



Delft University of Technology

Forecasting House Prices through Credit Conditions A Bayesian Approach

van der Drift, Rosa; de Haan, Jan; Boelhouwer, Peter

DOI

[10.1007/s10614-023-10542-9](https://doi.org/10.1007/s10614-023-10542-9)

Publication date

2024

Document Version

Final published version

Published in

Computational Economics

Citation (APA)

van der Drift, R., de Haan, J., & Boelhouwer, P. (2024). Forecasting House Prices through Credit Conditions: A Bayesian Approach. *Computational Economics*, 64(6), 3381-3405. <https://doi.org/10.1007/s10614-023-10542-9>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



Forecasting House Prices through Credit Conditions: A Bayesian Approach

Rosa van der Drift¹ · Jan de Haan¹ · Peter Boelhouwer¹

Accepted: 13 December 2023 / Published online: 4 March 2024
© The Author(s) 2024

Abstract

As housing development and housing market policies involve many long-term decisions, improving house price predictions could benefit the functioning of the housing market. Therefore, in this paper, we investigate how house price predictions can be improved. In particular, the merits of Bayesian estimation techniques in enhancing house price predictions are examined in this study. We compare the pseudo out-of-sample forecasting power of three Bayesian models—a Bayesian vector autoregression in levels (BVAR-l), a Bayesian vector autoregression in differences (BVAR-d), and a Bayesian vector error correction model (BVECM)—and their non-Bayesian counterparts. These techniques are compared using a theoretical model that predicts the borrowing capacity of credit-constrained and unconstrained households to affect house prices. The findings indicate that the Bayesian models outperform their non-Bayesian counterparts, and within the class of Bayesian models, the BVAR-d is found to be more accurate than the BVAR-l. For the two winning Bayesian models, i.e., the BVECM and the BVAR-d, the difference in forecasting power is more ambiguous; which model prevails depends on the desired forecasting horizon and the state of the economy. Hence, both Bayesian models may be considered when conducting research on house prices.

Keywords Forecasting · House prices · Bayesian VAR · Bayesian VECM · Cointegration

JEL Classification C11 · E32 · E37 · G12 · R21 · R31

✉ Rosa van der Drift
R.vanderDrift@tudelft.nl

Jan de Haan
J.deHaan-4@tudelft.nl

Peter Boelhouwer
P.J.Boelhouwer@tudelft.nl

¹ Department of MBE, Delft University of Technology, Delft, The Netherlands

1 Introduction

House price predictions are a fundamental part of informed decision-making in the built environment. For instance, the construction of houses is a time-consuming process.¹ Therefore, when contemplating whether to embark on a house-building project, one must consider future prices rather than current ones. The same rationale applies to housing market policies; the implementation of policies is a substantial undertaking, and their effects do not materialize immediately. Consequently, it may be prudent for policymakers to intervene or consider how to intervene before problems emerge. Moreover, house price forecasts can potentially reduce market uncertainty by providing guidance to the market at times when house prices are at a turning point. This is beneficial for banks and their regulators as it enables them to determine better capital requirements, thereby improving risk management. As such, in this paper, we compare different house price forecasts and test which provides more accurate house price predictions.

A frequently used method to forecast house prices is the ordinary least squares (OLS) technique. However, the problem with this estimation technique, for the use of forecasting house prices, is that it minimizes the sum of squares. While this technique results in the best fitting model (the model with the highest possible R^2), it can cause issues for forecasting as OLS models tend to explain variation that cannot be explained: they may overfit the data. That is, the model tries to attribute all variation to the model's variables, while some variation observed in the data simply cannot be explained. This issue is especially relevant in the context of complex models that involve a large number of variables. Nonetheless, even for relatively parsimonious models, overfitting can occur (Giannone et al., 2015). This overfitting problem manifests itself as a poor out-of-sample performance, and hence it is particularly undesirable when forecasting.

Bayesian models can reduce this overfitting problem by utilizing prior information on the behavior of time series.² Therefore, Bayesian estimation techniques are often preferred to their OLS counterparts (Bańbura et al., 2010; Giannone et al., 2015; Koop et al., 2010b). In the literature, three Bayesian model specifications are frequently used to predict economic time series: a Bayesian vector autoregression in differences (BVAR-d), a Bayesian vector autoregression in levels (BVAR-l), and a Bayesian vector error correction model (BVECM). A point of criticism expressed regarding the first model (the BVAR-d), is that it does not take long-run—or cointegrating—relationships into account.³ Since the second model (the BVAR-l) is presented in levels, it does allow for (implicit) cointegrating relationships. The model

¹ The time it costs to build a house depends, of course, on the regulatory context of the country (Oh & Yoon, 2020). Nonetheless, generally speaking, it ranges anywhere from three months to ten years (Meier, 2018; Michielsen et al., 2019; Oh & Yoon, 2020).

² It is worth noting that there are alternative methods that combat overfitting problems. Two popular ones are Ridge and Lasso regression (also called regularization methods). Note, however, that these two are, under some conditions, equivalent to Bayesian estimation. In this paper, we solely rely on Bayesian models due to their flexibility.

³ Note that modeling long-run relationships is especially important when predicting house prices, as the housing market is a sluggish market where prices exhibit long cycles and supply increases slowly.

does, however, include all potential long-run relationships and does not restrict variables to converge to their equilibrium in the long run. Consequently, the last model (the BVECM) might be more desirable than the previous two models as it includes a cointegration constraint that ties the variables together in the long run (Engle & Granger, 1987a).

The contribution of this study is twofold. Firstly, the literature lacks an empirical comparison of the forecasting ability of house prices for the three aforementioned Bayesian models. Although some studies have compared one or two of these Bayesian models to each other or their non-Bayesian counterparts, a systematic evaluation of all three models has not yet been undertaken. We conduct such a systematic comparison of the proposed Bayesian models and their non-Bayesian counterparts. Secondly, in the housing literature, a BVECM has, to the best of our knowledge, not been used to forecast house prices on a macroeconomic (or national) level, even though its model specification looks promising.⁴ Thus, our second contribution to the literature is to assess the merits of a BVECM to predict house prices on a national level.

Theoretical insights are required to incorporate long-run relationships into a BVECM. The long-run relationship we employ takes the importance of the financial sector into account by incorporating an equilibrium relationship between mortgage credit and house prices. This relationship is derived by Van der Drift et al. (2023) and relies on the notion that households spend a fixed fraction of their income on mortgage costs. The analysis is applied to the Dutch housing market as Dutch households generally purchase houses through mortgages, and thus, the proposed credit-based long-run equilibrium applies (Van der Drift et al., 2023).

The remainder of the paper is organized as follows. Section 2 briefly introduces the long-run model. Section 3 summarizes the econometric methods applied in this empirical analysis. Section 4 discusses how we evaluate each model's performance, Sect. 5 presents the results, and the last section concludes the paper.

2 The Long-Run Model

The literature contains an abundance of papers on potential equilibrium relationships for house prices.⁵ In this paper, we focus on one equilibrium relationship: the relationship between mortgage credit and house prices. This equilibrium can be explained by the fact that, in most countries, houses are bought through a mortgage, and consequently, mortgage conditions are an important driver of house prices. Not only does it drive demand, but it can be used to assess whether the housing market is overvalued. This works as follows: if households spend an increasingly larger

⁴ We are only familiar with the application of the BVECM in spatial/regional housing market models, and have not encountered papers that forecast house prices on a national level using a BVECM. Moreover, previous studies utilizing the BVECM employed standard priors and thereby neglected identification issues that can arise with the use of standard priors. Section 3.4 provides a more detailed discussion on this matter.

⁵ E.g., Fraser et al. (2008), Hort (1998), Hott and Monnin (2008), Iacoviello and Neri (2010), Korn and Yilmaz (2022), Liu et al. (2017), Malpezzi (1999), Mikhed and Zemcik (2009), Ozbakan et al. (2019), Scott (1990), and Tu et al. (2017).

share of their income on housing, it could indicate potential default risk. That is, households simply are unable to spend a larger and larger fraction of their income on housing, and consequently, the demand for housing, and hence house prices, are likely to decrease shortly.

The credit-based equilibrium relationship employed in this paper incorporates the recent extension developed by Van der Drift et al. (2023). In contrast to previous equilibrium relationships, this relationship takes two types of households into account: lower-income and higher-income households. The first type, lower-income households, are assumed to be constrained in their borrowing behavior. This constraint could be caused by either a credit rationing bank or stringent financial regulation. Either way, due to the constraint, borrowers can only spend a certain share of their monthly income on mortgage costs (i.e., these households face a debt-service-to-income cap). In the model, these maximum monthly mortgage payments are discounted to receive the total amount a lower-income household can spend on housing ($b_t^{\max,c}$).⁶

The other category of consumers, higher-income households, may not be subject to lending regulations. However, these households might exhibit a preference for allocating a specific proportion of their income towards mortgage payments. This can be conceptualized as a household budget plan where housing, the largest expense, receives a set amount and the remaining funds are allocated towards other expenses such as food, transportation, and leisure activities. Again, the monthly values are discounted to reflect their ability (or desire) to pay for housing ($b_t^{cd,u}$).⁷

In Van der Drift et al. (2023), it is formally derived that this long-run equilibrium can be presented by:

$$P_t = \gamma b_t^{\max,c} + (1 - \gamma)\theta b_t^{cd,u}. \quad (1)$$

Thus, from Eq. (1) it follows that the purchase price of housing (P_t) is explained by the maximum borrowing amount of constrained households ($b_t^{\max,c}$) and the ability to pay of unconstrained households ($b_t^{cd,u}$).⁸ Finally, γ reflects the spending share of constrained households and the parameter θ reflects the fraction of income high-income households spend on mortgage costs.⁹ An overview of the data sources used to estimate Eq. (1) is presented in Table 1 and the data is plotted in Fig. 1.

While this equilibrium condition may provide a useful benchmark, the reality of the housing market is more complex. The housing market is a sluggish market and, due to

⁶ The maximum borrowing amount of constrained households ($b_t^{\max,c}$) is expressed as a proportion (κ_t) of household income (y_t), discounted at the mortgage interest rate (i_t) over the duration of the mortgage (n). Therefore, the maximum borrowing amount is given by: $b_t^{\max,c} = \kappa_t y_t \frac{1 - (1+i_t)^{-n}}{i_t}$.

⁷ The ability to pay of unconstrained households ($b_t^{cd,u}$) is defined similarly as the maximum borrowing amount of constrained households. There are, however, two differences: the debt-service-to-income cap (κ) is not included, and the mortgage interest deduction rate (τ) is included in the calculation. This yields the following formula: $b_t^{cd,u} = \frac{i_t}{1 - (1+i_t)^{-n}} - i_t \tau_t$.

⁸ Note that housing supply is not included in the equilibrium outcome. This is the case as both types of consumers have unitary elastic demand. Consequently, an increase in housing supply will be offset by households demanding more housing. Please refer to Van der Drift et al. (2023) for more information on this topic.

⁹ The spending share is determined through empirical estimation, thus obviating the need for data on the relative consumption shares in the analysis.

Table 1 Data

Variable	Abbr	Source
Median house prices	P_t	NVM
Maximum borrowing amount of constrained households	$b_t^{max,c}$	CBS, DNB & NIBUD
Ability to pay of unconstrained households	$b_t^{cd,u}$	CBS, DNB & Tax

The acronyms NVM, DNB, NIBUD, Tax, and CBS refer respectively to the Netherlands' Association of Realtors, the Central Bank of the Netherlands, the National Institute for Family Financing, Dutch tax documents, and Statistics Netherlands

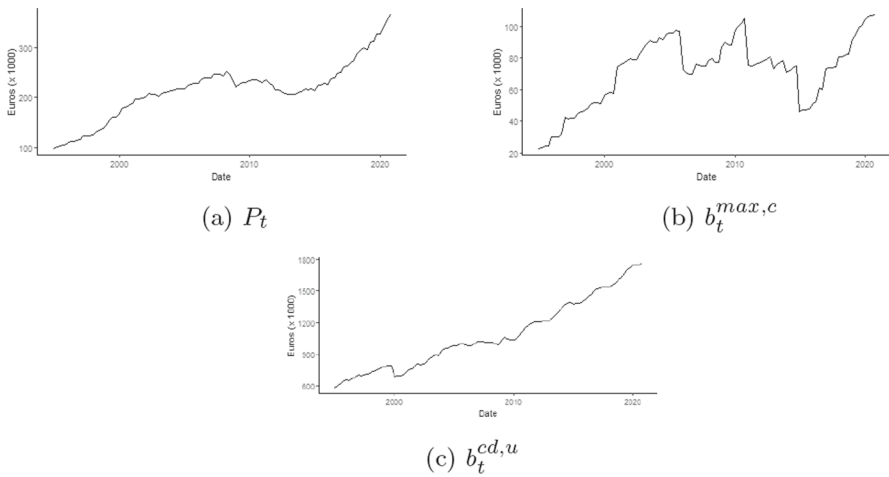


Fig. 1 Time series charts

speculative or psychological effects, it adjusts gradually to changing economic conditions. Consequently, house prices cannot be expected to always be at the long-run equilibrium relationship. As such, studies have found that lagged values of house prices yield substantial explanatory power over future values of house prices (Abraham & Hendershott, 1996; Hort, 1998; Malpezzi, 1999). Especially in the Netherlands, a country that displays minimal house price reversals, research has shown that house prices are past dependent (De Vries & Boelhouwer, 2009; Tu et al., 2017). To account for this phenomenon, we include four lags of each variable in the models.¹⁰

3 Methodology

In the following subsections, we discuss each of the empirical models whose predictive power will be evaluated in this paper.

¹⁰ One might argue that adding this many variables will hamper the model's efficiency. However, we would like to stress that variable selection is a less pronounced issue for Bayesian models as it effectively recognizes robust relationships between variables.

3.1 VAR

VARs are standard models to forecast time series, and they generally produce good forecasting results (Giannone et al., 2019). VARs can be viewed as a system of equations in which each dependent variable is explained by its own lags and the lags of explanatory variables, which themselves are the dependent variables in the system of equations (Sims, 1980). Due to this functional form, VARs are able to depict economic dependencies between own past values and lagged values of explanatory variables. Mathematically, the system of equations, in which all n variables are predicted, can be presented in the following way:

$$y_t = c + \sum_{i=1}^p \beta_i y_{t-i} + \varepsilon_t, \quad (2)$$

where y_t is an $n \times 1$ vector of variables. For the model presented in Sect. 2, y_t is equal to the vector P_t , $b_t^{mac,c}$, and $b_t^{cd,u}$. p reflects the number of lags that are included in the regression and, as discussed in Sect. 2, in our research this term is equal to four. β_i is an $n \times n$ matrix of parameters, and ε_t is an $n \times 1$ vector of disturbance terms or ‘noise’.

A prerequisite for the estimation of the above-presented VAR is that the analyzed time series are stationary.¹¹ House prices—and macroeconomic variables in general—usually fail to meet this requirement (Clayton, 1997; Koop et al., 2005; Wu et al., 2017). These time series are often non-stationary, and OLS-based inference breaks down; the autoregressive coefficients are biased and the t distributions are no longer approximately normally distributed.¹² Moreover, variables can appear to be related when in fact they are not, i.e., we might run a spurious regression.

In order to avoid invalid estimation, the data can be transformed into a stationary process by taking the differences between consecutive observations. If this first (second, third, etcetera) difference is stationary, we can run a VAR in differences (i.e., a VAR-d) without violating estimation assumptions. If we let Δ denote the difference operator, the VAR-d has the following functional form:

$$\Delta y_t = c + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \varepsilon_t. \quad (3)$$

For the model that will be estimated, i.e. the model described in the previous section, Eq. (3) can be written as:

¹¹ A stationary variable is one whose mean, variance, and covariance remain constant over time. Non-stationary variables, on the other hand, may exhibit trends, cycles or random walks.

¹² Note, that if variables are cointegrated that the OLS estimators are unbiased. In fact, the estimators are superconsistent; i.e., they converge to their true values at a faster rate than is the case for stationary variables (Stock, 1987) Nevertheless, if there is not “sufficient cointegration” their asymptotic distribution may be non-standard (Toda & Phillips, 1994).

$$\begin{aligned}
 \Delta P_t &= c_1 + \sum_{i=1}^4 \beta_{1,1,i} \Delta P_{t-i} + \sum_{i=1}^4 \beta_{1,2,i} \Delta b_{t-i}^{\max,c} + \sum_{i=1}^4 \beta_{1,3,i} \Delta b_{t-i}^{cd,u} + \varepsilon_{1,t} \\
 \Delta b_t^{\max,c} &= c_2 + \sum_{i=1}^4 \beta_{2,1,i} \Delta P_{t-i} + \sum_{i=1}^4 \beta_{2,2,i} \Delta b_{t-i}^{\max,c} + \sum_{i=1}^4 \beta_{2,3,i} \Delta b_{t-i}^{cd,u} + \varepsilon_{2,t} \quad (4) \\
 \Delta b_t^{cd,u} &= c_3 + \sum_{i=1}^4 \beta_{3,1,i} \Delta P_{t-i} + \sum_{i=1}^4 \beta_{3,2,i} \Delta b_{t-i}^{\max,c} + \sum_{i=1}^4 \beta_{3,3,i} \Delta b_{t-i}^{cd,u} + \varepsilon_{3,t}
 \end{aligned}$$

Thus, from Eq. (4) it explicitly follows that each variable has its own equation. Thence, all explanatory variables in a VAR are endogenous variables, i.e., they are created within the model. This provides a major advantage in forecasting, as there is no need to obtain information on future values of the explanatory variables.

3.2 VECM

As discussed above, VARs generally have to be differenced in order for the variables to be stationary and the estimation to be valid. However, in case of a common stochastic trend (i.e., cointegration), we can provide a meaningful interpretation of the parameters in their levels without violating any of the estimation assumptions. When data are cointegrated there namely exist relationships between variables in levels, which can render these variables stationary without taking consecutive differences. So-called VECMs capture these relationships and, in the presence of cointegration, are expected to outperform VARs over longer forecasting horizons (Engle & Yoo, 1987).

These long-run relationships are especially important when predicting house prices, as the housing market tends to adapt sluggishly to changing economic conditions. Therefore, in the empirical housing market literature, models that capture these long-run relationships (VECMs) are often preferred over models that do not (VAR-d).¹³ Mathematically, a VECM can be presented by:

$$\Delta y_t = c + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \Pi y_{t-1} + \varepsilon_t. \quad (5)$$

The only difference with a VAR-d (Eq. 3) is the error correction term y_{t-1} , which captures how variables change if one of the variables departs from its equilibrium value. The term's parameter Π represents the effect of the error-correction term on the short-run variables. This term is a $n \times r$ vector, where r is called the cointegration rank. r can be used to demonstrate the relationship between a VECM, a VAR-d, and a VAR-1. This works as follows: if r is zero, there are no cointegrating relationships, and the model reduces to a VAR-d as presented in Eq. (3). If r is larger than zero but smaller than n , cointegration is present and the model is best represented

¹³ See for example: Abraham and Hendershott (1996), Brissimis and Vlassopoulos (2009), Clayton et al. (2019), Damen et al. (2016), Hort (1998), Leung (2014), Malpezzi (1999), Tu et al. (2017), and Tuluca et al. (2000).

by a VECM (Eq. 5).¹⁴ Yet, if $r = n$, the variables are stationary, no cointegration is present, and the model can be validly estimated by a VAR-1 (Eq. 2).

If the data are cointegrated (i.e., $0 < r < n$), the parameter Π can be further divided into two subterms: $\Pi = \alpha\beta'$. Where β' can be interpreted as the distance of the variables from their equilibrium and α and describes the speed at which variables converge back to the equilibrium. In vector form, this error-correction term is rather abstract, yet for the model that will be estimated,¹⁵ the VECM can be written as:

$$\begin{aligned}
 \Delta P_t &= c_1 + \sum_{i=1}^4 \beta_{1,1,i} \Delta P_{t-i} + \sum_{i=1}^4 \beta_{1,2,i} \Delta b_{t-i}^{\max,c} + \sum_{i=1}^4 \beta_{1,3,i} \Delta b_{t-i}^{cd,u} + \alpha_1 (P_{t-1} - \beta'_1 b_{t-1}^{\max,c} - \beta'_2 b_{t-1}^{cd,u}) + \varepsilon_{1,t} \\
 \Delta b_t^{\max,c} &= c_2 + \sum_{i=1}^4 \beta_{2,1,i} \Delta P_{t-i} + \sum_{i=1}^4 \beta_{2,2,i} \Delta b_{t-i}^{\max,c} + \sum_{i=1}^4 \beta_{2,3,i} \Delta b_{t-i}^{cd,u} + \alpha_2 (P_{t-1} - \beta'_1 b_{t-1}^{\max,c} - \beta'_2 b_{t-1}^{cd,u}) + \varepsilon_{2,t} \\
 \Delta b_t^{cd,u} &= c_3 + \sum_{i=1}^4 \beta_{3,1,i} \Delta P_{t-i} + \sum_{i=1}^4 \beta_{3,2,i} \Delta b_{t-i}^{\max,c} + \sum_{i=1}^4 \beta_{3,3,i} \Delta b_{t-i}^{cd,u} + \alpha_3 (P_{t-1} - \beta'_1 b_{t-1}^{\max,c} - \beta'_2 b_{t-1}^{cd,u}) + \varepsilon_{3,t}
 \end{aligned}
 \tag{6}$$

where $P_{t-1} - \beta'_1 b_{t-1}^{\max,c} - \beta'_2 b_{t-1}^{cd,u}$ reflects the error-correction term. This term reflects how much the previous period deviated from the model’s long-run equilibrium. Its coefficient (α), denotes how fast the equilibrium will be restored. Therefore, if variables are above (below) the long-run equilibrium they are expected to decrease (increase) shortly. Thus, in contrast to the VAR, this model incorporates an equilibrium mechanism and, when employed correctly, one can think of this equilibrium mechanism as a bubble buster.

3.3 BVAR

A drawback of the above-described models is that they contain numerous parameters that need to be estimated, while macroeconomic time series generally have a short length. To illustrate this argument: the simple three-variable VAR-d model presented in Eq. (4) already includes 39 parameters that have to be estimated. Particularly, when time series are relatively short, this can cause overfitting; i.e., a situation in which the in-sample fit of the model (i.e., the R^2) is excellent, but the model’s out-of-sample performance is poor.

Bayesian estimation techniques combat the overfitting problem by using Bayes’ rule to shrink the model’s parameters towards a benchmark that is known to have decent forecasting power (Doan et al., 1984; Litterman, 1979; Sims, 1980). This benchmark is called the prior and broadly speaking it reflects a loss function for implausible explanations. Therefore, the model does not optimize its in-sample fit

¹⁴ Where r reflects the number of cointegrating vectors: if this number is one, there exists one linear combination of the variables that is stationary, and if r is two, there exist two linearly independent combinations, et cetera.

¹⁵ In line with the results of the Johansen test (Table 3) we model a VECM with one cointegrating relationship ($r = 1$). Given that the model contains only one cointegrating relationship, we utilize the Engle-Granger two-step method (2OLS) to conduct the analysis.

completely, which reduces overfitting problems, decreases parameter uncertainty, and improves the model's forecasting accuracy.

The Bayesian estimation method can be used to estimate the VAR-d as presented in Eq. (3), but it can also validly estimate a VAR-l as presented in Eq. (2). Thus, in contrast to an OLS-based VAR, having stationary data is not a requirement to estimate a BVAR (Phillips, 1991; Sims et al., 1990). That is the case as Bayesian estimation has the same shape regardless of whether the data are stationary or non-stationary, and consequently, one can validly estimate a BVAR-l.¹⁶

There is no consensus in the economic literature as to whether one should employ a BVAR-l or a BVAR-d. However, in the housing literature, the BVAR-l is used relatively more often.¹⁷ The popularity of the BVAR-l over the BVAR-d can presumably be explained by the fact that estimation in levels does not result in the information loss that is caused by differencing the data. Thence, similarly to a VECM, a BVAR-l includes information on variables in their levels. Yet, while a (B)VECM contains only stationary combinations of variables in levels, a BVAR-l simply contains all variables in their levels.¹⁸ Thence, the BVAR-l contains more potential long-run relationships that have to be estimated and does not restrict variables to converge to their equilibrium in the long run, and consequently, it is less efficient compared to a (B)VECM (Engle & Yoo, 1987).¹⁹ Furthermore, the BVAR-l includes non-stationary combinations of variables, while the (B)VECM only includes stationary combinations of variables, some of these non-stationary combinations might seem valid, but could turn out to be spurious.

For Bayesian models, it is common practice to utilize the Minnesota prior due to its ability to effectively forecast time series (Koop, 2017; Litterman, 1979). However, it is important to note that the prior specification for a BVAR-l differs from that of a BVAR-d as the former includes non-stationary variables, while the latter includes stationary variables. For non-stationary variables, the Minnesota prior incorporates the wisdom, that these time series generally follow a random walk. Therefore, it sets the prior mean of the first lag of own variables to one and it is set to zero for any other parameters (Koop & Korobilis, 2010a; Lütkepohl, 2005). As a result, this prior essentially mimics a random walk, which tends to forecast non-stationary variables

¹⁶ Please refer to Sims and Uhlig (1991) for more information on this topic. Suffice it to say that for OLS, the conditional likelihood function can be considered as a function of the data given the parameters, which may not be standard in case of non-stationarity. In contrast, the Bayesian approach views the conditional likelihood function as a function of the parameters given the data, which is standard.

¹⁷ For example, Das et al. (2009), Gupta and Das (2010), and Wu et al. (2017) employ a BVAR-d to estimate house prices. While Cuestas (2017), Emiliozzi et al. (2018), Gupta and Das (2008), Gupta and Miller (2012b), Hassani et al. (2017) and Hanck and Prüser (2020) estimate a BVAR-l. Note, however, that Hassani et al. (2017) predicted home sales instead of house prices.

¹⁸ For the model estimated in this study, the cointegration rank (r) is one. As a result, only one cointegrating relationship is estimated in the (B)VECM. For the BVAR-l this restriction is not implied, and consequently, this model includes three potential long-run relationships ($r = n$).

¹⁹ Nonetheless, in the literature it is often argued that restrictions can also result in misspecification of the model. A more Bayesian approach would be to put an unlikely prior on these long-run relationships, rather than ruling them out completely. Please refer to Appendix 3 for more information on this topic.

quite well.²⁰ In contrast, for stationary variables, it is customary to set the prior mean of all variables to zero. This reduces the model's parameters to white noise, which is known to accurately forecast stationary time series (Koop & Korobilis, 2010a).

The main merit of the Minnesota prior, that hence applies to the BVAR-d as well as the BVAR-l, is that the degree to which variables are shrunk differs per parameter. More distant lags of parameters (e.g., y_{t-4}) are expected to yield less information than more recent lags (e.g., y_{t-1}). This insight is incorporated by putting a stronger prior on more distant lags, hence shrinking these parameters relatively more. Moreover, the Minnesota prior includes the sensible notion that most of the variation in each of the variables is accounted for by own lags. This is enforced by putting a stronger prior on lags of other variables than on lags of the variable itself.²¹

3.4 BVECM

In the previous section we mentioned how a (B)VAR-l is different from VECM-like models; in essence, a VECM is a restricted VAR-l. These restrictions tie the variables together in the long run and make the VECM more efficient than an unrestricted VAR. Thence, imposing restrictions could result in a better forecasting accuracy.²² Yet, the Bayesian version of a VECM, the BVECM, has, to the best of our knowledge, not been used to forecast house prices at a macro level. The BVECM has only been used in a regional (spatial) model to capture price spillovers across housing markets (Gupta & Miller, 2012a; Nneji et al., 2015) and to predict house prices at the regional level (Gupta & Miller, 2012b).²³ Thence, the literature lacks an application and, more importantly, an evaluation of the forecasting power of the BVECM to predict house prices at the macro level.

The estimation of a BVECM is highly similar to that of a BVAR. The only difference is the need for a prior on the error correction term Π . This is, however, complicated by the fact that Π involves a product of parameters $\Pi = \alpha\beta'$, which introduces identification issues.²⁴ Because of these identification issues using standard priors is problematic and can lead to an improper posterior distribution (Kleibergen & Van Dijk, 1994; Koop et al., 2005).

²⁰ It should be noted that the prior mean suggests that the variables are integrated of order one I(1) but not cointegrated. Although technically speaking, it does not eliminate the possibility of cointegration. In order to favor cointegration, additional priors have been developed. These priors are outlined and applied in Appendices B and C.

²¹ For both models, we follow common practice and set the tightness of the prior in accordance with Litterman (1986) and Cuestas (2017).

²² However, it is important to note that incorrectly imposing restrictions could lead to misspecification of the model, thereby reducing the model's forecasting performance.

²³ Moreover, these papers use standard priors and therefore do not consider the identification issues mentioned below, which could lead to incorrect estimation.

²⁴ In particular, please note that Π relies on a combination of α and β' which is not unique. Thence, a nonsingular matrix A can affect α and β' without affecting their product Π . Put differently; $\Pi = \alpha\beta'$ is equivalent to $\Pi = \alpha A A^{-1} \beta'$. In order to identify the model, one needs to put restrictions on α and β' . However, even if global restrictions are imposed, a local identification issue still occurs if $\alpha = 0$. Please refer to Kleibergen and Van Dijk (1994) and Villani (2005) for more information on this topic.

In order to avoid improper estimation, in the main body of the article, we use the cointegration space approach.²⁵ Since the cointegration space approach focuses on the space spanned by the vector of long-run parameters (β')—rather than the values of the vector—the prior is able to tackle the above-mentioned identification issues (Koop et al., 2005, 2010b; Villani, 2005).²⁶ Moreover, as is customary for this approach, we put a noninformative prior on the speed of convergence (α) (Koop et al., 2010b). And, in line with the estimation of our BVAR, for the short-run parameters (β) we employ a Minnesota prior with a prior mean of zero.

4 Model Evaluation

To evaluate each model's predictive ability, we assess their out-of-sample—rather than in-sample—performance. This is because, as already hinted at in the introduction, in-sample errors are likely to underestimate forecasting errors (Makridakis et al., 1982). Method selection and estimation are namely designed to optimize the fit of the model on historical data, but history is unlikely to repeat itself, at least exactly. Thence, overfitting problems will only become apparent when looking at the out-of-sample performance of the models.

The entire dataset ranges from 1995 Q1 to 2020 Q4. However, to assess the out-of-sample performance of the models, we split the data into two parts: (1) the training set, which is used to estimate the parameters, and (2) the test set, which is used for model evaluation. We produce 1–40-quarter-ahead forecasts, starting with the training that ranges from 1995 Q1 to 2005 Q4. To eliminate the possibility that an arbitrary choice of sample split affects the model's performance, we rely on a rolling-origin evaluation approach (Tashman, 2000). That is, we update the sample one quarter at a time, and repeat the procedure until the end of the sample in 2011 Q4. For each forecast, we calculate the Mean Absolute Percentage Error (MAPE) and we average them out to end up with our final result.²⁷

²⁵ In Appendix 3 we compare our results to those obtained by using a recently developed prior by Giannone et al. (2019).

²⁶ The cointegrating space $\varphi = sp(\beta')$ is the space spanned by β' . It is an r -plane in n -space in which the cointegrating vectors β' lie. Please refer to Koop et al. (2005) for a graphical representation of the cointegration space and vector. The central location of φ is $\varphi^H = sp(H)$ and its dispersion is controlled by τ .

²⁷ In this study, we employ the MAPE as the measure of forecasting accuracy, rather than the root mean squared forecasting error (RMSFE) due to the latter's arbitrary penalization of large errors. For further elaboration on this topic, the reader is referred to Tashman (2000) and/or Willmott and Matsuura (2005). Despite this, we have conducted a comparison of the results obtained using the MAPE with those obtained using the RMSFE and found that both measures yield similar conclusions. The results of this comparison are available upon request.

5 Results

In this section, we first present identification tests. Thereafter, the forecasting performance of the models is evaluated. Subsequently, the drivers of the two best-performing models are illustrated. Finally, the performance of these two winning models is assessed along the house price cycle.

5.1 Identification Tests

From Sect. 3, it followed that time series properties (i.e., stationarity and cointegration) are an important determinant of which empirical models can validly be estimated. Thence, the natural starting point of a forecasting exercise is to examine the time series properties of variables used herein.

We first assess whether the models' variables are stationary. Please recall that one needs stationarity to properly estimate a VAR, while for BVARs stationarity is not a requirement. We utilize two non-stationarity tests: the Augmented Dickey-Fuller (ADF) and the Phillips-Perron (PP) test, as well as one stationarity test: the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, to determine whether the variables are stationary. The test results are presented in Table 2, and they indicate that all variables are non-stationary.²⁸

Repeating the tests for the first-order difference, we found evidence that the first differences of these variables are stationary for all but one case. Specifically, only the ADF test for the first-order difference of house prices failed to reject the null hypothesis of non-stationarity. However, considering that all other tests indicate stationarity and the ADF test is known to have low power (Afriyie et al., 2020), it can be inferred that all the series are non-stationary but their first difference is stationary (i.e., all the time series are integrated of order one). Therefore, to prevent a spurious regression, we run the VAR in first differences.

Subsequently, the number of cointegrating relationships (or the cointegration rank) is tested. We need at least one cointegrating relationship to estimate a VECM. For a BVECM this test result is not a strict requirement, yet it is not uncommon to test for the number of cointegrating relationships. According to the results of the Johansen test presented in Table 3, both the trace and the eigenvalue test reject the null hypothesis of no cointegration, but the tests fail to reject the maximum of one cointegrating relationship at a 5% level. Thence, the Johansen test provides evidence on the existence of at least one cointegrating relationship and hence justifies the use of a (B)VECM.

²⁸ In addition to the tests presented in Table 2, we employed the Zivot & Andrews test to investigate stationarity while considering the possibility of a structural break. This additional analysis was necessary because some variables exhibited significant fluctuations, as evident from Fig. 1. Due to brevity, these results are available upon request. Nonetheless, it is important to mention that the results of the Zivot & Andrews test confirm that all variables in our study are non-stationary.

Table 2 Non-stationarity & stationarity tests

	Test statistic levels			Test statistic first difference		
	P_t	$b_t^{\max,c}$	$b_t^{cd,u}$	ΔP_t	$\Delta b_t^{\max,c}$	$\Delta b_t^{cd,u}$
ADF	-1.72	-2.18	-1.18	-1.86	-4.12***	-5.37***
PP- τ	-0.39	-2.09	-1.14	-8.12***	-9.96***	-7.43***
PP- α	-1.04	-7.51	-3.75	-93.73***	-103.20***	-71.43***
KPSS	1.67***	0.81***	2.10***	0.36	0.14	0.26

The table presents the test statistics of the Augmented Dickey-Fuller (ADF) test, the Phillips–Perron test (PP), and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. The tests include both a trend and an intercept. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 3 Johansen Cointegration test

	Trace test	Eigenvalue test
$r \leq 2$	4.81	4.81
$r \leq 1$	18.34*	13.53
$r = 0$	41.34***	23.00**

r reflects the number of cointegrating relationships. All tests include an intercept. Tests that include both trend and intercept were also conducted, however, as the trend parameter was found to be insignificant, it was removed from the analysis. The optimal lag orders for the tests were determined using the Akaike information criterion, with a maximum of six lags. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

5.2 Forecasting Power

This section compares the predictive ability of the models discussed in Sect. 2. Figure 2 shows the forecasting errors of the models. Note that the same information is presented in tabular format in Appendix 1.

For ease of comparison, each subfigure displays a maximum of three models. As such, the BVAR-d model presented in subfigure b is identical to the BVAR-d model presented in subfigures c and d. When looking at all subfigures combined, we see a familiar pattern; i.e., for short-term forecasts, the forecasting errors are relatively small, but when forecasting further ahead the errors increase significantly. When forecasting ten years ahead the errors are severe, the model with the smallest forecasting error is on average off by circa 17 %. Thence, such long-term forecasts provide little valuable information. Nonetheless, for this research, it is interesting to include them as it is often argued that the benefits of a (B)VECM become apparent over longer forecasting horizons.

The first subfigure compares the predictive power of the two non-Bayesian models: the VAR and the VECM. It follows from the figure that the VAR performs slightly better at one-quarter-ahead forecasts, yet for all other forecasting horizons, the VECM prevails. This result is in line with economic theory (or practice), which states that VECMs are expected to outperform VARs over longer forecasting periods

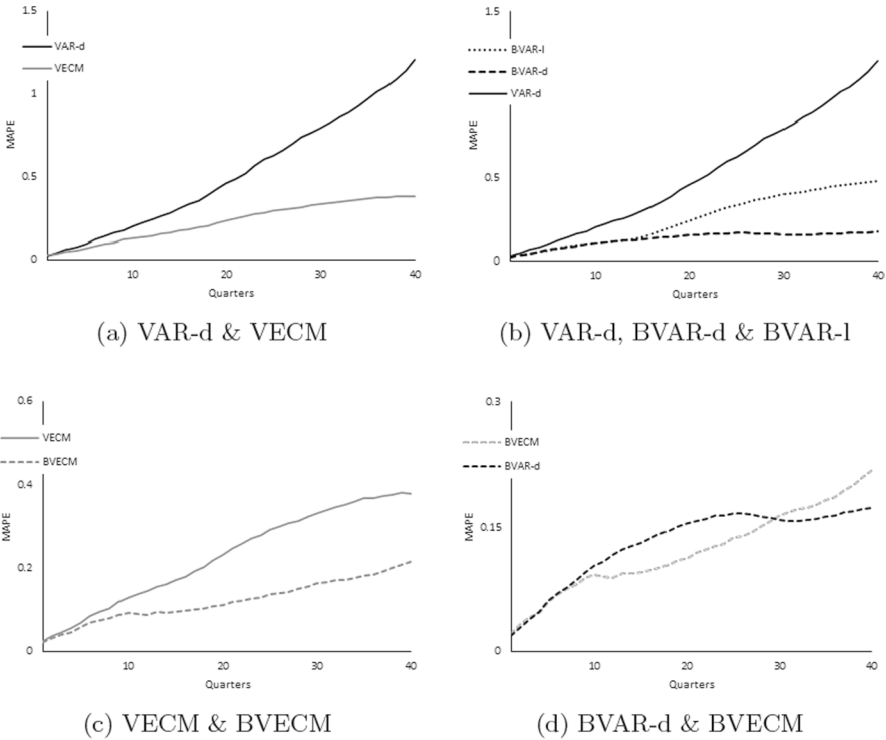


Fig. 2 Pseudo out-of-sample forecasting error

(Engle & Yoo, 1987). Note, however, that the VECM outperforms the VAR at relatively short forecasting horizons already. This is most likely the case as house prices do not drift off, they have a tendency to almost immediately move towards the long-run equilibrium (although it can still take a long time before the equilibrium is fully restored). Moreover, note that the difference in forecasting accuracy increases over the forecasting horizon, indicating that long-run relationships become more important when forecasting further ahead. Thus, modeling long-run relationships in VECM-form is especially beneficial when forecasting house prices more than one-quarter ahead, and this benefit becomes even more apparent on longer forecasting horizons.

Subsequently, in subfigure b, the forecasting accuracy of the VAR-d, BVAR-d, and BVAR-l are compared. At all forecasting horizons, the BVAR-d and BVAR-l outperform the VAR-d. When forecasting further ahead, this difference, in terms of forecasting accuracy, increases. Thus, the two Bayesian VARs always outperform the OLS VAR, indicating that using prior information improves the forecasting performance of the models significantly. When comparing the two BVARs, it is evident that their forecasting performance is comparable for predictions up to three and a half years ahead. However, as the forecasting horizon extends, the BVAR-d prevails over the BVAR-l. Thus, although a non-stationary BVAR is allowed, it does not seem to beat a stationary BVAR.

Subfigure c compares the VECM with the BVECM. For this comparison, we also find that at all forecasting horizons, the Bayesian model outperforms its non-Bayesian counterpart. However, in contrast to the VAR models (panel b), the difference in terms of forecasting accuracy is smaller. Moreover, the difference between the VECM and the BVECM stagnates after eight years (about 32 quarters). Overall, when we combine the results of panels b & c, we conclude that for both the VECM-like and VAR-like models, the Bayesian version outperforms its OLS counterpart, indicating that shrinking the models' parameters towards a parsimonious benchmark is beneficial. Therefore, when forecasting house prices, Bayesian estimation techniques would be preferred over their OLS counterparts at all forecasting horizons.

Finally, in Panel d we compare the two winning Bayesian models: the BVECM and the BVAR-d. For these models, the results are less clear-cut: for forecasts up until three quarters ahead the BVAR-d slightly outperforms the BVECM, for forecasts from three quarters to seven years ahead the BVECM prevails, yet for forecasts further ahead the BVAR-d wins again. Since the results are more comparable, we use the Diebold–Mariano (DM) test in Appendix 1 to determine whether the two forecasts differ. Although the difference between the two models is not significant for all forecasting periods, the results of the DM test suggest that the BVAR-d outperforms in the short-run, the BVECM performs best for medium-term forecasts, and the BVAR-d wins again for long-term forecasts.²⁹ These results are seemingly in line with the notion that the benefits of a BVECM become apparent over longer forecasting horizons. However, in contrast to this notion, for long-term forecasts, the BVAR-d prevails.³⁰ This anomaly can most probably be explained by the fact that we fail to estimate the error-correction term adequately when forecasting this far ahead, leading to a poorer performance of the BVECM compared to the BVAR-d. All in all, for short-run forecasts, the BVAR-d appears to be the preferable option, while for medium-term forecasts the BVECM outperforms, and for long-term forecasts the BVAR-d is the more sensible choice.

5.3 Forecast Error Variance Decomposition

From the previous section, it followed that neither the BVECM nor the BVAR-d exhibited a clear advantage in terms of forecasting accuracy. Therefore, this subsection dives deeper into the differences between the BVECM and the BVAR-d. This exercise aims to provide a better understanding of what drives both models. In particular, we are interested in how much the models' variables explain the variability of house prices over time. We formally assess each variable's contribution by analyzing the effect of shocks in the models' variables. This exercise is called a forecast

²⁹ In particular, the DM test indicates that for one-quarter-ahead forecasts, the BVAR-d forecasts significantly better at a 10% level of significance. On the other hand, the BVECM performs significantly better for a range of 11–21-quarters-ahead forecasts, and even at a 1% level of significance for three- to four-and-a-half-years-ahead forecasts. However, for ten-years-ahead forecasts, the BVAR-d forecasts demonstrate better performance at a 10% level of significance.

³⁰ Thus, while for the OLS models the benefits of incorporating a long-run relationship into the model are relatively large, this is to a lesser extent the case for Bayesian models.

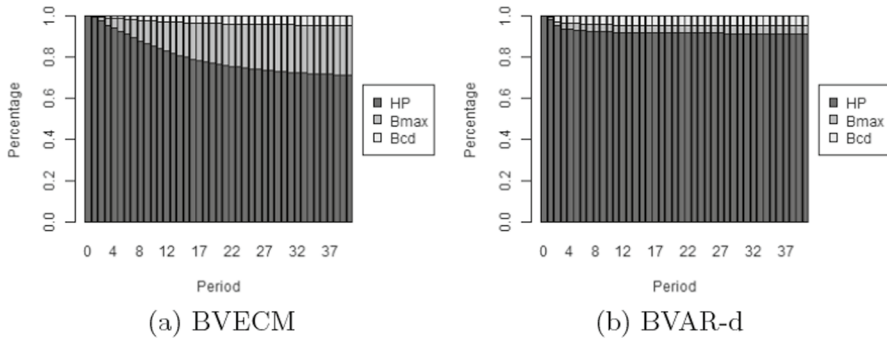


Fig. 3 Forecast error variance decomposition of house prices

error variance decomposition (FEVD), and it essentially reflects how important a shock is in explaining the variations of the variables in the model.³¹

Figure 3 reflects the FEVD of house prices for the BVECM and the BVAR-d. Up to one year ahead, the two subfigures are comparable. That is, variations in house prices are mainly determined by their past value. Thereafter, for the BVECM, the borrowing capacity of constrained households takes up a significant part of the explanation of house prices. Or, phrased differently, the error correction term gradually kicks in. For the BVAR-d this is not the case, this model mainly depends on variations in house prices. In particular, for the BVAR-d for forecasts ten years ahead, 90% of the variation in house prices comes from shocks in house prices, while this is only 70% for the BVECM. Thence, the BVAR-d resembles a simple autoregressive model that predicts that house prices tomorrow are merely explained by house prices today. The BVECM, on the other hand, incorporates an equilibrium relationship that includes the effect of changes in the borrowing capacity on house prices.

5.4 Forecasting Power Along the House Price Cycle

Having a better understanding of the differences between the BVAR-d and the BVECM, we finally assess how well both perform during different economic seasons. Or, phrased differently, we compare the forecasting accuracy of the winning models over the house price cycle. This analysis is performed as forecasters might be more interested in predicting the timing of peaks and troughs (i.e., turning points) rather than the aggregated performance of a model over the boom-bust cycle (i.e., Sect. 5.2).

Nevertheless, since house prices experienced at most two turning points during the out-of-sample period, we do not attempt a formal analysis of the competing models' ability to predict turning points. Instead, Our primary objective is to compare the forecasting accuracy of the models throughout the house price cycle. In particular, for this exercise, four different forecasting horizons are considered: one

³¹ The FEVD is needed as—due to the vector-lag form of the model—responses are not directly obvious from the models' estimates. That is, within the model everything depends on everything therefore, the most convenient way to assess responses to shocks is graphically.

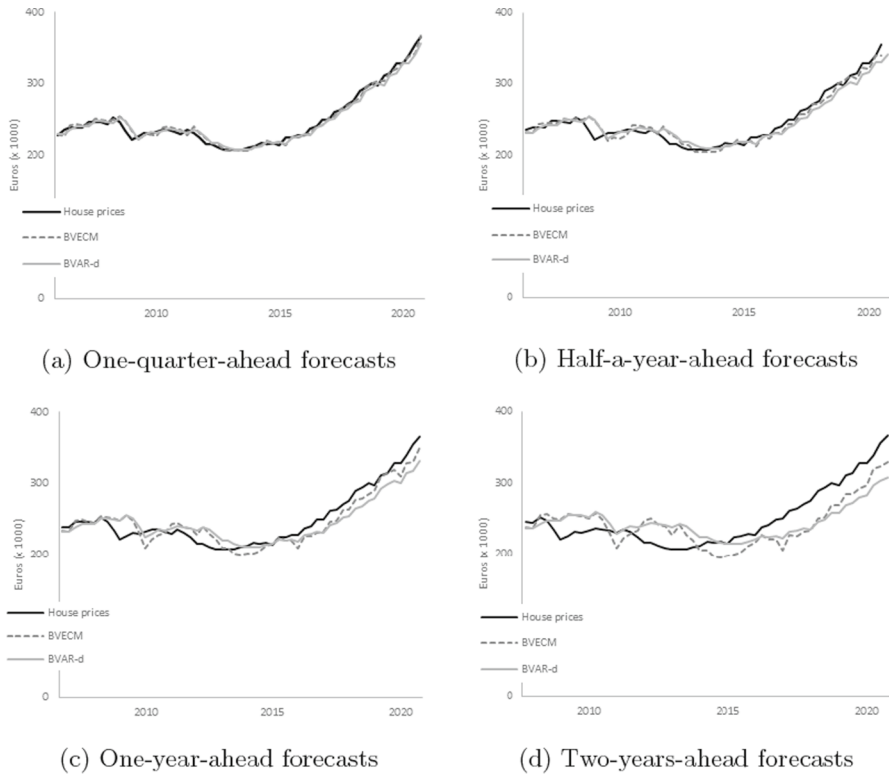


Fig. 4 Pseudo out-of-sample forecasts and actual house prices

quarter ahead, half a year ahead, one year ahead, and two years ahead. We stop at two-years-ahead forecasts as predicting turning points is a difficult matter, and consequently, one can simply not expect a model to predict a turning point that lies more than two years ahead. Subsequently, these predictions are plotted against their actual outcomes to assess which model performs best when.

The first panel of Fig. 4 shows the results for one-quarter-ahead forecasts. Differences between the BVECM and BVAR-d are small. For the BVECM we see, however, slightly more outliers than for the BVAR-d. This is true over the entire sample and is not concentrated at a specific event or economic season, indicating the BVAR-d is the preferable option for one-quarter-ahead forecasts.

For half-a-year-ahead forecasts, the results are similar, but the differences between the BVECM and the BVAR-d are slightly larger. The BVAR-d seems to be better equipped to predict the temporary revival of house prices after the financial crisis (i.e., 2009). The BVECM, on the other hand, performs slightly better in the period of fast-rising house prices of 2015–2020. The latter is most likely the case as house prices increased rapidly in that period due to decreases in the interest rate. The lower interest rate increased households’ borrowing capacity ($b_t^{\max,c}$ & $b_t^{cd,u}$) and this allowed them to buy a more expensive home; as such, these house price

increases are explained by their fundamentals (i.e., the long-run relationships). This equilibrium mechanism is incorporated in the BVECM, while the BVAR-d fails to disentangle speculative price movements from these founded price movements.

For one- and two-years-ahead forecasts, the differences between the models become even more apparent. From the two subfigures, it still follows that the BVECM exaggerated the bust caused by the financial recession. Strikingly, however, the revival of the housing market in 2014, is better captured by the BVECM than by the BVAR-d. The BVECM namely correctly pointed to a price increase in 2014, while the BVAR-d still predicted a decrease. In fact, the BVECM was able to predict the exact timing of this turning point. Thus, during a crisis, a BVAR-d may be more appropriate, while the BVECM might provide more accurate predictions after the crisis.

6 Conclusion

This study investigated five empirical methods to predict house prices in the Netherlands. In particular, OLS-based estimation methods were compared to their Bayesian counterparts. The Bayesian models were expected to outperform, as they are able to tackle overfitting problems. Our findings substantiate this hypothesis: Bayesian models have a better forecasting performance compared to the OLS models. Notably, this finding was observed for the relatively parsimonious three-variable model, indicating that overfitting can occur even with a small number of variables. Therefore, when predicting house prices, it might be wise to opt for Bayesian models, particularly when the sample is relatively small and/or the model contains many parameters.

Within the class of Bayesian models, we compared three variants: a BVAR in levels, a BVAR in differences, and a BVECM. While the BVAR in levels is commonly utilized in the housing literature, it becomes apparent that the BVAR in differences consistently exhibits better performance than the BVAR in levels. As for the remaining models, namely the BVAR in differences and the BVECM, the results are less clear-cut. Which model performs best in terms of forecasting accuracy seems to depend on two things: (i) the forecasting horizon, and (ii) the economic season. The BVECM incorporates long-run relationships and this seems to be particularly beneficial for medium-term forecasts. However, in periods of market uncertainty, such as during a recession, incorporating these long-run equilibrium relationships might not provide as much insightful information. Instead, a more useful approach could involve predicting that tomorrow will likely resemble today. Incorporating long-run equilibrium relationships seems, however, especially valuable if one seeks to predict when the recession will end, or to forecast house prices in normal times (i.e., post-crisis). Therefore, a researcher interested in predicting house prices might consider applying both Bayesian models and

carefully assess the models' forecasting results in light of the current economic state and the desired forecasting horizon.

Nonetheless, it is important to acknowledge that while our study sheds light on the relative performance of these Bayesian models, the practical applicability of such models during crisis periods should be approached with caution. The lack of reliable prior information during crises and the potential for rapid and unforeseen changes in economic conditions could impact the models' forecasting performance. Future research could explore alternative modeling approaches (e.g., regime-switching models) to better adapt to sudden economic shocks and identify optimal forecasting strategies during crisis periods.

Furthermore, it is essential to emphasize that this paper solely focuses on the Dutch housing market. In the Netherlands, mortgage debt is substantial, rendering mortgage conditions pivotal in determining house prices. Therefore, the long-term relationship used in the (B)VECM incorporates this insight. Our results support this equilibrium; in the short-run house prices are mainly driven by their lagged value, yet, this effect slowly fades away and the borrowing capacities take up a part of explaining house prices. Nevertheless, it is important to acknowledge that in countries where a greater share of properties are purchased outright or with minimal mortgage financing, the impact of mortgage conditions on house prices might be limited. For such contexts, integrating an alternate long-term relationship into a BVECM—potentially one based on wealth—could offer more insight. However, further research is necessary to delve into this subject.

Appendix 1. Forecasting Errors

Table 4 shows the forecasting errors as presented in Fig. 2. Each row represents a different forecasting horizon, expressed in quarterly intervals. While not displayed in Fig. 2, the column labeled "RW" showcases the forecasting errors for a random walk model with drift. This random walk model serves as a baseline or reference point for evaluating the effectiveness of more complex forecasting models.

The column labeled "Min" indicates the model with the smallest forecasting error. As highlighted in the main text, the winning model is consistently either the BVAR-d or the BVECM. To test the differences in forecasting performance between the BVECM and BVAR-d, the "DM" column presents the results of the Diebold-Mariano test for absolute forecasting errors using Newey-West errors (Diebold & Mariano, 1995). The null hypothesis posits that the two models exhibit equivalent forecasting accuracy, while the alternative hypothesis suggests differences in their forecasting accuracies.

Table 4 Pseudo out-of-sample forecasting error

Quarter	RW	VAR-d	VECM	BVAR-1	BVAR-d	BVECM	MIN	DM
1	0.021	0.022	0.026	0.022	0.019	0.022	BVAR-d	- 1.80*
2	0.033	0.043	0.037	0.032	0.029	0.034	BVAR-d	- 1.19
3	0.046	0.062	0.047	0.039	0.039	0.041	BVAR-d	- 0.53
4	0.062	0.076	0.057	0.048	0.048	0.047	BVECM	0.18
5	0.080	0.095	0.070	0.060	0.061	0.060	BVECM	0.06
6	0.096	0.121	0.085	0.072	0.070	0.069	BVECM	0.04
7	0.113	0.145	0.096	0.081	0.078	0.075	BVECM	0.25
8	0.130	0.162	0.105	0.088	0.086	0.081	BVECM	0.59
9	0.147	0.177	0.119	0.098	0.096	0.089	BVECM	0.83
10	0.163	0.200	0.131	0.105	0.103	0.093	BVECM	1.18
11	0.181	0.223	0.137	0.113	0.108	0.091	BVECM	2.22**
12	0.198	0.241	0.145	0.119	0.116	0.089	BVECM	2.89***
13	0.212	0.259	0.155	0.126	0.123	0.095	BVECM	2.86***
14	0.223	0.278	0.163	0.132	0.127	0.095	BVECM	2.93***
15	0.234	0.305	0.173	0.146	0.131	0.096	BVECM	3.02***
16	0.247	0.331	0.181	0.164	0.136	0.098	BVECM	3.28***
17	0.257	0.356	0.193	0.182	0.142	0.101	BVECM	3.15***
18	0.268	0.389	0.204	0.202	0.146	0.104	BVECM	2.81***
19	0.278	0.427	0.219	0.223	0.151	0.110	BVECM	2.45**
20	0.288	0.461	0.232	0.243	0.155	0.112	BVECM	2.30**
21	0.297	0.485	0.248	0.263	0.158	0.120	BVECM	1.87*
22	0.303	0.519	0.260	0.282	0.161	0.123	BVECM	1.65
23	0.310	0.562	0.273	0.303	0.164	0.127	BVECM	1.39
24	0.315	0.601	0.281	0.320	0.164	0.130	BVECM	1.13
25	0.320	0.626	0.294	0.337	0.167	0.137	BVECM	0.87
26	0.321	0.658	0.302	0.351	0.166	0.140	BVECM	0.71
27	0.320	0.698	0.309	0.367	0.166	0.144	BVECM	0.46
28	0.316	0.736	0.315	0.378	0.163	0.151	BVECM	0.09
29	0.311	0.763	0.326	0.391	0.161	0.157	BVECM	- 0.23
30	0.305	0.790	0.333	0.400	0.159	0.164	BVAR-d	- 0.56
31	0.300	0.824	0.341	0.410	0.158	0.168	BVAR-d	- 0.72
32	0.292	0.862	0.348	0.418	0.157	0.172	BVAR-d	- 0.85
33	0.287	0.891	0.354	0.428	0.159	0.174	BVAR-d	- 0.84
34	0.282	0.928	0.362	0.437	0.160	0.178	BVAR-d	- 0.92
35	0.277	0.969	0.368	0.447	0.163	0.183	BVAR-d	- 1.02
36	0.272	1.014	0.370	0.456	0.165	0.187	BVAR-d	- 1.09
37	0.268	1.046	0.375	0.463	0.168	0.194	BVAR-d	- 1.25
38	0.263	1.087	0.378	0.470	0.170	0.201	BVAR-d	- 1.46
39	0.260	1.138	0.382	0.476	0.172	0.210	BVAR-d	- 1.68
40	0.257	1.205	0.381	0.482	0.174	0.218	BVAR-d	- 1.91*

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Appendix 2. The Combination Prior

Doan et al. (1984) and Sims (1993) suggested supplementing the Minnesota prior of a BVAR-I with additional priors that favor unit roots and cointegration. These additional priors are respectively known as the “sum-of-coefficients” and “dummy-initial-observation” prior and they were motivated by the desire to prevent an overly large share of the variation in the data from being explained by the deterministic component.

In the literature, it is common practice to combine the Minnesota prior with the sum-of-coefficients prior and the dummy-initial-observation prior, which is known as the “combination prior” (Giannone et al., 2015). Since this combination prior aligns with the common belief that macroeconomic data often exhibit unit roots and cointegration, it tends to improve the forecasting accuracy of a BVAR-I.

In the remaining part of this section, a model with combination prior, as proposed by Giannone et al. (2015), is estimated and compared to the forecasting results of the two winning Bayesian models, namely the BVAR-d and BVECM. When comparing the forecasting results of the model with the combination prior to those of the BVECM and BVAR-d, we find that their forecasting accuracy is highly similar for forecasts up to one year ahead. See Fig. 5. However, in two instances, the combination prior demonstrates a slight advantage in forecasting power compared to the other models. Nevertheless, it is important to note that, as per the DM-test, this advantage does not translate into significantly better predictions in these cases. As we extend the forecasting horizon, the performance of the model with the combination prior deteriorates, and both the BVECM and BVAR-d significantly outperform it. Therefore, for this forecasting exercise, a model with a combination prior provides little improvement compared to a BVAR-d and/or BVECM.

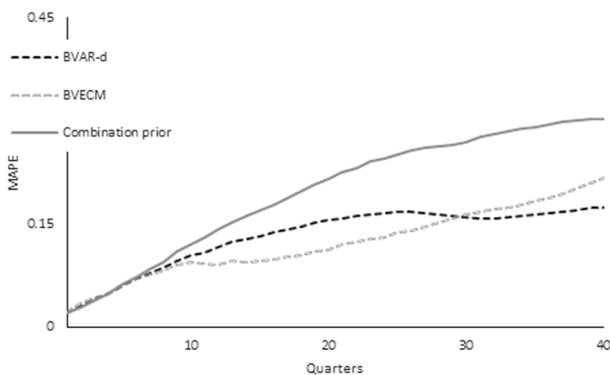


Fig. 5 Pseudo out-of-sample forecasting error, combination prior

Appendix 3. Priors for the Long Run

According to Giannone et al. (2019), the BVECM specification used in the main text is too restrictive. The authors state that this model is not flexible enough as in practice it is difficult to identify whether data are stationary or integrated. The authors include these insights into the so-called “prior for the long run” (PLR). This model can be thought of as a BVAR which includes the BVECM specification as a special case.³² Unlike the BVECM, the PLR includes all potential long-run relationships and does not exclude any potentially non-stationary long-run relationships. However, the PLR utilizes economic theory to assign higher weights to more plausible long-run relationships and lower weights to less plausible relationships.

In this section, the forecasting performance of a PLR is compared with the two winning Bayesian models: the BVAR-d and BVECM.³³ From Fig. 6, it is evident that up until a 2.5-year forecasting horizon, the forecasting performance of the models is comparable. However, in five instances, the PLR demonstrates a slight advantage in forecasting power compared to the other models. Nevertheless, it is important to note that, as per the DM-test, this advantage does not translate into significantly better predictions in these five cases. As we extend the forecasting horizon, an interesting shift occurs. The PLR model begins to underperform significantly when compared to the BVECM and the BVAR-d. Thus, compared to the BVECM and the BVAR-d, the PLR yields little improvement. Therefore, one might argue that, in this case, it would be wiser to rely on a BVAR-d and a BVECM for forecasting house prices instead of combining them into one model (i.e., the PLR).

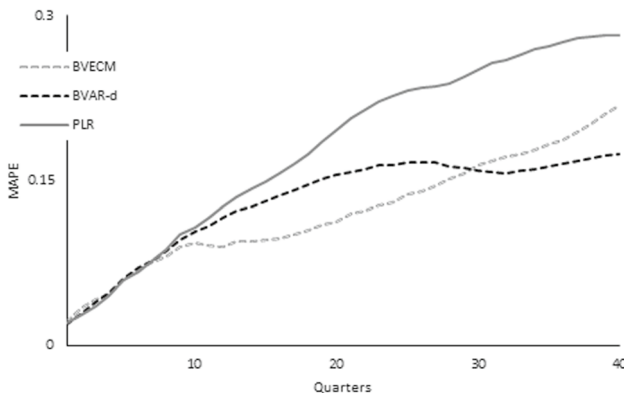


Fig. 6 Pseudo out-of-sample forecasting error, PLR

³² Please note that this prior is highly related to the sum-of-coefficients prior. In fact, Giannone et al. (2019) point out that under some conditions, the PLR simplifies to the sum-of-coefficients prior.

³³ I.e., a BVECM with priors on the cointegration space as proposed in Koop et al. (2010b).

Acknowledgements The authors would like to thank the two anonymous referees during the peer-review of the paper for their thoughtful suggestions and comments.

Funding No funding was received to assist with the preparation of this manuscript.

Conflict of interest The authors have no competing interests to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abraham, J. M., & Hendershott, P. H. (1996). Bubbles in metropolitan housing markets. *Journal of Housing Research*, 7(2), 191–207.
- Afriyie, J. K., Twumasi-Ankrah, S., Gyamfi, K. B., Arthur, D., & Pels, W. A. (2020). Evaluating the performance of unit root tests in single time series processes. *Mathematics and Statistics*, 8(6), 656–664. <https://doi.org/10.13189/ms.2020.080605>
- Bañbura, M., Giannone, D., & Reichlin, L. (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1), 71–92. <https://doi.org/10.1002/jae.1137>
- Brissimis, S. N., & Vlassopoulos, T. (2009). The interaction between mortgage financing and housing prices in Greece. *The Journal of Real Estate Finance and Economics*, 39(2), 146–164. <https://doi.org/10.1007/s11146-008-9109-3>
- Clayton, J. (1997). Are housing price cycles driven by irrational expectations? *The Journal of Real Estate Finance and Economics*, 14, 341–363. <https://doi.org/10.1023/A:1007766714810>
- Clayton, J., Ling, D. C., & Naranjo, A. (2009). Commercial real estate valuation: Fundamentals versus investor sentiment. *The Journal of Real Estate Finance and Economics*, 38(1), 5–37. <https://doi.org/10.1007/s11146-008-9130-6>
- Cuestas, J. C. (2017). House prices and capital inflows in Spain during the boom: Evidence from a cointegrated VAR and a structural Bayesian VAR. *Journal of Housing Economics*, 37, 22–28. <https://doi.org/10.1016/j.jhe.2017.04.002>
- Damen, S., Vastmans, F., & Buyst, E. (2016). The effect of mortgage interest deduction and mortgage characteristics on house prices. *Journal of Housing Economics*, 34, 15–29. <https://doi.org/10.1016/j.jhe.2016.06.002>
- Das, S., Gupta, R., & Kabundi, A. (2009). Could we have predicted the recent downturn in the South African housing market? *Journal of Housing Economics*, 18(4), 325–335. <https://doi.org/10.1016/j.jhe.2009.04.004>
- De Vries, P., & Boelhouwer, P. J. (2009). Equilibrium between interest payments and income in the housing market. *Journal of Housing and the Built Environment*, 24(1), 19–29. <https://doi.org/10.1007/s10901-008-9131-z>
- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263. <https://doi.org/10.1080/07350015.1995.10524599>
- Doan, T., Litterman, R., & Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, 3(1), 1–100. <https://doi.org/10.1080/07474938408800053>
- Emiliozzi, S., Guglielminetti, E., & Loberto, M. (2018). Forecasting house prices in Italy (Occasional Paper 463). *Bank of Italy*. <https://doi.org/10.2139/ssrn.3429819>
- Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 55(2), 251–276. <https://doi.org/10.2307/1913236>

- Engle, R. F., & Yoo, B. S. (1987). Forecasting and testing in co-integrated systems. *Journal of Econometrics*, 35(1), 143–159. [https://doi.org/10.1016/0304-4076\(87\)90085-6](https://doi.org/10.1016/0304-4076(87)90085-6)
- Fraser, P., Hoesli, M., & McAlevey, L. (2008). House prices and bubbles in New Zealand. *The Journal of Real Estate Finance and Economics*, 37(1), 71–91. <https://doi.org/10.1007/s11146-007-9060-8>
- Giannone, D., Lenza, M., & Primiceri, G. E. (2015). Prior selection for vector autoregressions. *Review of Economics and Statistics*, 97(2), 436–451. https://doi.org/10.1162/REST_a_00483
- Giannone, D., Lenza, M., & Primiceri, G. E. (2019). Priors for the long run. *Journal of the American Statistical Association*, 114(526), 565–580. <https://doi.org/10.1080/01621459.2018.1483826>
- Gupta, R., & Das, S. (2008). Spatial Bayesian methods of forecasting house prices in six metropolitan areas of South Africa. *South African Journal of Economics*, 76(2), 298–313. <https://doi.org/10.1111/j.1813-6982.2008.00191.x>
- Gupta, R., & Das, S. (2010). Predicting downturns in the US housing market: a Bayesian approach. *The Journal of Real Estate Finance and Economics*, 41(3), 294–319. <https://doi.org/10.1007/s11146-008-9163-x>
- Gupta, R., & Miller, S. M. (2012). “Ripple effects” and forecasting home prices in Los Angeles, Las Vegas, and Phoenix. *The Annals of Regional Science*, 48(3), 763–782. <https://doi.org/10.1007/s00168-010-0416-2>
- Gupta, R., & Miller, S. M. (2012). The time-series properties of house prices: A case study of the Southern California market. *The Journal of Real Estate Finance and Economics*, 44(3), 339–361. <https://doi.org/10.1007/s11146-010-9234-7>
- Hanck, C., & Prüser, J. (2020). House prices and interest rates: Bayesian evidence from Germany. *Applied Economics*, 52(28), 3073–3089. <https://doi.org/10.1080/00036846.2019.1705242>
- Hassani, H., Ghodsi, Z., Gupta, R., & Segnon, M. (2017). Forecasting home sales in the four census regions and the aggregate US economy using singular spectrum analysis. *Computational Economics*, 49(1), 83–97. <https://doi.org/10.1007/s10614-015-9548-x>
- Hort, K. (1998). The determinants of urban house price fluctuations in Sweden 1968–1994. *Journal of Housing Economics*, 7(2), 93–120. <https://doi.org/10.1006/jhec.1998.0225>
- Hott, C., & Monnin, P. (2008). Fundamental real estate prices: An empirical estimation with international data. *The Journal of Real Estate Finance and Economics*, 36(4), 427–450. <https://doi.org/10.1007/s11146-007-9097-8>
- Iacoviello, M., & Neri, S. (2010). Housing market spillovers: Evidence from an estimated DSGE model. *American Economic Journal: Macroeconomics*, 2(2), 125–164. <https://doi.org/10.1257/mac.2.2.125>
- Kleibergen, F., & Van Dijk, H. K. (1994). On the shape of the likelihood/posterior in cointegration models. *Econometric Theory*, 10(3–4), 514–551. <https://doi.org/10.1017/S0266466600008653>
- Koop, G. (2017). Bayesian methods for empirical macroeconomics with big data. *Review of Economic Analysis*, 9(1), 33–56.
- Koop, G., & Korobilis, D. (2010). Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends in Econometrics*, 3(4), 267–358. <https://doi.org/10.1561/080000013>
- Koop, G., León-González, R., & Strachan, R. W. (2010). Efficient posterior simulation for cointegrated models with priors on the cointegration space. *Econometric Reviews*, 29(2), 224–242. <https://doi.org/10.1080/07474930903382208>
- Koop, G., Strachan, R., Van Dijk, H. K., & Villani, M. (2005). Bayesian approaches to cointegration. In T. C. Mills & K. Patterson (Eds.), *Palgrave handbook of theoretical econometrics*. Palgrave MacMillan.
- Korn, R., & Yilmaz, B. (2022). House prices as a result of trading activities: A patient trader model. *Computational Economics*, 60(1), 281–303. <https://doi.org/10.1007/s10614-021-10149-y>
- Leung, C. K. Y. (2014). Error correction dynamics of House prices: An equilibrium benchmark. *Journal of Housing Economics*, 25, 75–95. <https://doi.org/10.1016/j.jhe.2014.05.001>
- Litterman, R. B. (1979). *Techniques of forecasting using vector autoregressions* (Working Paper 115). Federal Reserve Bank of Minneapolis. <https://researchdatabase.minneapolisfed.org/downloads/v118rd614>
- Litterman, R. B. (1986). Forecasting with Bayesian vector autoregressions: Five years of experience. *Journal of Business and Economic Statistics*, 4(1), 25–38. <https://doi.org/10.1080/07350015.1986.10509491>
- Liu, R., Hui, E. C. M., Lv, J., & Chen, Y. (2017). What drives housing markets: Fundamentals or bubbles? *The Journal of Real Estate Finance and Economics*, 55(4), 395–415. <https://doi.org/10.1007/s11146-016-9565-0>

- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.
- Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., & Winkler, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting*, 1(2), 111–153. <https://doi.org/10.1002/for.3980010202>
- Malpezzi, S. (1999). A simple error correction model of house prices. *Journal of Housing Economics*, 8(1), 27–62. <https://doi.org/10.1006/jhec.1999.0240>
- Meier, M. (2018). *Supply chain disruptions, time to build, and the business cycle*. University of Mannheim.
- Michielsen, T., Groot S., & Veenstra J. (2019). *Het bouwproces van nieuwe woningen*. Netherlands Bureau for Economic Policy Analysis (CPB).
- Mikhed, V., & Zemcik, P. (2009). Do house prices reflect fundamentals? Aggregate and panel data evidence. *Journal of Housing Economics*, 18(2), 140–149. <https://doi.org/10.1016/j.jhe.2009.03.001>
- Nneji, O., Brooks, C., & Ward, C. W. (2015). Speculative bubble spillovers across regional housing markets. *Land Economics*, 91(3), 516–535. <https://doi.org/10.3368/le.91.3.516>
- Oh, H., & Yoon, C. (2020). Time to build and the real-options channel of residential investment. *Journal of Financial Economics*, 135(1), 255–269. <https://doi.org/10.1016/j.jfineco.2018.10.019>
- Ozbakan, T. A., Kale, S., & Dikmen, I. (2019). Exploring house price dynamics: An agent-based simulation with behavioral heterogeneity. *Computational Economics*, 54(2), 783–807. <https://doi.org/10.1007/s10614-018-9850-5>
- Phillips, P. C. (1991). Optimal inference in cointegrated systems. *Econometrica: Journal of the Econometric Society*, 59(2), 283–306. <https://doi.org/10.2307/2938258>
- Scott, L. O. (1990). Do prices reflect market fundamentals in real estate markets? *The Journal of Real Estate Finance and Economics*, 3, 5–23. <https://doi.org/10.1007/BF00153703>
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, 48(1), 1–48. <https://doi.org/10.2307/1912017>
- Sims, C. A., Stock, J. H., & Watson, M. W. (1990). Inference in linear time series models with some unit roots. *Econometrica: Journal of the Econometric Society*, 58(1), 113–144. <https://doi.org/10.2307/2938337>
- Sims, C. A., & Uhlig, H. (1991). Understanding unit rooters: A helicopter tour. *Econometrica: Journal of the Econometric Society*, 59(6), 1591–1599. <https://doi.org/10.2307/2938280>
- Sims, C. A. (1993). A nine-variable probabilistic macroeconomic forecasting model. In R. J. Gordon (Ed.), *Business cycles, indicators, and forecasting* (pp. 179–212). University of Chicago Press.
- Stock, J. H. (1987). Asymptotic properties of least squares estimators of cointegrating vectors. *Econometrica: Journal of the Econometric Society*, 55(5), 1035–1056. <https://doi.org/10.2307/1911260>
- Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: An analysis and review. *International Journal of Forecasting*, 16(4), 437–450. [https://doi.org/10.1016/S0169-2070\(00\)00065-0](https://doi.org/10.1016/S0169-2070(00)00065-0)
- Toda, H. Y., & Phillips, P. C. (1994). Vector autoregression and causality: A theoretical overview and simulation study. *Econometric Reviews*, 13(2), 259–285. <https://doi.org/10.1080/07474939408800286>
- Tu, Q., De Haan, J., & Boelhouwer, P. J. (2017). The mismatch between conventional house price modeling and regulated markets: Insights from The Netherlands. *Journal of Housing and the Built Environment*, 32(3), 599–619. <https://doi.org/10.1007/s10901-016-9529-y>
- Tuluca, S. A., Myer, F. N., & Webb, J. R. (2000). Dynamics of private and public real estate markets. *The Journal of Real Estate Finance and Economics*, 21(3), 279–296. <https://doi.org/10.1023/A:1012055920332>
- Van der Drift, R., De Haan, J., & Boelhouwer, P. J. (2023). Mortgage credit and house prices: The housing market equilibrium revisited. *Economic Modelling* (Advance online publication). <https://doi.org/10.1016/j.econmod.2022.106136>
- Villani, M. (2005). Bayesian reference analysis of cointegration. *Econometric Theory*, 21(2), 326–357. <https://doi.org/10.1017/S026646660505019X>
- Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1), 79–82. <https://doi.org/10.3354/cr030079>
- Wu, T., Cheng, M., & Wong, K. (2017). Bayesian analysis of Hong Kong's housing price dynamics. *Pacific Economic Review*, 22(3), 312–331. <https://doi.org/10.1111/1468-0106.12232>