

# House Price Forecasting from Investment Perspectives

Song Shi <sup>1,\*</sup>, Vince Mangioni <sup>1</sup>, Xin Janet Ge <sup>1</sup>, Shanaka Herath <sup>1</sup>, Fethi Rabhi <sup>2</sup> and Rachida Ouyssse <sup>3</sup>

<sup>1</sup> School of the Built Environment, University of Technology Sydney, Sydney, NSW 2007, Australia; Vincent.Mangioni@uts.edu.au (V.M.); Xinjanet.ge@uts.edu.au (X.J.G.); Shanaka.herath@uts.edu.au (S.H.)

<sup>2</sup> School of Computer Science and Engineering, University of New South Wales, Sydney, NSW 2052, Australia; f.rabhi@unsw.edu.au

<sup>3</sup> School of Economics, University of New South Wales, Sydney, NSW 2052, Australia; rouysse@unsw.edu.au

\* Correspondence: song.shi@uts.edu.au

**Abstract:** Housing market dynamics have primarily shifted from consumption- to investment-driven in many countries, including Australia. Building on investment theory, we investigated market dynamics by placing investment demand at the center using the error correction model (ECM). We found that house prices, rents, and interest rates are cointegrated in the long run under the present value investment framework. Other economic factors such as population growth, unemployment, migration, construction activities, and bank lending were also important determinants of the housing market dynamics. Our forecasting results show that the Sydney housing market will continue to grow with no significant price decline in the foreseeable future.

**Keywords:** housing market dynamics; price forecast; error correction model (ECM); Sydney housing market

**Citation:** Shi, S.; Mangioni, V.; Ge, X.J.; Herath, S.; Rabhi, F.; Ouyssse, R. House Price Forecasting from Investment Perspectives. *Land* **2021**, *10*, 1009. <https://doi.org/10.3390/land10101009>

Academic Editor: Sarel Cilliers

Received: 20 August 2021

Accepted: 23 September 2021

Published: 26 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

A prosperous and healthy housing market is critical for the modern economy. Yet, there are crucial research gaps, and, therefore, there is a need to better understand housing market dynamics. One of the most critical observations we made in past decades is that housing market dynamics have primarily shifted from consumption- to investment-driven in many countries, including Australia. Investment activities (including speculation on capital gains) have been the center of housing demand. More importantly, this investment-driven housing demand is believed to continue in the foreseeable future. With firm guidance from investment theory, this paper examines housing market dynamics through a lens that places investment demand at the center of housing market dynamics using an augmented error correction model (ECM). Since the proposed ECM model contains information set on both current and past values of house prices, the results are useful in a one-step-ahead price forecast which is important in real estate markets.

The ECM method has been widely applied to house price forecasting in many countries such as the US [1], Netherlands [2], Euro areas [3], New Zealand [4], and Australia [5]. As pointed out by Wheaton, Chervachidze, and Nechayev [6], the method can handle both the stationarity and endogeneity problems that plague the time-series data for local housing price analysis. However, the empirical application of an ECM requires researchers to select possible “determinants” of the local housing market and distinguish their long- or short-term relationship in the market. Traditional reduced-form studies of housing market supply and demand struggle to provide firm guidance from theory on the selection of housing determinants, which means that the ECM specifications are often ad hoc in practice. The novelty of our approach is to specify an investment ECM model based on the long-run relationship among house prices, rents, and interest rates and a variety of short-term impacts from other economic variables. To the best of our knowledge, this is

the first paper distinguishing between the investment and consumption determinants in an ECM specification of house price analysis.

We found that the key market drivers for housing demand are prices, rents, and interest rates. There is a long-run relationship between housing prices, rents, and interest rates. House prices are also influenced by other local and nationwide economic variables such as population growth, unemployment rates, migration numbers, construction costs, bank mortgage lending, and spatial correlation among Australian capital cities. These economic variables directly affect house prices in the short-term and indirectly affect house prices through rents and interest rates in the long term.

Housing is both a consumption good and an investment asset. For many Australians, the great Australian dream is to own their family homes, leading to a better, more secure, and prosperous life. Nowadays, the rising housing prices in many Australian regions, especially in large capital cities such as Sydney and Melbourne, are making it increasingly difficult for many to achieve this dream. Over the last 40 years, housing prices in Sydney have increased more than thirty times; in contrast, household incomes have merely increased by about ten times. Although the Sydney population has almost doubled during the same period and household formations are becoming smaller, these do not fully explain why Australian house prices have priced out so many hard-working ordinary Australians. Yanotti and Wright [7] studied the determining factors of purchasing investment properties through analyzing mortgage finance data and identified where the investment stock is physically located. They found a mismatched investment of residential property and rental demand in Australia.

Traditional house price models derived from reduced-form structural supply and demand equations are valid but have a misplaced investment demand. In 1989, Mankiw and Weil [8] published their famous paper titled "The Baby Boom, The Baby Bust, and the Housing Market". By examining significant demographic changes in the US housing market, they concluded that real housing prices would fall substantially in the future. Thirty years later, global house prices, including in the US and Australia, have increased considerably rather than decreased. Why do their conclusions, based on demographic changes, have proved to be inaccurate? The answer largely depends on how we understand and interpret housing market dynamics.

In housing finance, making a deposit and then repaying the mortgage principal over 20 to 30 years represents an opportunity cost of ownership that is valued by the present value of future benefits. Shiller [9] argued that house prices should be equal to the current discounted value of future rents. This present value framework borrowed from finance literature applies to both investment and consumption purposes. For owner-occupied houses, there is an implicit rent called a user cost of housing paid by the owner [10–16]. Since the accurate estimation of a user cost can be difficult or even problematic [17], researchers tend to use market rents as a proxy for user cost in practice. Compared to the estimation of user cost, market rents are more readily observable in a marketplace. Shi, Jou, and Tripe [4] applied the present value model to investigate how the real rental rate and interest rate affect the real housing price. Using a similar asset pricing approach, Bourassa, Hoesli, and Oikarinen [18] found that a price-to-rent ratio measure performs well in housing market dynamics analysis.

To test the validity of our proposed investment modeling, we undertook a cointegration and error correction analysis for fifteen years of house prices in the Greater Sydney Area between 2004 and 2018. We first used data from 2004 to 2013 to develop the model and then tested the model's forecast performance in an out-of-sample test during 2014 and 2018. The results showed an average forecast error of 2.2%, or between AUD 19,000 and AUD 23,000, compared to the average house price in the Greater Sydney Area of AUD 863,000 during the testing period. Moreover, the model showed a superior forecast performance compared to the alternative models. We also forecasted house prices in the Greater Sydney Area for the foreseeable future, that is, from 2019 to 2030. The results indicated that house prices should bottom out in 2019 and would continue to rise in the

future. The chance for a significant downward market price adjustment is less likely. Note that our price forecasting was based on the data up to 2018, i.e., the best available information at the time of writing this paper. Although we cannot forecast future events like a pandemic, their economic impacts on the housing market are easy to incorporate in our models via the changes of interest rates and other economic variables in the specified dynamic system.

In the next section, we review the theoretical base for the modeling. We then outline the econometric tools used, describe the data, discuss the results, and provide the conclusions.

## 2. Theoretical Base for Modeling

The theoretical basis follows the present-value model, which is commonly used to estimate the value of stocks in the share market to estimate the values for housing. Under the proposed present-value theory, housing is valued as a financial asset where price relates to the expected future cash flows discounted to the present by using an expected discount rate. For owner-occupied housing, there is an implicit rent (called a user cost or economic rent) paid by the homeowners; for investment properties, there is market rent paid by the renters. Thus, housing price  $P$  at time  $t$  can be written as follows:

$$P_t = E_t \left[ \sum_{i=1}^n \frac{D_{t+i}}{(1+R)^i} \right] + E_t \left[ \frac{P_{t+n}}{(1+R)^n} \right] \quad (1)$$

where  $D_t$  is the dividend or cash flow at time  $t$  and  $R$  is the discount rate.  $E$  denotes the expected value.

In the finance literature, the first term is often called the fundamental value, and the second term is the price bubble. When  $n$  is sufficiently large, the second term converges to zero. The model implies that the current asset price is simply the sum of all expected present values of future cash flows, discounted at a constant rate. The model has been widely applied in finance to value shares.

In the property market, a simplified version of the above present model, called a capitalization approach, is used to value income-producing properties. However, while the assumption of a constant expected income  $D$  and discount rate  $R$  is analytically convenient, they contradict the evidence that both the expected income and the investor's expected rate of return vary over time. Campbell and Shiller [19,20] suggested a log-linear present-value model with time-varying expected returns, where a log asset price at time  $t$  is written as follows:

$$p_t = \frac{k}{1-\rho} + E_t \left[ \sum_{j=0}^n \rho^j \left[ (1-\rho) d_{t+1+j} - r_{t+1+j} \right] \right] + E_t \left[ \rho^j p_{t+j} \right] \quad (2)$$

where  $\rho = 1 / (1 + \exp(\overline{d-p}))$ ,  $\overline{d-p}$  is the average log dividend-price ratio,  $k = -\log(\rho) - (1-\rho) \log\left(\frac{1}{\rho} - 1\right)$ , and  $r_{t+1} = \log(P_{t+1} + D_{t+1}) - \log(P_t)$ .

When the time horizon  $n$  increases to infinity, the third term, which is the discounted expected value of the asset price, shrinks to zero. Accordingly, the current asset price can be presented as follows:

$$p_t = \frac{k}{1-\rho} + E_t \left[ \sum_{j=0}^{\infty} \rho^j \left[ (1-\rho) d_{t+1+j} - r_{t+1+j} \right] \right] \quad (3)$$

This equation can be rewritten in terms of the log dividend-price ratio, which is:

$$d_t - p_t = -\frac{k}{1-\rho} + E_t \left[ \sum_{j=0}^{\infty} \rho^j \left[ \Delta d_{t+1+j} + r_{t+1+j} \right] \right] \quad (4)$$

The above equation (4) is referred to by Campbell and Shiller [19,20] as the dividend-ratio model. This equation implies that the log dividend-price ratio should be stationary, provided the changes in log dividends and the expected stock return are stationary. Where the log dividend-price ratio is nonstationary, it is very likely that the expected stock return is nonstationary (highly persistent), even when the above present-value model is valid.

### 3. Econometric Tools for Model Development

The Augmented Dickey–Fuller (ADF) unit root test is applied to test for the stationarity of time-series variables. As expected, they are integrated order one I(1) processes.<sup>1</sup> Because house prices, rents, and interest rates are I(1) variables, they are further tested for cointegration under the proposed present-value framework.

Cointegration tests are widely used in time-series econometrics. In empirical tests, cointegration and unit root tests between stock prices and dividends give mixed findings depending on the time period studied. Using the annual US stock market data from 1871 to 1986, Campbell and Shiller [21] found that stock prices and dividends are not cointegrated. The deviation between prices and dividends is quite persistent. On the other hand, Diba and Grossman [22] indicated a possible cointegration relationship between stock prices and dividends for the US stock market. In terms of the housing market, Gallin [23] found the log rent-price ratio is stationary by using aggregated quarterly data for the US housing market. When Brooks, Katsaris, McGough, and Tsolacos [24] examined the monthly prices of UK equity-traded property stocks from 1986 to 1998, they found that prices and rents are not cointegrated over the sample period.

The error correction model (ECM) is the next application step in this study under the proposed present-value framework. The ECM model is based on the assumptions that house prices, rents, and interest rates are cointegrated and that house prices are affected by both long-run cointegration and short-run dynamics among prices, rents, and interest rates. Since house prices are affected by other factors and tend to be seasonal, key micro- and macro-economic variables and seasonal quarterly dummy variables are included as external control variables in the ECM model. The identified key economic variables include population and immigration growth, building costs, building supply, the amount of mortgage lending, and housing price growth in Australian capital cities. They are selected for the proposed ECM model according to their statistical significance in the modeling. To be specific, the Johansen cointegration test and the ECM model were written as follows:

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + \beta_1 y_{t-1} + \gamma_1 z_{t-1} + \delta_1 \mu_{t-1} + E_t + S_t + \varepsilon_t \quad (5)$$

where  $x$ ,  $y$ , and  $z$  are log price, rent, and interest rate in the first differences,  $\alpha_0$  is a constant,  $\mu$  is the error correction term,  $E$  is a vector of key micro and macroeconomic variables such as population, employment, immigration, bank lending, construction activities, etc. (see Table 1 for exact variables used in Equation (5)),  $S$  is a vector for seasonal dummies, and  $\varepsilon$  is white noise.

**Table 1.** Definition of variables.

Variable	Definition	Sources
SYD_HP_MEDIAN_H	Median house prices in the Greater Sydney Area	SIRCA
SYD_RENT_H	Median house rents in the Greater Sydney Area	SIRCA
SYD_HP_MEDIAN_U	Median unit prices in the Greater Sydney Area	SIRCA
SYD_RENT_U	Median unit rents in the Greater Sydney Area	SIRCA
AUS_INT_90D	Australia 90-day bill rate	Datastream
AUS_INT_10YRF	Australian ten-year government bond yield	Datastream

AUS_LEN_DF	Bank mortgage lending for all dwelling finance in Australia	ABS
AUS_CONF	Building construction cost index of Australia	ABS
NSW_PPGF	New South Wales population growth	ABS
SYD_SUP_DF	Total number of dwelling (all types) supply in the Greater Sydney Area	NSW Planning & Environment
SYD_NOMF	Net migration numbers in the Greater Sydney Area	ABS
ECC_PCF	Residential Property Price Index percentage change from the corresponding quarter of the previous year-weighted average of eight capital cities	ABS
Variable Transformation		
D(LOG(SYD_HP_MEDI AN_H))	Changes in median house prices	
D(LOG(SYD_RENT_H))	Changes in median house rents	
D(LOG(SYD_HP_MEDI AN_U))	Changes in median unit prices	
D(LOG(SYD_RENT_U))	Changes in median unit rents	
D(LOG(AUS_LEN_DF))	Growth rate of bank mortgage lending	
D(LOG(AUS_CONF))	Changes in building construction costs	
D(LOG(NSW_PPGF))	Changes in population growth	
D(LOG(SYD_NOMF))	Changes in net migration numbers	

Note: D denotes the first difference and LOG represents the natural logarithm.

For simplicity, we excluded any potential structure breaks or regime shifts in the above models. This is justified as our main objective is forecasting rather than econometric modeling. The literature showed that forecasts are not substantially affected by the presence of structural breaks [25]. Elliott and Muller [26,27] further showed that gains from modeling a structural break might be offset by imprecisely estimated break dates and post-break parameters. Therefore, ignoring a break may lead to more accurate forecasts [28].

#### 4. Variable Selection and Description

This analysis focused on the Greater Sydney Area, which covers 35 LGAs as defined by the Australian Bureau of Statistics (ABS).<sup>2</sup> The sample period spanned from Q3 2004 to Q4 2018. The reason the analysis started in 2004 is mainly due to data availability. For example, there is no reliable rental time-series data prior to 2004. All time-series data are in quarterly frequencies and collected from various data sources. In total, 50 variables were considered in the proposed ECM model, covering a broad range of national, regional, and local economies. Some key variables and data sources are listed below and detailed in Table 1.<sup>3</sup>

- (1) Australian Bureau of Statistics (ABS). This includes, for example, time-series data on regional population and immigration growth, nationwide bank mortgage lending, building construction cost, and price changes in other capital cities
- (2) Securities Industry Research Center of Asia-Pacific (SIRCA). This includes, for example, time-series data on median prices and rents for houses and units in the Greater Sydney Area. We assume that the mix and quality of quarterly residential property sales for houses and units (i.e., the number of bedrooms, bathrooms, car parks, lot size, land tenure, a mix of high- and low-value properties, and transaction types, etc.) are relatively stable over time. The impact of forced sales on reported

median price indices as discussed in Renigier-Biłozor, Walacik, Żróbek, and d'Amato [29] is small in this study.

- (3) DataStream. This includes, for example, market data on Australia's interest rates and government bond yields
- (4) NSW Planning & Environment. This includes, for example, time-series data related to local housing and land supply.

All the variables and forecasting results are expressed here in nominal terms, i.e., they are the actual value at the time of the analysis without adjusting for inflation. Depending on the purpose of their research, economists choose to use either real or nominal terms in their analyses. However, there is no consensus in the literature on what inflation indices should be used to deflate the variables. To avoid introducing unnecessary measurement errors in deflating variables for the analysis and later converting prices back to nominal values for the forecasting, nominal values are adopted in this market dynamic analysis.

## 5. Model Development and Testing Results

### 5.1. Model Development

Table 2 shows the ECM modeling results for the Greater Sydney Area between 2004 and 2018. Column (1) contains the results for houses, and column (2) is for units. The adjusted R-squared is 0.843 for houses, which means that a variation of 84.3% in quarterly house price changes has been explained by the ECM house price model. For units, the adjusted R-squared is 0.812. Thus, both house and unit price changes are well modeled over the sample period.

Under the present-value framework, asset prices, rents, and interest rates are expected to be cointegrated in the long run. The cointegration results in Table 2 show prices, rents, and interest rates are indeed cointegrated.<sup>4</sup> Both house and unit price changes are positively related to rent changes but negatively related to interest rate changes. For every one percent increase in rents, it increases house prices by about 0.262%, while for every one percent increase in interest rates, it decreases house prices by 0.582%. In contrast, unit prices are more sensitive to rent and less sensitive to interest rate changes. For every one percent increase in rents, it increases unit prices by 0.5%, while for every one percent increase in interest rates, it decreases unit prices by about 0.412%. One possible explanation for this is that houses are more expensive than units. Thus, house owners are more sensitive to interest rate changes than unit owners. Another explanation is that units are more likely to be held for investment purposes. Investors focus more on income instead of interest rate changes because interest costs are tax-deductible for property investment.

The long-term relationship, as indicated by the cointegration term *CointEq1*, is negative (−0.048 for houses and −0.046 for units) and statistically significant, which means the speed of adjustment from a short-run towards a long-run equilibrium in the Sydney housing market is about 4.6–4.8 percent each quarter. The results indicate a very slow price adjustment process in the housing market.

Short-term price dynamics also affect price changes. Table 2 shows that the current period price change is statistically significant in relation to the last period price change and the price change in other Australian capital cities (apart from Sydney). One percentage point increase in the last period price change increases the current period price change by 0.536% and 0.639% for houses and units, respectively. Although other economic factors are not statistically significant in the price model, they are important in rent and interest rate models. For example, rents are found to be negatively related to the last period's net migration figure in New South Wales and the new dwelling supply in Sydney. Interest rates, however, are found to relate positively to the last period's government bond yields, construction costs, net migration figure, and dwelling supply. Interest rates are also negatively related to Australian mortgage lending.<sup>5</sup>

Overall, both house and unit price equations are well supported by the ECM modeling under the present-value framework. Our results show that prices are mostly driven by the short-run dynamics in the market, and the price adjustment process to its long-run equilibrium is very slow. It takes about 20 quarters for the housing market to adjust towards its long-run equilibrium. The results imply a self-fulfilling phenomenon in the Sydney housing market, which is consistent with the bubble literature on the Australian and New Zealand housing markets [5,30,31].

**Table 2.** Model development—Error correction model estimates, Q3 2004–Q4 2018.

	(1)	(2)
	House Price Changes	Unit Price Changes
CointEq1	−0.048 *** (0.014)	−0.046 *** (0.015)
D(LOG(SYD_HP_MEDIAN(-1)))	0.536 *** (0.071)	0.639 *** (0.076)
D(LOG(SYD_RENT(-1)))	0.009 (0.107)	−0.007 (0.054)
D(LOG(AUS_INT_90D(-1)))	0.017 (0.021)	−0.007 (0.014)
Constant	−0.002 (0.003)	−0.000 (0.003)
D(LOG(AUS_INT_10YRF))	−0.003 (0.015)	−0.001 (0.011)
D(LOG(AUS_LEN_DF))	0.007 (0.034)	0.011 (0.025)
D(LOG(AUS_CONF))	−0.291 (0.269)	−0.210 (0.215)
D(LOG(NSW_PPGF))		−0.003 (0.004)
D(LOG(SYD_SUP_DF))	0.003 (0.008)	−0.006 (0.007)
D(LOG(SYD_NOMF))	0.042 (0.036)	0.052 * (0.028)
ECC_PCF	0.002 *** (0.000)	0.001 *** (0.000)
Seasonal dummy	Yes	Yes
Adj. R-squared	0.843	0.812
Sum sq. resids	0.003	0.006
S.E. equation	0.009	0.012
F-statistic	24.567	2.966
Log likelihood	200.977	184.466
Akaike AIC	−6.447	−5.878
Schwarz SC	−5.95	−5.38
Mean dependent	0.012	0.008
S.D. dependent	0.022	0.014
The cointegration equation results		
LOG(SYD_HP_MEDIAN_H(-1))	1.000	1.000
LOG(SYD_RENT_H(-1))	−0.262 (0.249)	−0.500 (0.117)
LOG(AUS_INT_90D(-1))	0.582 (0.077) ***	0.412 (0.043) ***
Constant	−12.475	−10.609

Notes: Descriptions of the explanatory variables are provided in Table 1. Standard errors are shown in ( ). \*, \*\*, and \*\*\* denote significance levels at the 10%, 5%, and 1% level, respectively.

### 5.2. Out-of-Sample Testing, Underlying Variable Assumptions, and Measures of Forecasting Accuracy

To validate the proposed ECM model forecasting performance, we used an in-sample (training) data period from 2004 to 2013 to estimate the model and a pseudo-out of-sample period between 2014 and 2018 to test its forecasting performance. In other words, we used the first 10 years of data to build the model and compared its forecasted prices to the actual prices in the next five years from 2014 to 2018. The ECM models developed for the periods between 2004 and 2013 are included in Appendix A.

Since other economic variables are included as exogenous inputs in the proposed ECM model, we needed to make some assumptions about those variables in the out-of-sample testing period before forecasting prices. One approach was to take them as observable. In this case, forecasting errors between the forecasted and actual prices would have to be due to the ECM model itself rather than any estimation errors introduced from the other forecasted variables. However, this take-as-given approach is not realistic. In reality, we simply do not know those economic variables in advance. Thus, we introduced a conditional forecasting method of using time-series techniques to forecast those underlying economic variables. These advanced time-series forecasting techniques include auto-regressive integrated moving average (ARIMA) and the auto-regressive (AR) methods. Four statistical measures were used to evaluate a forecast performance and the statistical difference between the forecasted prices and actual prices. They are the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Theil inequality coefficient. While the first three statistics are based on quadratic loss and average errors, the Theil index measures the proximity between the forecasts examining the ability of the forecast series to match the mean and the variance of the actual series. The model forecast evaluation results are shown in Table 3.

**Table 3.** Out-of-sample testing, Q1 2014–Q4 2018.

Variable	obs.	RMSE	MAE	MAPE	Theil
Panel A: Actual					
Median house prices	20	22,578	19,003	2.207	0.013
Median unit prices	20	31,923	23,047	3.141	0.024
Panel B: Statistical models					
ARIMA model					
Median house prices	20	53,771	45,449	5.368	0.032
Median unit prices	20	24,593	20,107	2.299	0.014
AR(1) model					
Median house prices	20	59,000	49,203	5.776	0.035
Median unit prices	20	18,484	14,111	2.068	0.014
AR(4) model					
Median house prices	20	70,685	59,990	7.172	0.042
Median unit prices	20	21,640	18,154	2.729	0.016
AR(8) model					
Median house prices	20	64,694	54,936	6.534	0.038
Median unit prices	20	21,370	18,047	2.688	0.016
AR(16) model					
Median house prices	20	46,217	39,780	4.712	0.027
Median unit prices	20	25,551	19,379	2.836	0.019
AR(20) model					
Median house prices	20	37,305	31,007	3.677	0.022
Median unit prices	20	30,753	23,981	3.397	0.023

Notes: RMSE denotes the Root Mean Square Error; MAE is for the Mean Absolute Error; MAPE is for the Mean Absolute Percentage Error; and Theil is for the Theil Inequality Coefficient.



Panel A of Table 3 shows the model forecast evaluation results for houses when underlying variables are taken as given. Between the forecasted and actual quarterly house prices, the RMSE is AUD 22,578, the MAE is AUD 19,003, the MAPE is 2.207%, and the Theil for inequity measure is 0.013. The average house price in the Greater Sydney Area is about AUD 863,000 during the forecasting period between 2014 and 2018. Thus, the forecasted quarterly house prices closely track the actual house prices, which is demonstrated in Figure 1.

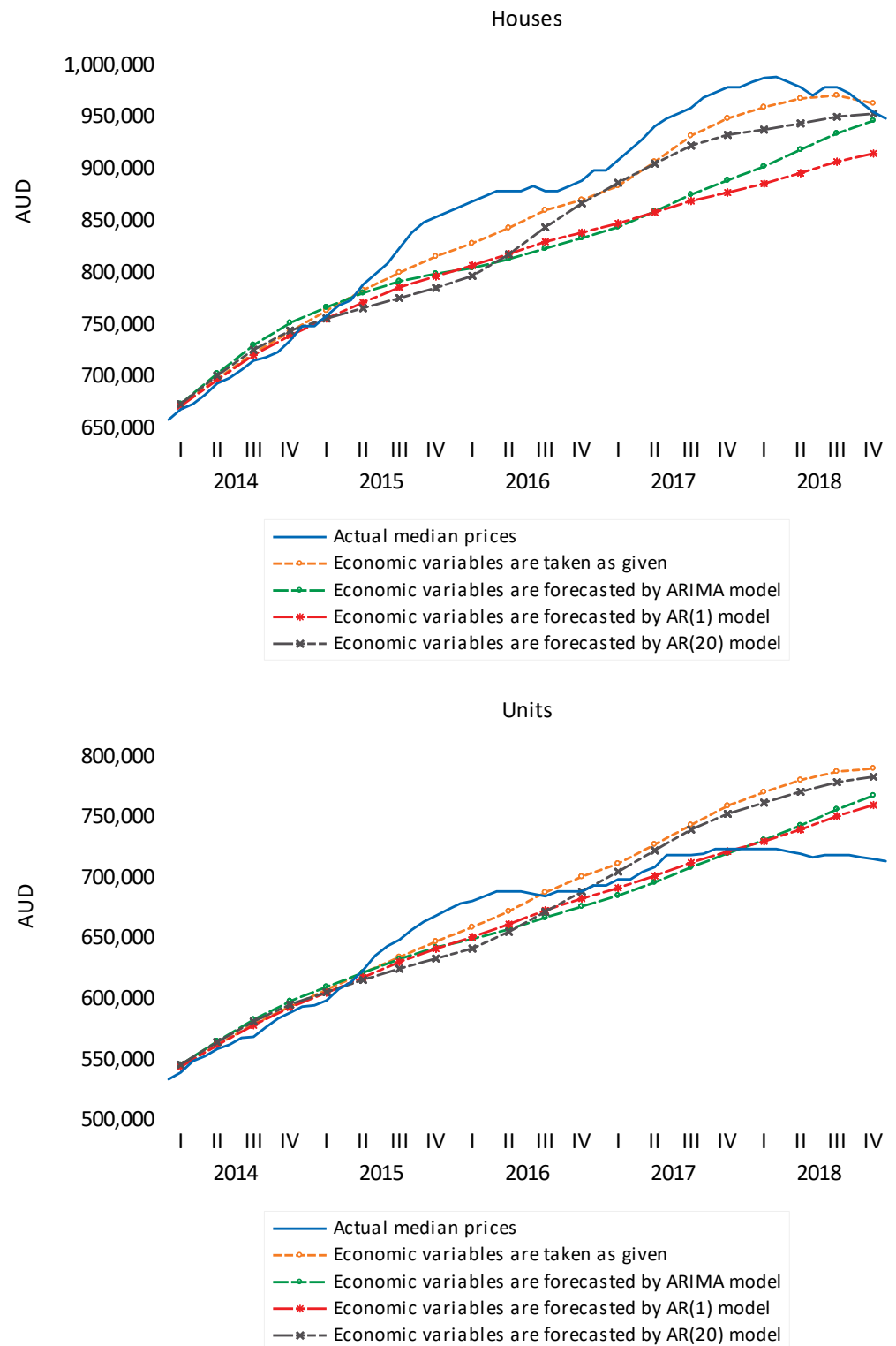


Figure 1. Out-of-Sample forecast, Q1 2014–Q4 2018.

Panel B of Table 3 shows the results of conditional forecast when the underlying economic variables are unknown and must be forecasted. Two popular and widely-applied time-series forecasting techniques (ARIMA and AR models) were considered in this analysis. The results show that using an AR(1) process in the underlying economic variable forecast produced the best forecasting results for units, while an AR(20) process produced the best forecasting results for houses. The ARIMA model produced somewhat forecasting results in between an AR(1) and AR(20) process. The results showed that the forecasting performance of developed models is conditional on the assumed knowledge of underlying variables. As AR(1) and AR(20) processes tend to set the boundary of forecasting, we chose to rely on the more sophisticated ARIMA model and selected AR(1) and AR(20) processes for future price forecasting in the next section. The results provide some insights into forecasting for decision-making under different underlying variable assumptions.

### 5.3. Forecasting Results

Table 4 presents the forecast median house and unit prices for the Greater Sydney Area from 2019 to 2030. These prices are forecast-based on the model developed using the data between 2004 and 2018 described in Section 4. The underlying economic variables are estimated using the ARIMA and selected AR models. In other words, these are fitted values to each one of the macroeconomic variables,  $E_t$ .

**Table 4.** Forecasted median prices, Q1 2019–Q4 2028.

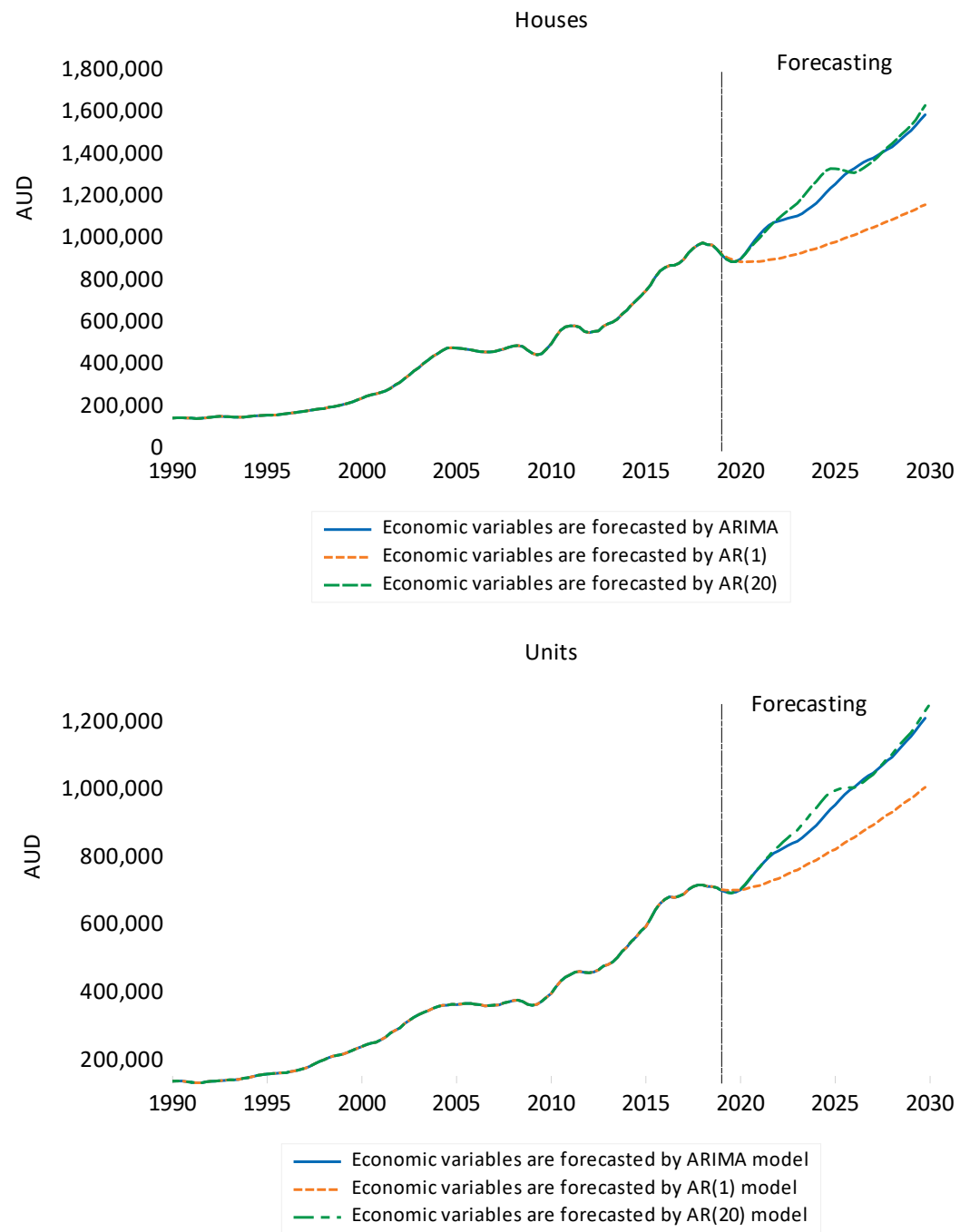
Time Period	Houses			Units		
	ARIMA	AR(1)	AR(20)	ARIMA	AR(1)	AR(20)
2019Q1	931,067	935,696	930,709	708,278	711,058	707,385
2019Q2	911,194	919,994	910,228	703,633	709,259	702,693
2019Q3	899,668	910,522	898,726	701,262	709,549	701,167
2019Q4	899,592	902,946	898,575	704,042	709,965	705,222
2020Q1	911,688	897,870	910,480	710,599	709,785	712,396
2020Q2	937,604	896,699	935,054	725,269	712,940	727,899
2020Q3	969,586	898,580	963,080	742,813	717,425	744,982
2020Q4	1,000,454	899,203	988,440	761,172	721,061	762,201
2021Q1	1,026,564	899,777	1,010,434	776,647	723,352	776,808
2021Q2	1,051,763	902,546	1,035,673	793,422	728,505	794,303
2021Q3	1,072,154	907,377	1,061,955	807,664	734,725	810,899
2021Q4	1,084,812	910,416	1,085,048	818,901	739,903	827,012
2022Q1	1,090,764	913,157	1,103,916	825,515	743,606	838,728
2022Q2	1,097,922	918,003	1,122,849	833,694	750,172	852,274
2022Q3	1,104,906	924,863	1,141,454	841,290	757,768	864,616
2022Q4	1,110,291	929,794	1,158,450	848,505	764,205	877,111
2023Q1	1,115,409	934,287	1,176,081	854,240	769,020	887,132
2023Q2	1,127,265	940,781	1,201,009	864,403	776,693	902,642
2023Q3	1,143,098	949,182	1,229,239	876,219	785,338	919,601
2023Q4	1,159,630	955,449	1,255,977	888,902	792,691	937,265
2024Q1	1,176,506	961,117	1,280,240	900,605	798,278	951,949
2024Q2	1,199,588	968,711	1,307,132	916,464	806,760	970,516
2024Q3	1,224,766	978,157	1,329,884	933,008	816,191	986,871
2024Q4	1,247,488	985,306	1,341,254	948,795	824,226	999,107
2025Q1	1,267,094	991,752	1,340,653	961,684	830,379	1,003,737
2025Q2	1,289,900	1,000,113	1,336,449	977,148	839,503	1,009,010
2025Q3	1,311,892	1,010,324	1,330,797	991,885	849,581	1,011,236
2025Q4	1,328,916	1,018,109	1,324,326	1,004,629	858,177	1,012,546

2026Q1	1,341,300	1,025,120	1,321,241	1,013,773	864,786	1,012,437
2026Q2	1,356,635	1,034,069	1,329,277	1,025,616	874,466	1,019,592
2026Q3	1,371,711	1,044,895	1,343,665	1,037,221	885,119	1,029,243
2026Q4	1,382,851	1,053,181	1,359,345	1,047,595	894,211	1,040,725
2027Q1	1,390,980	1,060,639	1,374,828	1,055,374	901,217	1,050,439
2027Q2	1,404,237	1,070,076	1,396,359	1,067,094	911,409	1,066,013
2027Q3	1,419,395	1,081,436	1,420,529	1,079,755	922,604	1,082,358
2027Q4	1,432,427	1,090,149	1,441,346	1,092,156	932,161	1,098,464
2028Q1	1,444,141	1,097,988	1,459,068	1,102,630	939,534	1,110,928
2028Q2	1,462,484	1,107,862	1,480,598	1,117,673	950,221	1,128,060
2028Q3	1,483,687	1,119,714	1,503,627	1,133,929	961,946	1,144,919
2028Q4	1,503,083	1,128,815	1,524,299	1,149,803	971,957	1,161,090
2029Q1	1,521,148	1,137,001	1,544,480	1,163,369	979,686	1,174,553
2029Q2	1,545,758	1,147,287	1,572,678	1,181,283	990,865	1,194,720
2029Q3	1,572,689	1,159,614	1,606,490	1,199,862	1,003,123	1,216,369
2029Q4	1,596,568	1,169,086	1,640,703	1,217,315	1,013,591	1,238,525

Notes: the prices are forecasted based on various statistical assumptions of exogenous variables in the error correction model (ECM). Prices are in nominal terms.

For both houses and units, the best-predicted price pattern was conditional on a high order serial correlation in the underlying variable assumptions such as an AR(20) process, while the worst predicted price path was found if the underlying economic variables follow a low order serial correlation such as an AR(1) process. The results from the ARIMA variable assumptions were in between. Figure 2 shows the future price movements based on different underlying variable assumptions.

It shows that housing prices are more likely to follow a high-growth scenario, as indicated by the ARIMA underlying variable assumptions; the market will bottom out in 2019–2020 and continue to grow in the future. Under this scenario, the median price will reach AUD 1,596,568 for houses and AUD 1,217,315 for units in 2030, presenting an annual compounding growth rate of 5%. This forecast is supported by the alternative variable assumptions following an AR(20) process. Even in a low-growth scenario, as indicated in the AR(1) variable assumptions, there will be no significant price decline in the foreseeable future; rather, prices will mostly be flat or in a slow-growth mode with an average annual growth rate of 2% in the next decade.



**Figure 2.** Forecasted Median Prices, 2019Q1–2029Q4.

#### 5.4. Robustness Check

##### 5.4.1. Unconditional Forecast

The forecasting performance of the developed ECM model depends on the forecasting of underlying economic factors. This raises some concerns about the usefulness of the proposed ECM model. Alternatively, we used the unconditional forecast of the standard AR(1) process and ECM model without other economic factors to validate the results. The advantage of using an AR(1) process for house price forecasting is that it is simple and does not require any other variables in the modeling. As the standard ECM model only depends on prices, rents, and interest rates, there is no need to forecast other economic variables in this situation. Using the data from 2004 to 2013, we developed the standard

AR(1) and ECM models and placed them in 5-year out-of-sample testing between 2014 and 2018 for performance comparison. The results are presented in Appendix B.

Our results show that the forecasting performance of a standard ECM model without underlying economic variables is less accurate than a standard AR(1) forecast, despite the ECM model having a higher adjusted R-squared (0.78) than the AR(1) process (0.62) in the model development period between 2004 and 2013. The findings were in line with the forecasting literature, stating that models, which fit the historical data well, do not necessarily perform better than others in forecasting. It is not surprising that our proposed conditional ECM model is superior to the unconditional forecast models such as the standard AR(1) and ECM models. Our findings stress the importance of including other economic variables in the ECM model development and, thereafter, forecasting.

#### 5.4.2. Forecasted Median Rents

Our proposed ECM model can also be used to predict future rents. To check the validity of our price model, we further checked whether the predicted rents were in line with market expectations. The predicted rental price movements are presented in Appendix C. Our results show that the rental markets for both houses and units will continue to creep up in the foreseeable future. For houses, median rents are forecasted to increase from AUD 550 in 2019 to about AUD 760 per week in 2030. For units, the rental prices are forecasted to increase from AUD 530 in 2019 to AUD 840 per week in 2030, all in nominal prices. Our model shows a faster rental growth for units than houses. This could be due to the demand shift in the housing market as owning becomes unaffordable for many people, especially for young people who choose to rent rather than own. As a result, the rental demands in the unit market grow at a faster pace than that of the housing market.

#### 5.4.3. Forecasted Interest Rates

Interest rates (90-day bill rates) are a key variable in the proposed ECM forecasting model. Similar to the rent, we checked what the “predicted” interest rates look like. The predicted interest rate graph is included in Appendix D. Our model forecasts a continually downward interest rate curve that falls to near zero in the future, and virtually there is no difference between rates predicted in the house and unit markets. This forecast is in line with the recent interest rate cuts and the forecast by the Reserve Bank of Australia (RBA) that interest rates could stay low ‘for a long, long time’ [32]. Therefore, our model is robust in relation to the interest rate check. Note that interest rates, rents, and house prices are endogenously determined in the ECM model, which means that we can only take robustness checks on their predicted values but cannot conduct stress tests related to the increase or decrease in these variables in the future period.

## 6. Conclusions

Housing market dynamics have shifted from traditional consumption-motivated to investment-driven in many developed countries, including Australia, over past decades. Traditional reduced-form studies of housing market supply and demand are ad hoc without firm guidance from theory and struggle to confront the challenges brought about by the dramatically increasing importance of the housing market for both policymakers and the public in their decision-making. Using sophisticated econometric tools available for modern time-series analysis and forecasting, we explored the housing market dynamics by placing the investment demand at the center guided by the present value investment theory. We found that the principal variables in the housing market dynamics (long-run relationship) are prices, rents, and interest rates. Meanwhile, other influencing variables, such as bank mortgage lending, building construction costs, population growth, dwelling supply, and net migration, affected the market in the short run via rent and interest rates. We further demonstrated that the proposed investment model has a superior forecasting

performance compared to alternative models. The results showed that the Sydney housing market is more likely to follow a high-growth scenario in the foreseeable future with an average compounding growth rate of 5% per annum.

As the interest rate is a national variable set by the Reserve Bank of Australia (RBA), an important policy implication from this research is if the State Government wants to stabilize housing prices to confront housing affordability, rent control for residential properties is a sensible way to do it. Alternative tools and policies could include increasing the building supply, restricting migration, and putting in place various kinds of purchase restriction controls. From the perspective of the Central Government, the most effective way to slow the housing market is via macro-prudential tools such as placing a loan-to-value ratio restriction on housing purchases.

Importantly, we need to point out that our forecast results are based on the time-series data up to 2018. The model parameters could be changed when new information is available. One advantage of our ECM forecast is that the model can be automated for updates in an IT platform designed to manage house price forecasts. Of course, the accuracy of our forecast depends on our understanding of the housing market dynamics and econometric tools adopted in this study.

**Author Contributions:** Conceptualization, S.S.; methodology, S.S.; software, S.S.; validation, S.S. and R.O.; formal analysis, S.S.; investigation, S.S., V.M., X.J.G., S.H., F.R. and R.O.; resources, S.S. and V.M.; data curation, S.S.; writing—original draft preparation, S.S.; writing—review and editing, S.S., V.M., X.J.G., S.H., F.R. and R.O.; project administration, S.S. and V.M.; funding acquisition, V.M., S.S., X.J.G., S.H. and F.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** NSW Landcom.

**Acknowledgments:** We thank Yang Zhang for their assistance with this research. We acknowledge Landcom, the NSW Government’s land and property development organization, for funding this project.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Error Correction Model Estimates for the Period between Q3 2004 and Q4 2013.

	(1)	(2)
	House Price Changes	Unit Price Changes
CointEq1	−0.084 *** (0.025)	−0.074 ** (0.030)
D(LOG(SYD_HP_MEDIAN(-1)))	0.577 *** (0.090)	0.674 *** (0.105)
D(LOG(SYD_RENT(-1)))	0.104 (0.148)	−0.016 (0.057)
D(LOG(AUS_INT_90D(-1)))	0.039 (0.033)	−0.008 (0.021)
Constant	−0.003 (0.005)	0.000 (0.004)
D(LOG(AUS_INT_10YRF))	−0.017 (0.028)	−0.015 (0.019)
D(LOG(AUS_LEN_DF))	−0.006 (0.043)	−0.002 (0.029)
D(LOG(AUS_CONF))	−0.143 (0.3540)	−0.073 (0.278)
D(LOG(NSW_PPGF))		−0.006

		(0.005)
D(LOG(SYD_SUP_DF))	0.011 (0.012)	-0.000 (0.009)
D(LOG(SYD_NOMF))	0.072 (0.045)	0.076 ** (0.033)
ECC_PCF	0.001 * (0.000)	0.000 (0.000)
Seasonal dummy	Yes	Yes
Adj. R-squared	0.827	0.806
Sum sq. resids	0.002	0.001
S.E. equation	0.010	0.007
F-statistic	14.650	11.954
Log likelihood	131.030	146.459
Akaike AIC	-6.16	-6.919
Schwarz SC	-5.556	-6.272
Mean dependent	0.008	0.009
S.D. dependent	0.023	0.015
The cointegration equation results		
LOG(SYD_HP_MEDIAN_H(-1))	1.000	1.000
LOG(SYD_RENT_H(-1))	-0.373 (0.202) *	-0.550 (0.079) ***
LOG(AUS_INT_90D(-1))	0.427 (0.102) ***	0.280 (0.042) ***
Constant	-11.588	-10.113

Notes: Descriptions of the explanatory variables are provided in Table 1. Standard errors are shown in (). \*, \*\*, and \*\*\* denote significance levels at the 10%, 5%, and 1% level, respectively.

## Appendix B

**Table A2.** Forecasting Performance by Alternative AR(1) and ECM Models.

Models	obs.	RMSE	MAE	MAPE	Theil
Panel A: Houses					
AR(1)	20	84,743	74,407	8.179	0.051
ECM without economic variables	20	91,834	78,256	9.626	0.055
Panel B: Units					
AR(1)	20	27,524	22,353	3.256	0.021
ECM without economic variables	20	46,490	39,064	6.140	0.036

Note: We first developed a price model based on the standard AR(1) and ECM techniques using the available data between 2004 and 2013; we then used the developed models to forecast future house and unit prices and compare them to the actual prices. The AR(1) model depends only on the lagged value of prices, while the ECM model depends only on the lagged value of prices, rents, and interest rates. No other economic variables are included in the AR(1) or ECM modeling.

Appendix C

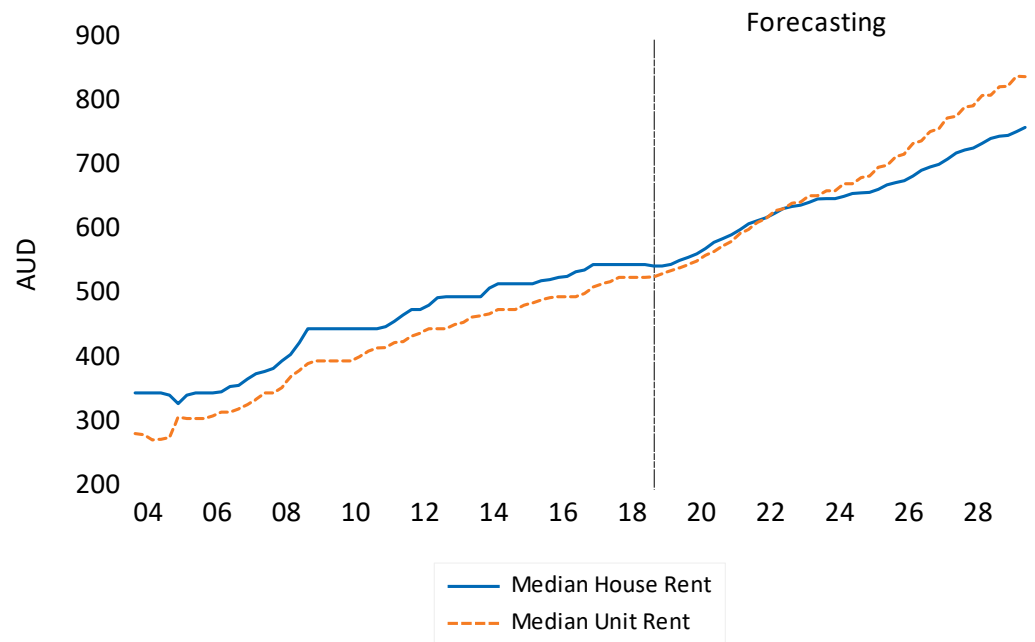


Figure A1. Forecasted Median Rents.

Appendix D



Figure A2. Forecasted Interest Rates.

Notes

- <sup>1</sup> A time series process with a unit root (a random walk).
- <sup>2</sup> The statistical area of the Greater Sydney Area is maintained by the Australian Bureau of Statistics. For 35 LGAs and their geographic locations and boundaries, please go to the Australian Bureau of Statistics website at: <https://dbr.abs.gov.au/index.html> (accessed on 25 July 2021).
- <sup>3</sup> The authors went through a wide range of data collection in this study. As not all variables are available or statistically significant in our model, we only report the key variables used in the ECM model, as shown in Table 1. Please contact the corresponding author for a complete list of variables collected in this study.
- <sup>4</sup> Results of Johansen’s cointegration test are available on request from the corresponding author.
- <sup>5</sup> Results are available on request by contacting the corresponding author.



## References

1. Malpezzi, S. A simple error correction model of house prices. *J. Hous. Econ.* **1999**, *8*, 27–62.
2. Francke, M.K.; Vujic, S.; Vos, G.A. *Evaluation of House Price Models Using an ECM Approach: The Case of the Netherlands*; Methodological Paper No. 2009-05; Ortec Finance Research Center: Rotterdam, The Netherlands, 2009.
3. Gattini, L.; Hiebert, P. *Forecasting and Assessing Euro Area House Prices through the Lens of Key Fundamentals*; Working Paper Series No. 1249; European Central Bank: Frankfurt, Germany, 2010.
4. Shi, S.; Jou, J.B.; Tripe, D. Can interest rates really control house prices? Effectiveness and implications for macro prudential policy. *J. Bank. Financ.* **2014**, *47*, 15–28.
5. Shi, S.; Valadkhani, A.; Smyth, R.; Vahid, F. Dating the timeline of house price bubbles in Australian capital cities. *Econ. Rec.* **2016**, *92*, 590–605.
6. Wheaton, W.; Chervachidze, S.; Nechayev, G. *Error Correction Models of MSA Housing “Supply” Elasticities: Implications for Price Recovery*; Working Paper Series No. 07-15; Center for Real Estate, Massachusetts Institute of Technology: Cambridge, MA, USA, 2014.
7. Yanotti, M.B.; Wright, D. Residential property in Australia: Mismatched investment and rental demand. *Hous. Stud.* **2021**, doi:10.1080/02673037.2021.1929858.
8. Mankiw, N.G.; Weil, D.N. The baby boom, the baby bust, and the housing market. *Reg. Sci. Urban Econ.* **1989**, *19*, 235–258.
9. Shiller, R. Long-term perspective on the current boom in home prices. *Econ. Voice* **2006**, *3*, 1–11.
10. Engsted, T.; Nielsen, B. Testing for rational bubbles in a coexplosive vector autoregression. *Econ. J.* **2012**, *15*, 226–254.
11. Engsted, T.; Hviid, S.J.; Pedersen, T.Q. Explosive bubbles in house prices? Evidence from the OECD countries. *J. Int. Financ. Mark. Inst. Money* **2016**, *40*, 14–25.
12. Gyourko, J.; Sinai, T. The spatial distribution of housing-related ordinary income tax benefits. *Real Estate Econ.* **2003**, *31*, 527–575.
13. Hendershott, P.H.; Slemrod, J. Taxes and the user cost of capital for owner-occupied housing. *J. Am. Real Estate Urban Econ. Assoc.* **1983**, *10*, 375–393.
14. Himmelberg, C.; Mayer, C.; Sinai, T. Assessing high house prices: Bubbles, fundamentals, and misperceptions. *J. Econ. Perspect.* **2005**, *19*, 67–92.
15. Meese, R.; Wallace, N. Testing the present value relation for housing prices: Should I leave my house in San Francisco? *J. Urban Econ.* **1994**, *35*, 245–266.
16. Poterba, J. Tax subsidies to owner-occupied housing: An asset market approach. *Q. J. Econ.* **1984**, *99*, 729–752.
17. Case, K.E.; Shiller, R.J. The efficiency of the market for single family homes. *Am. Econ. Rev.* **1989**, *79*, 125–137.
18. Bourassa, S.C.; Hoesli, M.; Oikarinen, E. Measuring house price bubbles. *Real Estate Econ.* **2016**, *47*, 534–563.
19. Campbell, J.Y.; Shiller, R.J. The dividend-price ratio and expectations of future dividends and discount factors. *Rev. Financ. Stud.* **1988**, *1*, 195–227.
20. Campbell, J.Y.; Shiller, R.J. Stock prices, earnings, and expected dividends. *J. Financ.* **1988**, *43*, 661–676.
21. Campbell, J.Y.; Shiller, R.J. Cointegration and tests of present value models. *J. Political Econ.* **1987**, *95*, 1062–1088.
22. Diba, B.T.; Grossman, H.I. Explosive rational bubbles in stock prices? *Am. Econ. Rev.* **1988**, *78*, 520–530.
23. Gallin, J. The long-run relationship between house prices and rents. *Real Estate Econ.* **2008**, *36*, 635–658.
24. Brooks, C.; Katsaris, A.; McGough, T.; Tsolacos, S. Testing for bubbles in real estate price cycles. *J. Prop. Res.* **2001**, *18*, 341–346.
25. Stock, J.H.; Watson, M.W. Evidence on structural instability in macroeconomic time series relations. *J. Bus. Econ. Stat.* **1996**, *14*, 11–30.
26. Elliott, G.; Muller, U.K. Confidence sets for the date of a single break in linear time series regressions. *J. Econ.* **2007**, *141*, 1196–1218.
27. Elliott, G.; Muller, U.K. Pre and post break parameter inference. *J. Econ.* **2014**, *180*, 141–157.
28. Pesaran, M.H.; Pick, A.; Timmermann, A. Variable selection, estimation and inference for multi-period forecasting problems. *J. Econ.* **2011**, *164*, 173–187.
29. Renigier-Biłozor, M.; Walacik, M.; Żróbek, S.; d’Amato, M. Forced sale discount on property market—How to assess it? *Land Use Policy* **2018**, *78*, 104–115.
30. Berry, M.; Dalton, T. Housing prices and policy dilemmas: A peculiarly Australian problem? *Urban Policy Res.* **2004**, *22*, 69–91.
31. Shi, S.; Kabir, M.H. Catch animal spirits in auction: Evidence from New Zealand property market. *Real Estate Econ.* **2018**, *46*, 59–84.
32. Commins, P. Lowe Warns Rates Could Stay Low ‘for a Long, Long Time’. *The Australian Financial Review*. Available online: <https://www.afr.com/policy/economy/lowe-warns-rates-could-stay-low-for-a-long-long-time-20190925-p52uun> (accessed on 25 September 2019).