 33(4), pp. 509-532.
- Cogley, T., & T.J. Sargent (2005). *Drifts and volatilities: monetary policies and outcomes in the post WWII US*. Review of Economic Dynamics, Vol. 8(2), pp. 262-302.
- Das, S., R. Gupta, & A. Kabundi (2011). *Forecasting regional house price inflation: a comparison between dynamic factor models and vector autoregressive models*. Journal of Forecasting, Vol. 30(2), pp. 288-302.
- Drought, S., & Ch. McDonald (2011). *Forecasting house price inflation: a model combination approach*. Reserve Bank of New Zealand, Discussion Paper series, 2011/07.
- Eickmeier, S., & B. Hofmann (2010). *Monetary policy, housing booms and financial (im)balances*. Deutsche Bundesbank, Discussion Paper Series 1: Economic Studies, No 07/2010.
- Emiris, M. (2016). *A dynamic factor model for forecasting house price in Belgium*. National Bank of Belgium Working Paper research, No 313.
- Jarociński, M., & F. Smets (2008). *House prices and the stance of monetary policy*. European Central Bank, Working Paper No 891.
- de Haas, R. & I. de Greef (2000). *Housing prices, bank lending, and monetary policy*. De Nederlandsche Bank Research Series Supervision Paper, No 31.
- Gerdemesmer, D., A. Lenarčič, & B. Roffia (2012). *An alternative method for identifying booms and busts in the euro area housing market*. European Central Bank, Working Paper Series, No 1493.
- Greiber, C., & R. Setzer (2007). *Money and housing – evidence for the euro area and the US*. Deutsche Bundesbank, Discussion Paper Series: Economic Studies, No 12/2007.

- Gupta, R., A. Kabundi, & S.M. Miller (2009). *Forecasting the US real house price index: structural and non-structural models with and without fundamentals*. University of Connecticut, Department of Economics Working Paper series, 2009-42.
- Gupta, R., M. Jurgilas, & A. Kabundi (2010). *The effect of monetary policy on real house price growth in South Africa: a factor-augmented vector autoregression (FAVAR) approach*. *Econometric Modelling*, Vol. 27(1), pp. 315-323.
- Gupta, R., & S.M. Miller (2009). *The time-series properties on housing prices: A case study of the Southern California Market*. University of Connecticut, Department of Economics, Working Paper series, 2009-10.
- Hyndman, R.J., & A.B. Koehler (2006). *Another look at measures of forecast accuracy*. *International Journal of Forecasting*, Vol. 22(4), pp. 679-688.
- Kajuth, F., Th. Knetsch, & N. Pinkwart (2013). *Assessing house prices in Germany: evidence from an estimated stock-flow model using regional data*. Deutsche Bundesbank Discussion Paper No. 46/2013.
- Karlsson, S. (2013). *Forecasting with Bayesian vector autoregressions*. *Handbook of Economic Forecasting*, Vol. 2 (Part B), pp. 791-897.
- Katrakilidis, C., & E. Trachanas (2012). *What drive housing price dynamics in Greece: new evidence from asymmetric ARDL cointegration*. *Economic Modelling*, Vol. 29(4), pp. 1064-1069.
- Kouwenberg, R., & R. Zwinkels (2014). *Forecasting the US housing market*. *International Journal of Forecasting*. Vol. 30, pp. 415-425.
- Litterman, R. (1986). *Forecasting with Bayesian vector autoregressions – five years of experience*. *Journal of Business and Economic Statistics*, Vol. 4(1), pp. 25-38.
- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.
- Meen, G. (2002). *The time-series behavior of house prices: a transatlantic divide?* *Journal of Housing Economics*, Vol. 11, pp. 1-23.
- an de Meulen, Ph., M. Micheli, & T. Schmidt (2011). *Forecasting house prices in Germany*. Ruhr Economic Papers, No. 294.
- Muellbauer, J., & A. Murphy (1997). *Booms and busts in the UK housing market*. *The Economic Journal*, Vol. 107(445), pp. 1701-1727.
- Oikarinen, E. (2008). *Interaction between housing prices and household borrowing in Finland*. ETLA Discussion Papers, No. 1145.
- Poterba, J.M. (1984). *Tax subsidies to owner-occupied housing: an asset-market approach*. *The Quarterly Journal of Economics*, Vol. 99(4), pp. 729-752.
- Primiceri, G.E. (2005). *Time varying structural vector autoregressions and monetary policy*. *The Review of Economic Studies*, Vol. 72(3), pp. 821-852.
- Rapach, D.E., & J.K. Strauss (2007). *Forecasting real housing price growth in eight district states*. Federal Reserve Bank of St. Louis, Regional Economic Development, Vol. 3(2), pp. 33-42.
- Risse, M., & M. Kern (2016). *Forecasting house price growth in the Euro area with dynamic model averaging*. *North American Journal of Economics and Finance*, Vol. 38, pp. 70-85.
- Robstad, Ø (2014). *House prices, credit and the effect of monetary policy in Norway: evidence from structural VAR models*. Norges Bank, Working Paper 05.
- Sims, C.A. (1980). *Macroeconomics and reality*. *Econometrica* Vol. 48, pp. 1-48.
- Stiglitz, J.E. (1990). *Symposium on Bubbles*. *Journal of Economic Perspectives*, Vol. 4(2), pp. 13-18.

Timmermann, A (2006). *Forecast Combinations*. Handbook of Economic Forecasting, Vol. 1 (chapter 4), pp. 135-196, Elsevier.

## APPENDIX A

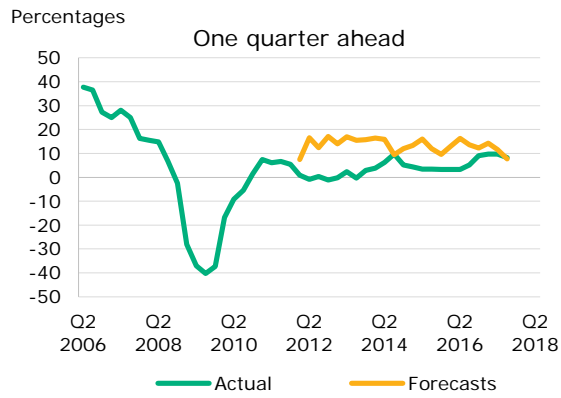
Table A1. Cointegrating relationship estimation results

Dependent variable:	<i>hpi</i>	
	Coefficients	Standard errors
<i>Variable</i>		
<i>Intercept</i>	-2.054*	0.399
<i>r</i>	-1.084*	0.376
<i>ccpi</i>	1.543*	0.083
<i>credInv</i>	0.205*	0.020
<i>debtRatio</i>	0.231*	0.029

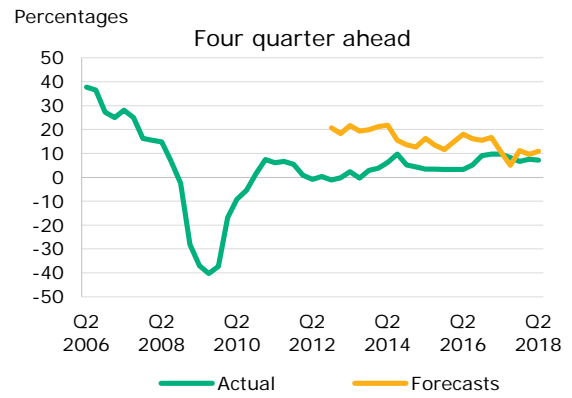
Notes: \* significant at the 0.05 level.

## APPENDIX B

Fig. B1. Out-of-sample forecasts of weighted ARDL model at different forecasting horizons



Sources: Statistics Lithuania, authors' calculations



Sources: Statistics Lithuania, authors' calculations

## APPENDIX C

Table C1. Forecasting accuracy of the naive forecasts and univariate ARMA models

Model	log level				q-o-q growth				y-o-y growth			
	h1	h4	h8	h10	h1	h4	h8	h10	h1	h4	h8	h10
Naive	0.018	0.047	0.101	0.132	0.025	0.080	0.156	0.191	0.024	<b>0.042</b>	<b>0.077</b>	<b>0.090</b>
AR(1)	0.020	0.054	0.130	0.173	<b>0.020</b>	<b>0.042</b>	<b>0.065</b>	<b>0.071</b>	0.041	0.053	0.113	0.153
AR(2)	0.022	0.070	0.174	0.232	<b>0.020</b>	<b>0.041</b>	<b>0.062</b>	<b>0.067</b>	0.040	<b>0.046</b>	<b>0.092</b>	<b>0.125</b>
AR(3)	0.022	0.075	0.195	0.261	<b>0.020</b>	<b>0.039</b>	<b>0.061</b>	<b>0.065</b>	0.038	<b>0.042</b>	<b>0.082</b>	<b>0.113</b>
AR(4)	0.022	0.075	0.194	0.261	<b>0.020</b>	<b>0.039</b>	<b>0.061</b>	<b>0.064</b>	0.032	<b>0.043</b>	<b>0.085</b>	<b>0.117</b>
ARMA(1,1)	0.021	0.060	0.146	0.195	<b>0.020</b>	<b>0.039</b>	<b>0.059</b>	<b>0.061</b>	0.040	<b>0.049</b>	<b>0.100</b>	<b>0.134</b>
ARMA(2,1)	0.022	0.074	0.192	0.259	<b>0.020</b>	<b>0.040</b>	<b>0.061</b>	<b>0.065</b>	0.037	<b>0.044</b>	<b>0.092</b>	<b>0.129</b>
ARMA(3,1)	0.023	0.074	0.189	0.254	<b>0.020</b>	<b>0.040</b>	<b>0.062</b>	<b>0.066</b>	0.036	<b>0.042</b>	<b>0.085</b>	<b>0.119</b>
ARMA(4,1)	0.022	0.075	0.191	0.255	<b>0.020</b>	<b>0.039</b>	<b>0.057</b>	<b>0.060</b>	0.031	<b>0.041</b>	<b>0.078</b>	<b>0.109</b>
ARMA(1,2)	0.022	0.064	0.157	0.209	<b>0.020</b>	<b>0.040</b>	<b>0.062</b>	<b>0.067</b>	0.033	<b>0.048</b>	<b>0.099</b>	<b>0.129</b>
ARMA(2,2)	0.023	0.071	0.177	0.236	<b>0.020</b>	<b>0.040</b>	<b>0.063</b>	<b>0.067</b>	0.029	<b>0.044</b>	<b>0.081</b>	<b>0.096</b>
ARMA(3,2)*	-	-	-	-	<b>0.020</b>	<b>0.040</b>	<b>0.060</b>	<b>0.063</b>	0.031	<b>0.046</b>	<b>0.089</b>	<b>0.104</b>
ARMA(4,2)	0.022	0.077	0.202	0.269	0.028	<b>0.043</b>	<b>0.064</b>	<b>0.065</b>	0.049	<b>0.047</b>	<b>0.089</b>	<b>0.140</b>

Notes: the table provides RMSE at different forecasting horizon. The bolded numbers show models where RMAE < 1.

\* Stationary model was not found for ARMA(3,2) specification.

Table C2. Forecasting accuracy of the VAR models

Model	log level				q-o-q growth				y-o-y growth			
	h1	h4	h8	h10	h1	h4	h8	h10	h1	h4	h8	h10
<b>Specification 1</b>												
VAR(1)	0.038	0.104	0.119	<b>0.105</b>	0.021	0.051	<b>0.080</b>	<b>0.093</b>	0.026	0.057	0.117	0.155
VAR(2)	0.031	0.115	0.168	0.163	0.021	0.062	0.117	0.144	0.033	0.077	0.190	0.264
VAR(3)	0.021	0.071	0.156	0.176	0.022	0.065	0.147	0.190	0.037	0.094	0.211	0.278
VAR(4)	0.022	0.049	<b>0.085</b>	<b>0.100</b>	0.027	0.062	0.138	0.185	0.039	0.106	0.232	0.300
<b>Specification 2</b>												
VAR(1)	0.020	<b>0.041</b>	<b>0.049</b>	<b>0.049</b>	0.029	0.065	<b>0.093</b>	<b>0.104</b>	0.027	0.055	<b>0.099</b>	<b>0.122</b>
VAR(2)	0.028	0.058	<b>0.071</b>	<b>0.072</b>	0.025	0.055	<b>0.090</b>	<b>0.103</b>	0.046	0.093	0.164	0.223
VAR(3)	0.028	<b>0.046</b>	<b>0.059</b>	<b>0.060</b>	0.042	0.079	0.148	0.189	0.044	0.098	0.133	<b>0.152</b>
VAR(4)	0.039	0.062	<b>0.085</b>	<b>0.077</b>	0.043	0.080	0.143	0.168	0.062	0.131	0.190	0.241
<b>Specification 3</b>												
VAR(1)	0.021	<b>0.042</b>	<b>0.050</b>	<b>0.057</b>	0.024	0.068	<b>0.117</b>	<b>0.131</b>	0.050	0.103	0.188	0.246
VAR(2)	0.031	0.057	<b>0.099</b>	<b>0.121</b>	0.031	0.074	<b>0.129</b>	<b>0.142</b>	0.057	0.095	0.160	0.205
VAR(3)	0.035	0.059	<b>0.083</b>	<b>0.079</b>	0.037	0.098	0.172	0.191	0.062	0.136	0.197	0.223
VAR(4)	0.044	0.104	0.163	<b>0.173</b>	0.043	0.119	0.286	0.391	0.068	0.136	0.290	0.298
<b>Specification 4</b>												
VAR(1)	0.038	0.119	0.179	0.184	0.021	0.052	<b>0.084</b>	<b>0.094</b>	0.031	0.054	<b>0.080</b>	<b>0.094</b>
VAR(2)	0.031	0.102	0.177	0.192	0.020	<b>0.047</b>	<b>0.078</b>	<b>0.084</b>	0.039	0.060	<b>0.109</b>	<b>0.132</b>
VAR(3)	0.025	0.066	0.162	0.191	0.026	0.055	<b>0.090</b>	<b>0.090</b>	0.048	0.088	0.156	0.199
VAR(4)	0.043	0.089	0.184	0.190	0.028	0.059	<b>0.113</b>	<b>0.132</b>	0.061	0.119	0.208	0.250
<b>Specification 5</b>												
VAR(1)	0.036	0.117	0.182	0.189	0.024	0.071	<b>0.128</b>	<b>0.143</b>	0.037	0.072	0.118	<b>0.145</b>
VAR(2)	0.030	0.094	0.142	<b>0.133</b>	0.024	0.071	<b>0.120</b>	<b>0.129</b>	0.047	0.082	0.147	0.181
VAR(3)	0.029	0.063	<b>0.114</b>	<b>0.139</b>	0.029	0.074	0.138	<b>0.170</b>	0.060	0.099	0.170	0.206
VAR(4)	0.052	0.125	0.234	0.261	0.031	0.074	0.161	0.219	0.060	0.237	0.506	0.577
<b>Specification 6</b>												
VAR(1)	0.021	<b>0.043</b>	<b>0.049</b>	<b>0.047</b>	0.024	0.052	<b>0.079</b>	<b>0.090</b>	0.030	0.055	<b>0.087</b>	<b>0.100</b>
VAR(2)	0.027	0.059	<b>0.069</b>	<b>0.066</b>	0.023	<b>0.045</b>	<b>0.066</b>	<b>0.066</b>	0.035	0.078	0.154	0.216
VAR(3)	0.029	<b>0.050</b>	<b>0.068</b>	<b>0.065</b>	0.032	0.060	0.119	<b>0.153</b>	0.056	0.125	0.240	0.307
VAR(4)	0.038	0.060	<b>0.085</b>	<b>0.074</b>	0.033	0.072	0.127	<b>0.136</b>	0.060	0.147	0.301	0.390
<b>Specification 7</b>												
VAR(1)	0.021	<b>0.042</b>	<b>0.061</b>	<b>0.074</b>	0.023	<b>0.060</b>	<b>0.104</b>	<b>0.119</b>	0.039	0.078	0.130	<b>0.159</b>
VAR(2)	0.026	0.067	0.192	0.245	0.029	0.070	<b>0.116</b>	<b>0.126</b>	0.052	0.089	0.169	0.240
VAR(3)	0.032	0.070	0.165	<b>0.188</b>	0.036	0.088	0.155	0.175	0.072	0.119	0.192	0.233
VAR(4)	0.045	0.181	0.406	0.481	0.046	0.151	0.386	0.525	0.088	0.134	0.252	0.236

Notes: the table provides RMSE at different forecasting horizon. The bolded numbers show models where RMAE < 1.



Table C3. Forecasting accuracy of the BVAR models

Model	log level				q-o-q growth				y-o-y growth			
	h1	h4	h8	h10	h1	h4	h8	h10	h1	h4	h8	h10
<b>Specification 1</b>												
BVAR(1)	0.024	0.052	<b>0.051</b>	<b>0.045</b>	0.021	<b>0.046</b>	<b>0.074</b>	<b>0.090</b>	0.023	<b>0.045</b>	<b>0.091</b>	<b>0.114</b>
BVAR(2)	0.025	0.057	<b>0.058</b>	<b>0.050</b>	0.021	<b>0.043</b>	<b>0.056</b>	<b>0.059</b>	0.025	0.053	0.118	0.153
BVAR(3)	0.025	0.062	<b>0.077</b>	<b>0.070</b>	0.021	<b>0.047</b>	<b>0.081</b>	<b>0.101</b>	0.024	0.055	0.120	0.161
BVAR(4)	0.021	<b>0.041</b>	<b>0.054</b>	<b>0.055</b>	0.021	<b>0.048</b>	<b>0.087</b>	<b>0.111</b>	0.024	<b>0.050</b>	<b>0.106</b>	0.142
<b>Specification 2</b>												
BVAR(1)	<b>0.018</b>	<b>0.041</b>	<b>0.085</b>	<b>0.111</b>	0.021	<b>0.044</b>	<b>0.067</b>	<b>0.079</b>	0.022	<b>0.037</b>	<b>0.057</b>	<b>0.063</b>
BVAR(2)	<b>0.018</b>	<b>0.040</b>	<b>0.082</b>	<b>0.107</b>	0.021	<b>0.043</b>	<b>0.057</b>	<b>0.060</b>	0.022	<b>0.039</b>	<b>0.069</b>	<b>0.081</b>
BVAR(3)	<b>0.018</b>	<b>0.039</b>	<b>0.079</b>	<b>0.103</b>	0.021	<b>0.051</b>	<b>0.101</b>	<b>0.129</b>	0.022	<b>0.039</b>	<b>0.065</b>	<b>0.077</b>
BVAR(4)	<b>0.019</b>	<b>0.038</b>	<b>0.074</b>	<b>0.096</b>	0.021	<b>0.046</b>	<b>0.085</b>	<b>0.109</b>	0.023	<b>0.040</b>	<b>0.061</b>	<b>0.072</b>
<b>Specification 3</b>												
BVAR(1)	<b>0.017</b>	<b>0.039</b>	<b>0.081</b>	<b>0.108</b>	0.021	0.052	<b>0.087</b>	<b>0.103</b>	0.023	<b>0.038</b>	<b>0.053</b>	<b>0.063</b>
BVAR(2)	<b>0.018</b>	<b>0.042</b>	<b>0.092</b>	<b>0.125</b>	0.021	<b>0.044</b>	<b>0.062</b>	<b>0.066</b>	0.023	<b>0.040</b>	<b>0.064</b>	<b>0.075</b>
BVAR(3)	<b>0.018</b>	<b>0.039</b>	<b>0.090</b>	<b>0.118</b>	0.021	<b>0.050</b>	<b>0.099</b>	<b>0.126</b>	0.027	0.051	<b>0.084</b>	<b>0.098</b>
BVAR(4)	0.019	<b>0.046</b>	0.104	0.136	0.021	<b>0.047</b>	<b>0.078</b>	<b>0.096</b>	0.028	0.071	0.139	0.164
<b>Specification 4</b>												
BVAR(1)	<b>0.019</b>	<b>0.034</b>	<b>0.063</b>	<b>0.088</b>	0.021	<b>0.046</b>	<b>0.072</b>	<b>0.081</b>	0.022	<b>0.038</b>	<b>0.057</b>	<b>0.064</b>
BVAR(2)	<b>0.019</b>	<b>0.036</b>	<b>0.060</b>	<b>0.081</b>	0.021	<b>0.043</b>	<b>0.057</b>	<b>0.059</b>	0.023	<b>0.043</b>	<b>0.077</b>	<b>0.093</b>
BVAR(3)	<b>0.020</b>	<b>0.042</b>	<b>0.082</b>	<b>0.109</b>	0.021	<b>0.052</b>	0.108	0.142	0.023	<b>0.044</b>	<b>0.075</b>	<b>0.091</b>
BVAR(4)	<b>0.019</b>	<b>0.037</b>	<b>0.076</b>	<b>0.103</b>	0.021	<b>0.050</b>	<b>0.101</b>	<b>0.136</b>	0.023	<b>0.040</b>	<b>0.065</b>	<b>0.078</b>
<b>Specification 5</b>												
BVAR(1)	<b>0.018</b>	<b>0.028</b>	<b>0.056</b>	<b>0.080</b>	0.021	0.051	<b>0.087</b>	<b>0.102</b>	0.024	<b>0.040</b>	<b>0.064</b>	<b>0.081</b>
BVAR(2)	<b>0.018</b>	<b>0.032</b>	<b>0.061</b>	<b>0.086</b>	0.021	<b>0.044</b>	<b>0.062</b>	<b>0.065</b>	0.024	<b>0.044</b>	<b>0.076</b>	<b>0.091</b>
BVAR(3)	<b>0.019</b>	<b>0.035</b>	<b>0.075</b>	<b>0.102</b>	0.021	<b>0.048</b>	<b>0.097</b>	<b>0.125</b>	0.027	0.052	<b>0.088</b>	<b>0.103</b>
BVAR(4)	0.020	<b>0.046</b>	0.111	0.150	0.021	<b>0.048</b>	<b>0.081</b>	<b>0.100</b>	0.027	0.058	<b>0.106</b>	<b>0.126</b>
<b>Specification 6</b>												
BVAR(1)	<b>0.019</b>	0.048	<b>0.098</b>	<b>0.127</b>	0.021	<b>0.045</b>	<b>0.070</b>	<b>0.083</b>	0.023	<b>0.036</b>	<b>0.054</b>	<b>0.057</b>
BVAR(2)	0.019	0.047	<b>0.096</b>	<b>0.124</b>	0.021	<b>0.042</b>	<b>0.056</b>	<b>0.058</b>	0.022	<b>0.039</b>	<b>0.066</b>	<b>0.077</b>
BVAR(3)	0.019	<b>0.045</b>	<b>0.093</b>	<b>0.118</b>	0.021	<b>0.047</b>	<b>0.088</b>	<b>0.111</b>	0.022	<b>0.038</b>	<b>0.065</b>	<b>0.075</b>
BVAR(4)	<b>0.019</b>	<b>0.041</b>	<b>0.080</b>	<b>0.102</b>	0.021	<b>0.044</b>	<b>0.076</b>	<b>0.096</b>	0.023	<b>0.038</b>	<b>0.057</b>	<b>0.067</b>
<b>Specification 7</b>												
BVAR(1)	<b>0.019</b>	0.047	<b>0.098</b>	<b>0.127</b>	0.021	0.051	<b>0.086</b>	<b>0.101</b>	0.023	<b>0.038</b>	<b>0.054</b>	<b>0.061</b>
BVAR(2)	0.019	0.048	0.105	0.139	0.021	<b>0.044</b>	<b>0.060</b>	<b>0.064</b>	0.023	<b>0.041</b>	<b>0.066</b>	<b>0.076</b>
BVAR(3)	<b>0.018</b>	<b>0.046</b>	0.109	0.142	0.021	<b>0.049</b>	<b>0.094</b>	<b>0.117</b>	0.026	<b>0.047</b>	<b>0.074</b>	<b>0.083</b>
BVAR(4)	<b>0.019</b>	<b>0.047</b>	0.111	0.145	0.021	<b>0.047</b>	<b>0.079</b>	<b>0.098</b>	0.026	<b>0.047</b>	<b>0.078</b>	<b>0.089</b>

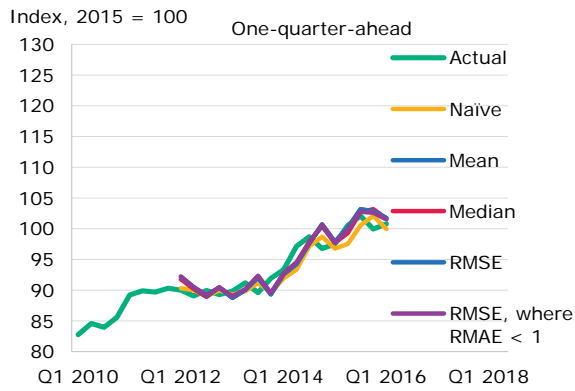
Notes: the table provides RMSE at different forecasting horizon. The bolded numbers show models where RMAE < 1.

Table C4. Forecasting accuracy of the BVAR-SV models

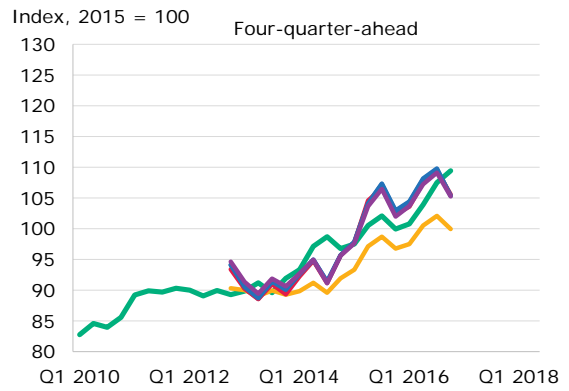
Model	log level				q-o-q growth				y-o-y growth			
	h1	h4	h8	h10	h1	h4	h8	h10	h1	h4	h8	h10
<b>Specification 1</b>												
BVAR-SV(1)	0.030	0.083	<b>0.120</b>	<b>0.129</b>	0.020	0.051	<b>0.101</b>	<b>0.131</b>	0.022	<b>0.036</b>	<b>0.062</b>	<b>0.069</b>
BVAR-SV(2)	0.027	0.079	<b>0.116</b>	<b>0.125</b>	0.021	<b>0.044</b>	<b>0.086</b>	<b>0.112</b>	0.021	<b>0.040</b>	<b>0.084</b>	<b>0.107</b>
BVAR-SV(3)	0.025	0.068	<b>0.107</b>	<b>0.117</b>	0.020	<b>0.033</b>	<b>0.054</b>	<b>0.061</b>	0.023	<b>0.037</b>	<b>0.073</b>	<b>0.093</b>
BVAR-SV(4)	0.021	<b>0.046</b>	<b>0.074</b>	<b>0.086</b>	0.021	<b>0.034</b>	<b>0.051</b>	<b>0.056</b>	0.023	<b>0.036</b>	<b>0.067</b>	<b>0.085</b>
<b>Specification 2</b>												
BVAR-SV(1)	<b>0.017</b>	<b>0.033</b>	<b>0.065</b>	<b>0.080</b>	0.019	<b>0.039</b>	<b>0.081</b>	<b>0.108</b>	0.023	<b>0.042</b>	<b>0.071</b>	<b>0.084</b>
BVAR-SV(2)	<b>0.018</b>	<b>0.035</b>	<b>0.067</b>	<b>0.084</b>	0.019	<b>0.036</b>	<b>0.064</b>	<b>0.081</b>	0.025	0.053	<b>0.104</b>	<b>0.133</b>
BVAR-SV(3)	<b>0.019</b>	<b>0.036</b>	<b>0.073</b>	<b>0.093</b>	0.019	<b>0.035</b>	<b>0.054</b>	<b>0.063</b>	0.026	0.056	0.114	0.151
BVAR-SV(4)	<b>0.019</b>	<b>0.034</b>	<b>0.066</b>	<b>0.084</b>	0.020	<b>0.037</b>	<b>0.050</b>	<b>0.054</b>	0.026	0.054	<b>0.098</b>	<b>0.123</b>
<b>Specification 3</b>												
BVAR-SV(1)	<b>0.019</b>	<b>0.034</b>	<b>0.041</b>	<b>0.041</b>	0.020	0.055	<b>0.112</b>	<b>0.141</b>	0.027	<b>0.046</b>	<b>0.072</b>	<b>0.088</b>
BVAR-SV(2)	0.021	<b>0.044</b>	<b>0.054</b>	<b>0.051</b>	0.020	0.054	<b>0.105</b>	<b>0.132</b>	0.028	<b>0.048</b>	<b>0.082</b>	<b>0.107</b>
BVAR-SV(3)	0.021	<b>0.044</b>	<b>0.055</b>	<b>0.051</b>	0.020	<b>0.039</b>	<b>0.061</b>	<b>0.066</b>	0.030	0.055	<b>0.097</b>	<b>0.123</b>
BVAR-SV(4)	0.021	<b>0.044</b>	<b>0.056</b>	<b>0.054</b>	0.021	<b>0.042</b>	<b>0.063</b>	<b>0.069</b>	0.036	0.069	<b>0.117</b>	<b>0.152</b>
<b>Specification 4</b>												
BVAR-SV(1)	<b>0.021</b>	<b>0.044</b>	<b>0.060</b>	<b>0.066</b>	0.021	<b>0.042</b>	<b>0.069</b>	<b>0.086</b>	0.023	<b>0.041</b>	<b>0.073</b>	<b>0.089</b>
BVAR-SV(2)	<b>0.020</b>	<b>0.047</b>	<b>0.058</b>	<b>0.063</b>	0.020	<b>0.039</b>	<b>0.056</b>	<b>0.057</b>	0.023	<b>0.046</b>	<b>0.094</b>	<b>0.120</b>
BVAR-SV(3)	<b>0.021</b>	<b>0.047</b>	<b>0.059</b>	<b>0.065</b>	0.021	<b>0.049</b>	<b>0.100</b>	<b>0.126</b>	0.023	<b>0.048</b>	<b>0.096</b>	<b>0.121</b>
BVAR-SV(4)	0.021	<b>0.041</b>	<b>0.048</b>	<b>0.054</b>	0.021	<b>0.046</b>	<b>0.087</b>	<b>0.110</b>	0.025	<b>0.048</b>	<b>0.087</b>	<b>0.108</b>
<b>Specification 5</b>												
BVAR-SV(1)	<b>0.020</b>	<b>0.037</b>	<b>0.048</b>	<b>0.056</b>	0.022	0.053	<b>0.091</b>	<b>0.107</b>	0.028	0.051	<b>0.093</b>	<b>0.118</b>
BVAR-SV(2)	<b>0.019</b>	<b>0.039</b>	<b>0.045</b>	<b>0.048</b>	0.021	<b>0.044</b>	<b>0.067</b>	<b>0.069</b>	0.026	<b>0.048</b>	<b>0.091</b>	<b>0.116</b>
BVAR-SV(3)	0.021	<b>0.040</b>	<b>0.043</b>	<b>0.044</b>	0.021	<b>0.043</b>	<b>0.083</b>	<b>0.099</b>	0.028	0.051	<b>0.094</b>	<b>0.114</b>
BVAR-SV(4)	0.021	<b>0.039</b>	<b>0.039</b>	<b>0.041</b>	0.021	<b>0.047</b>	<b>0.079</b>	<b>0.092</b>	0.033	0.055	<b>0.090</b>	<b>0.108</b>
<b>Specification 6</b>												
BVAR-SV(1)	<b>0.018</b>	<b>0.030</b>	<b>0.056</b>	<b>0.068</b>	0.022	0.066	0.148	0.196	0.025	<b>0.039</b>	<b>0.060</b>	<b>0.068</b>
BVAR-SV(2)	<b>0.018</b>	<b>0.031</b>	<b>0.050</b>	<b>0.059</b>	0.021	0.061	0.146	0.196	0.027	0.053	<b>0.092</b>	<b>0.112</b>
BVAR-SV(3)	0.021	<b>0.032</b>	<b>0.056</b>	<b>0.068</b>	0.019	<b>0.031</b>	<b>0.049</b>	<b>0.056</b>	0.027	0.052	<b>0.092</b>	<b>0.112</b>
BVAR-SV(4)	0.021	<b>0.032</b>	<b>0.050</b>	<b>0.060</b>	0.021	<b>0.036</b>	<b>0.049</b>	<b>0.052</b>	0.035	0.068	<b>0.113</b>	<b>0.144</b>
<b>Specification 7</b>												
BVAR-SV(1)	<b>0.018</b>	<b>0.029</b>	<b>0.030</b>	<b>0.030</b>	0.023	0.079	0.172	0.222	0.031	0.055	<b>0.100</b>	<b>0.134</b>
BVAR-SV(2)	0.022	<b>0.045</b>	<b>0.045</b>	<b>0.037</b>	0.023	0.072	0.155	0.201	0.033	0.060	<b>0.106</b>	<b>0.132</b>
BVAR-SV(3)	0.027	0.061	<b>0.063</b>	<b>0.053</b>	0.020	<b>0.036</b>	<b>0.061</b>	<b>0.070</b>	0.036	0.067	<b>0.117</b>	<b>0.146</b>
BVAR-SV(4)	0.028	0.065	<b>0.071</b>	<b>0.065</b>	0.023	<b>0.049</b>	<b>0.074</b>	<b>0.079</b>	0.044	0.075	<b>0.119</b>	<b>0.144</b>

Notes: the table provides RMSE at different forecasting horizon. The bolded numbers show models where RMAE < 1.

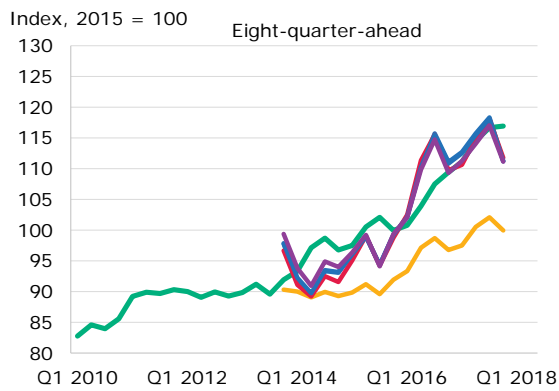
Figure C1. Forecasts of forecast combinations at different forecasting horizons



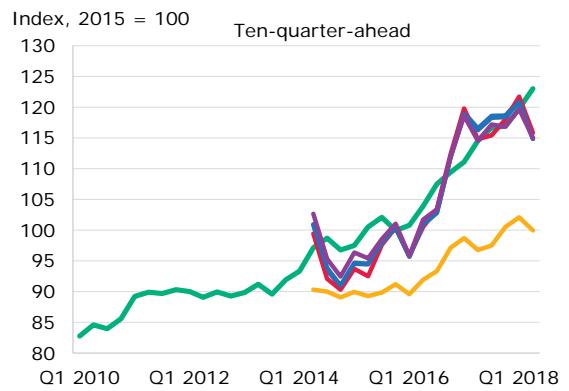
Sources: Statistics Lithuania, authors' calculations



Sources: Statistics Lithuania, authors' calculations



Sources: Statistics Lithuania, authors' calculations



Sources: Statistics Lithuania, authors' calculations