RUHR

ECONOMIC PAPERS

Philipp an de Meulen
Martin Micheli
Torsten Schmidt

Forecasting House Prices in Germany

RWI

#294

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Prof. Dr. Thomas K. Bauer
RUB, Department of Economics, Empirical Economics
Phone: +49 (0) 234/3 22 83 41, e-mail: thomas.bauer@rub.de

Prof. Dr. Wolfgang Leininger
Technische Universität Dortmund, Department of Economic and Social Sciences
Economics – Microeconomics
Phone: +49 (0) 231/7 55-3297, email: W.Leininger@wiso.uni-dortmund.de

Prof. Dr. Volker Clausen
University of Duisburg-Essen, Department of Economics
International Economics
Phone: +49 (0) 201/1 83-3655, e-mail: vclausen@vwl.uni-due.de

Prof. Dr. Christoph M. Schmidt
RWI, Phone: +49 (0) 201/81 49-227, e-mail: christoph.schmidt@rwi-essen.de

Editorial Office

Joachim Schmidt
RWI, Phone: +49 (0) 201/81 49-292, e-mail: joachim.schmidt@rwi-essen.de

Ruhr Economic Papers #294

Responsible Editor: Christoph M. Schmidt

ISSN 1864-4872 (online) – ISBN 978-3-86788-339-9

Ruhr Economic Papers #294

Philipp an de Meulen, Martin Micheli, and Torsten Schmidt

Forecasting House Prices in Germany

rwi

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

ISSN 1864-4872 (online)
ISBN 978-3-86788-339-9

Philipp an de Meulen, Martin Micheli, and Torsten Schmidt[1]

Forecasting House Prices in Germany

Abstract

In the academic debate there is a broad consensus that house price fluctuations have a substantial impact on financial stability and real economic activity. Therefore, it is important to have timely information on actual and expected house price developments. The aim of this paper is to measure the latest price movements in different real estate markets in Germany and forecast near-term price developments. Therefore we construct hedonic house price indices based on real estate advertisements on the internet platform ImmobilienScout24. Then, starting with a naive AR(p) model as a benchmark, we investigate whether VAR and ARDL models using additional macroeconomic information can improve the forecasting performance as measured by the mean squared forecast error (MSFE). While these models reduce the forecast error only slightly, forecast combination approaches enhance the predictive power considerably.

JEL Classification: C43, C53, R31

Keywords: House price forecasts; forecast combination; hedonic price index

November 2011

1 *All RWI. – We thank Thomas K. Bauer, Tobias Kitlinski, Simeon Vosen, and participants of the 1st Workshop on Real Estate Forecasting in Hamburg, Germany and the 3rd ReCapNet Conference in Mannheim, Germany for valuable comments.* *– All correspondence to Martin Micheli, RWI, Hohenzollernstr. 1–3, 45128 Essen, Germany. E-mail: martin.micheli@rwi-essen.de.*

1 Introduction

While real estate prices in Germany hardly moved during the recent financial crisis they started to increase markedly in 2010. The robust economic recovery, the expected expansionary monetary policy, and households' increasing desire for tangible assets and secure investment are potential reasons for this increase and might put ongoing upward pressure on real estate prices in the near future. If such real estate price increases accelerate, this may bear macroeconomic risks. The recent financial crisis has painfully demonstrated the effects of real estate price volatility on real economic activity. After a decade of increasing property prices in the US, housing prices began to decline sharply in 2007. As known to date, this led to tremendous deterioration of financial institutions' balance sheets, and triggered the subsequent financial crisis. In Germany, residential homes related mortgage debt outstanding was about 40% of Germany's GDP in 2009 (Financial Stability Board 2011). Hence, even small real estate price losses could increase stress in the financial sector which might trigger a credit crunch and push the economy into recession. Therefore, it is important to monitor real estate prices timely and to accurately predict future developments. It is the aim of this paper to perform these tasks, focusing on the German real estate market.

To shed some light onto the real estate market in Germany, we proceed in two steps. Based on a new dataset provided by the internet platform ImmobilienScout24, we construct house price indices for four different subsegments of the housing market, new as well as existing homes and apartments. Then, we use additional data and different techniques to forecast house price developments in the respective market segments over the period of six month.

The underlying dataset for constructing the house price indices contains offer prices and detailed information about the characteristics of the properties.[1] Since properties can only be considered as *incomplete* substitutes we have to take into account their heterogeneity in terms of e.g. quality, age and

[1] The use of offer prices has the advantage that they are available from the beginning of the buying process which often takes several months. These prices provide timely information about the housing market. However, the drawback is that offer and transaction prices differ and that this difference may vary over time.

location in order to construct an unbiased house price index. To calculate a quality adjusted price change, one has to rule out the effect of changing qualitative differences on the house price index. We account for that by using an hedonic regression approach, based on the seminal works by Lancaster (1966) and Rosen (1974).

While there do already exist several hedonic house price indices for German real estate markets – e.g. the HPX-hedonic of the Hypoport AG (Dübel and Iden 2008) and the vdp-Index published by The Association of German Pfandbrief Banks (vdp) (Hofer 2008) as well as the index constructed by Deutsche Bundesbank (Deutsche Bundesbank 2011), based on data provided by the BulwienGesa AG on 125 German cities – these indices often suffer from a low time frequency, substantial data delays, a lack of necessary information on property price determining characteristics or potential revisions. In turn, house price data provided by the internet platform ImmobilienScout24 provides us with a large amount of information both on house prices and property characteristics. Moreover, the data is available without any appreciable time lag, and the large number of observations allows us to construct an adjacent hedonic house price index on a monthly basis, which is not subject to index revisions. Hence, using the data of ImmobilienScout24, timely information about price movements in the housing sector is available. To attain information on future real estate prices we forecast the development of these indices.

Our selection of an appropriate forecasting approach is guided by the existing literature on house price forecasting and by data limitations. Forecasting house prices has gained renewed attention in empirical studies because of the recent collapse of house prices in the US. Almost all modern forecasting techniques, like ARIMA, ARDL, VAR, VECM, Bayesian VAR, Bayesian factor augmented VAR, DSGE and model combination approaches, have been used (Chen and Yu 2010; Gattini and Hiebert 2010; Gupta et al. 2009; Larson 2010; Rapach and Strauss 2007). Since only subgroups of models have been tested in a coherent framework so far, no model appears to be clearly superior to others. However, at least two lessons can be learned from these studies. Firstly, using additional economic variables improves the forecasting perfor-

mance of an autoregressive model. Secondly, Gupta et al. (2009) find that a BVAR with ten variables outperforms BVARS with 120 variables. These authors draw the conclusion that only a few fundamental variables are important for house price developments. In our forecasting exercise the model selection is restricted by the fact that house price data is only available since 2007. Because of these relatively short time-series, we decided to start our forecasting exercise with ARDL and VAR models as well as forecast combinations using a relatively small number of additional economic variables.
The outline of the paper is as follows: In the next section we describe the dataset and outline the construction of the house price indices based on hedonic price functions. In section 3 we present our forecasting models and compare the results of the different forecasting techniques. Section 4 concludes.

2 Index Construction

2.1 The hedonic regression approach

In general, hedonic price indices base on the notion that a good's value is determined by its observable characteristics expressed by the hedonic price function (Lancaster 1966; Rosen 1974; Triplett 2004; Maurer et al. 2004). By regressing house prices on the respective houses' characteristics, we can decompose property prices into different price determining components, with the estimated coefficients measuring the marginal effects of the respective characteristics on the corresponding house prices. Controlling for these characteristics is essential to make all individual properties comparable. Additionally, taking into account the heterogeneous attributes of properties, time fixed effects will be employed to measure the effect of the time the price announcement was made.[2]
To take into account that the coefficients may vary over time we slightly deviate from an de Meulen et al. (2011) and employ a *True Adjacent Pe-*

[2]For a detailed discussion on the construction of the employed house price index see an de Meulen et al. (2011). On the theory of hedonic price indices see Brachinger (2002).

riods Price Index. Therefore, we split our sample of T periods into $T-1$ subsamples, each consisting of 2 adjacent periods. To allow for time-variant coefficients, the hedonic price function (1) is estimated separately for each subsample $s = 1, 2, \ldots, T-1$. Hence, we conduct $T-1$ cross-section analyses. However, to control for the effect of time in each of these single regressions, we add a time dummy γ_i^s which equals zero for property i observed in the respective first period of the subsample and one, otherwise. Such an approach is referred to as the *True Adjacent Periods Price Index* (Brachinger 2002).
A representative hedonic price function for the adjacent periods t and $t+1$ is given by the following regression equation:

$$\ln p_i^s = \sum_{j=1}^{J} \beta_j^s X_{ji}^s + \tau^s \gamma_i^s + \sum_{k=1}^{K} \mu_k^s + \epsilon_i^s \ . \tag{1}$$

s refers to the subsample of observations advertised either in period t or $t+1$, where $t \in T-1$. $\ln p_i^s$ denotes the logarithm of the announced price of property i in periods t or $t+1$. X_{ji}^s reflects the realization of the j's price determining variable of house i offered in periods t or $t+1$. ϵ_i^s is a zero-mean error term. It is possible that a certain property occurs at several points in time, since an advertiser is likely to continue offering her property until it is sold. As this might lead to correlation in the error term, the calculation of standard errors will be based on a robust covariance matrix of the error term clustered at a regional level. To account for the effect of a property's location on its price, we include a set of K regional dummy variables.[3] The regression model further controls for the advertisement duration, since prices of properties with longer advertisement duration tend to be reduced over time.
The construction of the house price indices is based on house price data provided by the internet platform ImmobilienScout24. Our sample consists of an unbalanced panel of nationwide monthly observations on house prices and corresponding property characteristics for 58 periods from January 2007 to October 2011. Data on individual house prices and their respective attributes

[3]The regional classification coincides with the division of municipalities in Germany.

are derived from property offers individual advertisers place on the internet. Summary statistics for all variables are shown in Table 1.

To take into consideration the potential heterogeneity of different housing market segments, Equation (1) is estimated separately for the four different real estate categories: for new as well as existing apartments and houses. The respective samples consist of $1,228,071$ observations on new houses, $6,571,228$ observations on existing houses as well as $832,969$ and $6,450,082$ observations for new and existing apartments, respectively.

2.2 Regressors

The regressions for the four categories are based on different sets of variables.[4] For single apartments we include a property's age, the logarithm of its living space measured in m^2 as well as its number of rooms. Furthermore, the set of control variables consists of several dummy variables including information on the availability of a cellar, an escalator, a garden, a balcony as well as a built-in kitchen. A further dummy variable is included which equals one if the property is rented and zero otherwise. Last, another set of dummy variables controls for the property's category[5] and current condition as well as missing information on properties' age[6].

With the hedonic price function of new and existing homes, we again include houses' age, their logarithmic living space and surface area, the number of rooms, information on the availability of a cellar as well as different dummies describing the categories and condition[7] of offered houses.[8] For new homes,

[4]The regression results for the hedonic price equations are available upon request.

[5]We divide properties' categories into high- and low-quality objects, with the latter being omitted from the regressions, thus chosen as the reference group.

[6]New apartments are categorized into those being firstly occupied and those objects, that lack any information on their condition, with the latter being chosen as the reference group. With existing apartments, we group objects with an overall good and bad condition, with the former left out from the regression.

[7]As with apartments, new houses are categorized into those being firstly occupied and those objects, that lack any information on their condition, with the latter being chosen as the reference group. Accordingly, existing houses are grouped into those which are in good and bad condition, with the former left out from the estimation.

[8]Categorization is done with respect to detached houses, terraced houses and other types of houses, with terraced houses being omitted from the regressions.

an additional dummy variable is included, which is equal to one if the house is still under construction.

2.3 Index Construction

Based on the estimation results of Equation (1) for the four categories, we calculate the corresponding house price indices. Since property price indices are to assess the *quality adjusted* price inflation of the respective property category over time, we make use of the respective $T-1$ estimated coefficients $\tau^s \; \forall \; s = 1, \ldots, T-1$ to construct the indices. Note that, taking into account the heterogeneous attributes of properties, τ^s measures the effect of time on the house price level. We refer to the estimated coefficients as $\hat{\tau}^s$. The antilogarithm of $\hat{\tau}^s$ estimates the inflation of the house price level between t and $t+1$, while keeping the quality of houses constant. Accordingly, $e^{\hat{\tau}^{t+1}}$ estimates the house price inflation between $t+1$ and $t+2$. Multiplying $e^{\hat{\tau}^{t+1}}$ with $e^{\hat{\tau}^t}$ then gives an estimate of the inflation between t and $t+2$, which reflects the index value of period $t+2$. In general, the index value of a certain period $l \leq T$ is then constructed as follows:

$$\prod_{k=1}^{l-1} e^{\hat{\tau}^k} \; .$$

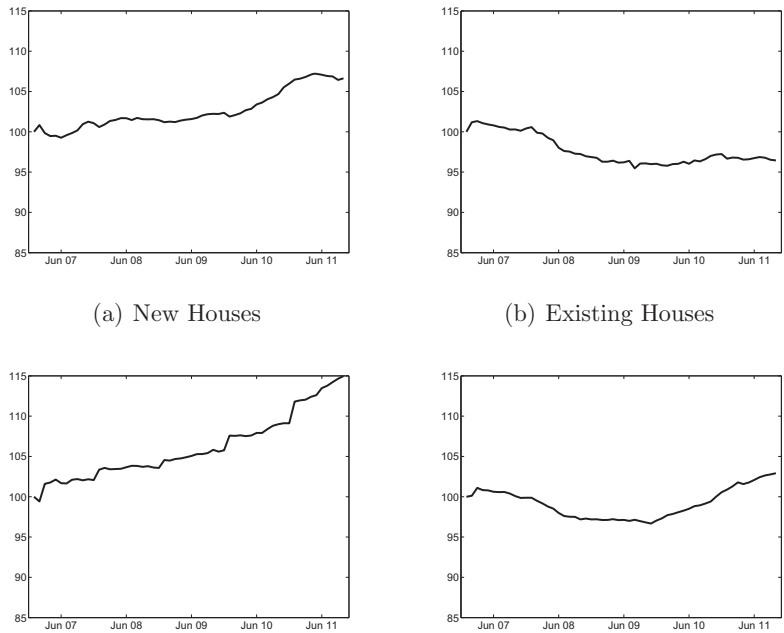
While $\hat{\tau}^t$ is an unbiased estimator of τ^t, $e^{\hat{\tau}^t}$ is a biased estimator of the antilogarithm of τ^t. To correct this bias, we add one-half of the variance of $\hat{\tau}^t$ to the estimated coefficient $\hat{\tau}^t$.[9] Moreover, to normalize the index series, we multiply each corrected value by 100, with the initial value of the series being set to 100. Then, the normalized and corrected index value of a certain period $l \leq T$ is constructed by

$$100 \cdot \prod_{k=1}^{l-1} e^{\left(\hat{\tau}^k + \frac{1}{2}\hat{\sigma}^2_{\tau^k}\right)} \; . \tag{2}$$

Figure 1 shows the price indices for the four categories of properties, calculated according to Equation (2). The price indices are in line with already

[9] See Goldberger (1968), Kennedy (1981), Teekens and Koerts (1972).

Figure 1: Housing price indices



(a) New Houses (b) Existing Houses

(c) New Apartments (d) Existing Apartments

existing indices mentioned in Section 1, see e.g. Deutsche Bundesbank (2010). It can be seen that among the segments of new buildings, the price level did apparently not suffer from the Great Recession in 2008 and 2009. For new houses as well as new apartments the price index has steadily increased since 2007. Moreover, in both segments, the price increase even accelerated in 2010. For new houses this did not change until recently, while the price index of new apartments started to decrease since May 2011. In contrast to new buildings, price indices for existing apartments and houses decreased between January 2007 and January 2010 before prices raised during 2010. As with new buildings, however, price inflation of houses did not continue in 2011 while prices even accelerated slightly for existing apartments. The different price developments of apartments and houses might be due to a demand shift in favor of apartments. The ongoing socio-economic trend towards a smaller

household size on average could be the reason for stronger price increases of apartments. Whether these developments will last or vanish in the near future will be subject of Section 3.

3 Forecasting Real Estate Prices

3.1 Forecasting Approach

As a natural starting point, we choose a simple autoregressive (AR) model as our benchmark to obtain forecasts of the development of the house prices. We compute the six month ahead forecast ($h = 6$) recursively, starting with $h = 1$, applying the following autoregressive function

$$y_{t+h} = \alpha_0 + \alpha_1 \delta_{jan} + \sum_{j=1}^{l_y} \beta_j y_{t+h-j} + \epsilon_{t+h} \ , \tag{3}$$

where y_{t+h} is the predicted logarithm of the house price level in $t + h$ based on data available until period t, with $h = 1, ..., 6$. y_{t+h} is explained by its l_y lagged values, where l_y is optimally chosen according to the Schwarz Information Criterion (SIC). Remarkably, each year in January, indices significantly increase. We conjecture that this may be due to the fact, that the number of offers is typically low in December – especially between Christmas and New Year's Day – and increases substantially in January. To control for this effect, we include a dummy variable δ_{jan} for the January of each year. Finally, ϵ_{t+h} denotes the forecast error.

Starting from the $AR(l_y)$ model, we control for 26 additional macroeconomic variables to potentially enhance the forecast accuracy, see Section 3.3. Indeed, Rapach and Strauss (2007), among others, have shown that accounting for additional variables improves accuracy of house price forecasts in other countries. To compute out of sample forecasts, we split the available sample with $T = 58$ observations into an in-sample period of $I = 36$ observations and an out-of-sample period of length $O = 22$. With a forecast horizon of six month, this allows us to evaluate 16 forecasts. We apply two different econometric models. Firstly, we recursively run ARDL regressions separately

using one of the 26 additional variables, x_i:

$$y_{t+h} = \alpha_0 + \alpha_1 \delta_{jan} + \sum_{j=1}^{l_{yi}} \beta_j y_{t+h-j} + \sum_{j=1}^{l_{x_i}} \gamma_j x_{i,t-j+1} + \epsilon_{t+h} \ . \tag{4}$$

Secondly, we calculate the six month ahead forecasts using vector autoregressive (VAR) models. We stick to the same procedure as with ARDL, estimating a bivariate VAR model for each indicator of the following form:

$$\begin{aligned} \begin{pmatrix} y_{t+h} \\ x_{t+h} \end{pmatrix} = \begin{pmatrix} \alpha_0^1 \\ \alpha_0^2 \end{pmatrix} + \begin{pmatrix} \alpha_1^1 \\ \alpha_1^2 \end{pmatrix} \cdot \delta_{jan} + \\ \sum_{j=1}^{l_{yi}} \begin{pmatrix} \theta_j^{11} & \theta_j^{12} \\ \theta_j^{21} & \theta_j^{22} \end{pmatrix} \cdot \begin{pmatrix} y_{t+h-j} \\ x_{t+h-j} \end{pmatrix} + \begin{pmatrix} \epsilon_{t+h}^1 \\ \epsilon_{t+h}^2 \end{pmatrix} \end{aligned} \tag{5}$$

Again, l_{yi} denotes the optimal lag length determined according to SIC. This gives us 26 additional forecasts for each period.

3.2 Forecast Combination

In a next step we combine the information of these various forecasts by pooling them. The literature has shown that forecast combination can enhances forecast accuracy considerably (Clemen 1989). We use three different types of combination techniques, similar to Rapach and Strauss (2007). A detailed discussion on forecast combinations can be found in Timmermann (2006).
Firstly, we employ simple combination approaches such as mean, median, and trimmed mean of our single variable six month ahead forecasts of the level of the real estate price levels. To compute the trimmed mean, we drop the lowest and highest five percent of our computed forecasts. We do this for both, the ARDL and the VAR forecasts.
Secondly, the forecasts are weighted with the forecast error of previous periods, see Stock and Watson (2004). We make use of two different weighting strategies, which represent the two extremes of weighting to previous forecast errors. We construct the individual forecast's weight from the previous period

only. The weights are given by the inverse of the fraction of the respective individual variable's mean squared forecast error in the previous period to the sum of mean squared forecast errors of all forecasts in the previous period. In addition, we do not only consider the previous period but compute the single weights from the forecast errors of all past periods.
Thirdly, we adopt the cluster approach (Aiolfi and Timmermann 2006). The single variable predictions are clustered into three groups, depending on their previous period's forecasting performance. Using only the best cluster, the forecast is the mean of the single equations' forecasts.

3.3 Explanatory Variables

As mentioned above, we refer to 26 potential predictors of the German real estate market. Among the set of variables included, there are 16 sentiment indicators by the European Commission. These indicators are based on consumer surveys asking for e.g. financial and economic situations of household as well as planned housing-related purchases and savings.[10] These indicators cover a wide range of economic information households base their consumption decisions on. Due to the fact that in Germany most residential real estate is owned or commissioned by households it is therefore likely that these indicators contain information on future house price developments.
In addition, we include three ifo indicators into our forecast models, which comprise the sentiment of German construction enterprises.[11] In contrast to household survey data covering the demand side, these variables provide information on the supply side of the German real estate market.
To not bias the house price forecast by the selection of potential predictors,

[10] In detail, these indicators are: Financial situation, last 12 month; Financial situation, next 12 month; General economic situation, last 12 month; General economic situation, next 12 month; Major purchases planned, last 12 month; Major purchases intended, currently; Price trend assessment, last 12 month; Price trend expectations, next 12 month; Savings planned, next 12 month; Savings intended, currently; Statement on financial situation of household; Consumer confidence; Unemployment expectations, next 12 month; Intention to buy a car within the next 12 month; Purchase or build a home within the next 12 month; Home improvements over the next 12 month.

[11] German Business Climate, Construction Industry; German Business Sentiment, Construction Industry; German Business Expectation, Construction Industry.

the set of sentiment indicators, which refer to private sector expectations, is complemented by macroeconomic data to have a broad-based and balanced selection of variables. Among the variables included, we differentiate between macroeconomic price data, which provides direct information on the housing price level as well as real construction data and the unemployment rate potentially affecting housing supply and demand, respectively.[12]
To simulate a real time analysis we only include information that is available at the 10th of the respective next month.

3.4 Forecasting Results

Table 2 reports the mean squared forecast errors (MSFE) for the benchmark AR model for all categories. To evaluate the forecasting performance of the single ARDL and VAR models with one additional variable, we compute the ratios of the respective models' MSFE to the MSFE of the AR process. These figures are reported in Table 2. A ratio smaller than one indicates a forecasting performance superior to the one of the AR process. With respect to the chosen out-of-sample period, the AR model predicts price developments in the housing stock accurately: the MSFE is only 0.71 index points. In turn, price developments of new homes can not be explained that well, reflected by a MSFE of 4.78 index points.
With regard to additional variables included in the analysis, only few enhance the forecasting performance substantially and none of them does so for all segments and all model specifications. However, the expected financial situation, which enhances forecast accuracy for all VAR specifications and all specifications in the apartment segments, seems to add relevant information to the forecasting equation. The forecasting performance can also be improved by including information on planned major purchases and intended savings as predictors. To test whether additional variables increase the explanatory power significantly, we make use of the modified Diebold-Mariano test. Table 2 reports the related p-values in squared brackets. Among the

[12] Consumer price; Consumer price, Housing rent (net); CDAX share price index; Interest rate of new mortgages; New orders, Construction; Building permits.

variables included, households' intended major purchases and savings highly affect future price developments in the real estate market. On the contrary, information on households' past and expected future general economic situation does not significantly improve the forecast of real estate prices. Moreover, neither the unemployment rate nor households' assessment of the past and future general economic situation seems to enhance the forecasting performance.

To improve the forecasting performance, we combine the various forecasts as described in section 3.2. The results are reported in Table 3. They are in line with the literature on forecast combination. Simple techniques increase forecast performance considerably. However, applying more sophisticated approaches further enhances predictive power. Particularly, those methods putting weight on the latest available forecast performance are significantly superior to the AR model. For the cluster approach the AR model is outperformed for all specifications. For six out of the eight forecasting approaches, the performance is significantly superior to the AR model on the 90% significance level.

4 Conclusions

The aim of this paper is to measure and forecast real estate prices in Germany. As there is no widely accepted real estate price index in Germany we first construct an index based on a dataset by ImmobilienScout24, Germany's largest internet platform for real estate prices. We construct true hedonic price indices for four categories of the housing market. Due to the extensive dataset, we are able to allow for time variation in the valuation of objects' characteristics over time.

Next, we predict the developments of the different price indices within the next six months. In order to do so, we run out of sample forecasts and evaluated the forecasts using the MSFE. Our findings suggest that the AR process can hardly be outperformed by individual models. Out of the 26 potential predictors included, only information on the financial situation of households, major purchases planned, and intended savings seem to entail

useful information on future housing prices.

However, the forecasting performance can further be increased using different forecast combination approaches. Simple combination approaches such as the mean or the median only slightly improve the performance. More sophisticated techniques as the cluster approach by Aiolfi and Timmermann (2006) substantially increase the predictive power, especially if weights are assigned according to the previous period's forecast errors only. The MSFE is decreased for all categories and both model approaches, for six out of eight approaches the predictive power is superior to the AR model on the 90% significance level.

References

Aiolfi, M. and A. Timmermann (2006). Persistence in forecasting performance and conditional combination strategies. *Journal of Econometrics 135*(1-2), 31 – 53.

an de Meulen, P., T. K. Bauer, M. Micheli, T. Schmidt, M. Kiefer, L.-H. Wilke, and S. Feuerschütte (2011). Ein hedonischer Immobilienpreisindex auf Basis von Internetdaten 2007 - 2011. *RWI Projektberichte*.

Brachinger, H. W. (2002). Statistical theory of hedonic price indices. *DQE Working Paper 01*, 54–55. Department of Quantitative Economics, University of Fribourg Switzerland.

Chen, Y.-C. and W.-C. Yu (2010). The Housing Price Forecasting and the Outbreak of the Financial Crisis: Evidence of the Arima Model. *National Cheng Kung University, Department of Economics Discussion Paper*.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting 5*(4), 559 – 583.

Dübel, H.-J. and S. Iden (2008). Hedonischer Immobilienpreisindex Deutschland: Isolierung qualitativer Hauspreismerkmale durch hedonische Regressionsanalyse aus Daten der Europace-Plattform (Hypoport AG) und Machbarkeit eines hedonischen Preisindexes für Deutschland. Technical report, Forschungsauftrag des Bundesamtes für Bauwesen und Raumordnung.

Deutsche Bundesbank (2010). Extended Investment Cycle with Stable Prices: Supply and Demand in the German Housing Market From a Longer-term Perspective. *Monthly Report 06*, 47–58.

Deutsche Bundesbank (2011). House Prices in 2010 in Germany. *Monthly Report 02*, 54–55.

Financial Stability Board (2011). Thematic Review on Mortgage Underwriting and Originating Practices. *Peer Review Report*.

Gattini, L. and P. Hiebert (2010). Forecasting and Assessing Euro Area House Prices Through the Lens of Key Fundamentals. *ECB Working Paper 1249*.

Goldberger, A. S. (1968). The Interpretation and Estimation of Cobb-Douglas Functions. *Econometrica 36*(3/4), 464–472.

Gupta, R., A. Kabundi, and S. M. Miller (2009). Forecasting the US Real House Price Index: Structural and Non-Structural Models with and without Fundamentals. *University of Pretoria, Department of Economics Working Paper 27*.

Hofer, T. (2008). vdp – Transaktionsdatenbank und – Immobilienpreisindex: Stand, Ergebnisse, Perspektiven. Technical report, Verband deutscher Pfandbriefbanken e.V.

Kennedy, P. E. (1981). Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations. *The American Economic Review 71*(4), p. 801.

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *The Journal of Political Economy 74*(2), 132–157.

Larson, W. D. (2010). Evaluating Alternative Methods of Forecasting House Prices: A Post-Crisis Reassessment. *RPF Working Paper 004*.

Maurer, R., M. Pitzer, and S. Sebastian (2004). Hedonic Price Indices for the Paris Housing Market. *Allgemeines Statistisches Archiv 88*(3), 303–326.

Rapach, D. E. and J. K. Strauss (2007). Forecasting Real Housing Price Growth in the Eighth District States. *Federal Reserve Bank of St. Louis Regional Economic Development 3*(2), 33–42.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *The Journal of Political Economy 82*(1), 666–677.

Stock, J. H. and M. W. Watson (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting 23*(6), 405–430.

Teekens, R. and J. Koerts (1972). Some Statistical Implications of the Log Transformation of Multiplicative Models. *Econometrica 40*(5), 793–819.

Timmermann, A. (2006). Chapter 4 forecast combinations. Volume 1 of *Handbook of Economic Forecasting*, pp. 135 – 196. Elsevier.

Triplett, J. (2004). Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products. OECD Science, Technology and Industry Working Papers 2004/9, OECD Publishing.

Table 1: Summary Statistics

	Existing Houses		New Houses		Existing Ap.		New Ap.	
	mean	*std*	*mean*	*std*	*mean*	*std*	*mean*	*std*
Price	258497.4	195995.8	274535.0	164536.8	138882.8	130993.2	273831.9	186092.0
Price (square meter)	1591.8	828.8	1990.9	770.6	1636.7	873.6	2645.3	918.0
Age	30.7	31.3	−0.1	0.7	29.1	30.1	−0.1	0.7
Living space	162.05	62.9	135.08	31.39	80.4	35.8	100.45	37.3
base area	795.5	856.02	489.32	335.69				
Rooms	5.7	1.8	4.8	1.0	2.9	1.1	3.3	1.0
Cellar	0.2	0.4	0.1	0.4	0.3	0.5	0.4	0.5
No cellar	0.3	0.5	0.3	0.5	0.3	0.5	0.2	0.4
Elevator					0.2	0.4	0.6	0.5
No elevator					0.2	0.2	0.1	0.1
Garden					0.2	0.4	0.3	0.5
No garden					0.1	0.3	0.1	0.3
Balcony					0.7	0.5	0.9	0.4
No balcony					0.1	0.3	0.0	0.1
Built-in kitchen					0.3	0.5	0.1	0.2
No built-in kitchen					0.1	0.3	0.2	0.4
Rent out					0.2	0.4	0.0	0.1
Detached house	0.6	0.5	0.6	0.5				
Terraced house	0.2	0.4	0.4	0.5				
Other houses	0.18	0.39	0.06	0.24				
Apartment(regular quality)					0.8	0.4	0.7	0.5
Apartment (high quality)					0.2	0.4	0.3	0.5
First occupation			0.7	0.4			0.8	0.4
Good condition	0.5	0.5			0.6	0.5		
Under construction			0.2	0.4			0.3	0.4
Observations	6,571,228		1,228,071		6,450,082		832,969	

Table 2: Forecast Results of the VAR and Individual ARDL Models

	Existing Houses		New Houses		Existing Ap.		New Ap.	
Predictor	VAR	$ARDL$	VAR	$ARDL$	VAR	$ARDL$	VAR	$ARDL$
AR MSFE	0.71		4.78		2.89		0.97	
German Business Climate; Construction Industry	1.52	1.17	1.56	1.76	0.93	0.71	1.77	1.47
					[0.34]	[0.05]		
German Business Sentiment; Construction Industry	1.99	2.15	1.33	1.48	0.63	0.42	1.10	0.91
					[0.03]	[0.03]		[0.36]
German Business Expectation; Construction Industry	0.77	0.58	1.45	1.59	1.41	1.13	1.88	1.66
	[0.30]	[0.17]						
Financial situation, last 12 months	1.38	1.70	1.12	1.25	0.27	0.42	0.68	0.61
					[< 0.01]	[< 0.01]	[0.09]	[0.08]
Financial situation, next 12 months	0.76	1.09	0.86	1.04	0.54	0.60	0.94	0.83
	[0.28]		[0.27]		[< 0.01]	[0.01]	[0.39]	[0.18]
General economic situation, last 12 months	1.26	1.10	1.12	1.11	1.16	0.83	2.16	2.03
						[0.17]		
General economic situation, next 12 months	1.35	1.03	1.20	1.31	1.02	0.98	1.35	1.25
						[0.44]		
Major purchases planned, next 12 months	0.91	0.91	1.00	1.08	0.53	0.81	0.92	0.91
	[0.44]	[0.43]			[< 0.01]	[0.06]	[0.41]	[0.39]
Major purchases intended, currently	0.75	0.57	0.86	0.92	0.70	0.60	1.16	1.20
	[0.25]	[0.13]	[0.01]	[0.17]	[< 0.01]	[< 0.01]		
Price trend assessment, last 12 months	0.80	0.60	0.82	0.88	0.80	0.65	1.62	1.62
	[0.12]	[0.07]	[0.03]	[0.21]	[0.11]	[< 0.01]		
Price trend expectations, next 12 months	1.08	1.23	1.00	1.33	0.99	0.88	1.27	1.38
					[0.47]	[0.01]		
Savings planned, next 12 months	1.49	1.04	1.44	1.43	0.69	0.97	1.50	1.07
					[0.01]	[0.41]		
Savings intended, currently	0.81	0.76	0.81	0.90	0.78	0.50	1.13	1.27
	[0.20]	[0.19]	[0.04]	[0.22]	[0.03]	[0.01]		

Table 2 (contd.)

	Existing Houses		New Houses		Existing Ap.		New Ap.	
Predictor	*VAR*	*ARDL*	*VAR*	*ARDL*	*VAR*	*ARDL*	*VAR*	*ARDL*
AR MSFE	0.71		4.78		2.89		0.97	
Statement on financial situation of household	1.65	1.95	1.16	1.17	0.28	0.46	0.88	1.03
					[< 0.01]	[< 0.01]	[0.42]	
Consumer confidence	1.24	1.03	1.23	1.31	1.16	0.99	1.29	1.50
						[0.46]		
Unemployment expectations, next 12 months	1.56	1.27	1.23	1.26	1.17	1.06	1.76	2.25
Intention to buy a car within the next 12 months	2.71	2.84	1.58	1.66	1.26	1.20	1.28	1.41
Purchase or build a home within the next 12 months	1.23	1.11	1.99	1.87	1.11	1.01	1.04	1.22
Home improvements over the next 12 months	0.51	0.49	1.20	1.25	1.37	1.19	1.24	1.35
	[0.06]	[0.09]						
Unemployment rate	1.94	1.66	1.14	1.25	1.14	1.44	1.66	1.59
Consumer price; Housing rent (net)	4.84	2.65	1.08	1.24	0.77	0.70	1.42	1.17
					[0.32]	[0.28]		
Consumer price	2.79	2.13	1.38	1.51	1.68	1.72	0.96	1.09
							[0.45]	
CDAX share price index	0.84	0.90	1.32	1.33	1.24	0.93	1.17	1.28
	[0.04]	[0.04]				[0.21]		
Interest rate of new mortgages	2.22	2.18	0.71	0.64	0.44	0.36	0.92	1.15
			[0.06]	[0.04]	[0.06]	[0.02]	[0.45]	
New orders; Construction	1.04	1.15	1.18	1.32	1.30	0.96	0.96	1.04
						[0.30]	[0.39]	
Building permits	1.08	1.37	1.06	1.35	0.93	0.79	0.99	1.08
					[0.34]	[0.11]	[0.49]	

Note: The entries in the first row report the MSFE for the autoregressive forecast model. The other rows give the ratio of the MSFE for the respective individual ARDL model using the predictor mentioned in the first column to the MSFE for the autoregressive forecast model. p-values of the Diebold-Mariano test are reported in squared brackets.

Table 3: Forecast Combination Results

	Existing Houses		New Houses		Existing Ap.		New Ap.	
Predictor	*VAR*	*ARDL*	*VAR*	*ARDL*	*VAR*	*ARDL*	*VAR*	*ARDL*
AR MSFE	0.71		4.78		2.89		0.97	
Mean	0.90	0.83	1.01	1.20	0.69	0.65	0.94	1.01
	[0.36]	[0.30]			[< 0.01]	[< 0.01]	[0.33]	
Median	0.92	0.81	1.10	1.24	0.82	0.75	0.92	0.99
	[0.39]	[0.26]			[< 0.01]	[< 0.01]	[0.27]	[0.47]
Trimmed Mean, 90%	0.89	0.83	1.10	1.20	0.72	0.67	0.94	1.00
	[0.36]	[0.30]			[< 0.01]	[< 0.01]	[0.33]	
Weighted MSFE, last period	0.53	0.42	0.83	0.96	0.35	0.32	0.78	0.94
	[0.08]	[0.07]	[0.09]	[0.35]	[< 0.01]	[0.01]	[0.22]	[0.43]
Weighted MSFE, entire history	0.75	0.63	1.41	1.36	0.37	0.25	1.21	1.38
	[0.19]	[0.17]			[< 0.01]	[< 0.01]		
Cluster(3)	0.45	0.40	0.75	0.82	0.34	0.41	0.89	0.91
	[0.07]	[0.06]	[0.04]	[0.10]	[< 0.01]	[< 0.01]	[0.39]	[0.39]

Note: The entries report the ratio of the MSFE for the respective combination forecast model given in the first column to the MSFE for the autoregressive forecast model. p-values of the Diebold-Mariano test are reported in squared brackets.