Housing Risk and Returns in Submarkets with Spatial Dependence and Heterogeneity

P. S. Morawakage¹ · G. Earl¹ · B. Liu¹ · E. Roca¹ · A. Omura¹

Accepted: 3 November 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Employing a recently developed panel econometric technique, first, we show that accounting for spatial dependence and heterogeneity yields more accurate risk factor coefficients and abnormal housing returns. Rather than systematic risks, idiosyncratic risks explain the variations in residential housing excess returns. After controlling for asset-specific and systematic risk factors, the positive and significant impact of the unobservable common factors on the excess returns suggests that speculative market forces drive the housing excess returns. Second, we then analyze the risks and returns of houses in affordable and expensive submarkets allowing for spatial dependence and heterogeneity. We find that houses in the affordable submarkets perform better than houses in the expensive submarkets. Thus, the potential demand for houses in the affordable submarket may aggravate the housing affordability crisis. Our study's results, therefore, encourage policymakers and investors to view the housing market as a collection of regional units and submarkets, but not as a single national market.

Keywords AMG · Housing risk-return · Spatial dependence · Speculation · Submarkets

JEL Codes C5 · G1 · I3 · R1 · R3

Introduction

Private residential housing is a real estate investment category that represents a significant part of global wealth. Thus, the accurate modeling of the risks and returns of housing investment is of interest to the investment circle and is a major concern

✉ P. S. Morawakage
prabath.morawakage@griffithuni.edu.au; morawakageps@kln.ac.lk

¹ Department of Accounting, Finance, and Economics, Griffith University, Nathan, Queensland 4111, Australia

in the academic literature. The housing asset market is unique by nature, as local factors such as crime rate, ethnicities, and amenities that vary across geographic locations substantially affect the level of investment risk (Kiefer, 2011; Koulizos, 2011; Peng, 2016). Local factor heterogeneity makes the idiosyncratic or asset-specific risks significant when determining the expected returns on residential housing investments (Case et al., 2011; Pedersen et al., 2014; Simlai, 2018). Hence, it is recommended that such idiosyncratic characteristics, especially those risks specific to each spatial unit, be considered when modeling housing risk and returns. Another unique aspect of residential housing pricing is the presence of spatial dependence among geographical areas (e.g., with the neighboring cities and suburbs) within a country, as the regions are interconnected (Costello et al., 2011; Hudson et al., 2018; Meen 1999; Nneji et al., 2015).

Analyzing the housing markets in different spatial units, therefore, provides important information not available at the aggregate level (Bangura & Lee, 2020; Beracha and Skiba, 2013). However, no study has previously considered spatial heterogeneity and dependence in disentangling the dynamic relationships between the risk-return of the housing markets among subregions. To close this important gap, the study at hand analyzes the housing market (based on the postcodes) in Brisbane, the capital city of the state of Queensland in Australia.

The spatial dependence and heterogeneous slope coefficients across cross-sections are two major problems in analyzing housing markets with panel data, as they produce biased estimations (Oikarinen et al., 2018). In this context, the standard asset pricing models of mainstream finance¹ are inappropriate for estimating the expected returns on residential housing investments (Hyun & Milcheva, 2018; Lusht, 1988). For this reason, we need to make appropriate adjustments to an existing asset pricing model to analyze the risk-return relationship of the residential housing market. In particular, our model is based, in spirit, on the Arbitrage Pricing Theory (APT) (Ross, 1976) and on a robust econometric technique called Augmented Mean Group (AMG) (Eberhardt & Teal, 2010). By considering additional risk factors, APT overcomes the single market factor limitation in the widely used Capital Asset Pricing Model (CAPM) introduced by Sharpe (1964). AMG estimation accounts for spatial dependence and heterogeneity. We are the first to apply the AMG technique to model the risk and returns considering spatial dependence and heterogeneity across different locations.

The Australian housing market has been attracting the attention of global and domestic investors due to its recent performance (Cox & Pavletich, 2020; Prequin, 2019). As a result, according to the Australian Bureau of Statistics,² the overall value of the private residential real estate in Australia had reached close to 8 trillion AUD by 2020, which is more than three times the overall value of the companies

¹ Standard asset pricing models including CAPM by Sharpe (1964), five factor model by Fama & French (2015).

² <https://www.abs.gov.au/statistics/economy/price-indexes-and-inflation/residential-property-price-index/es-eight-capital-cities/latest-release#key-statistics>

listed in the Australian Stock Exchange (ASX).³ However, the significant increase in overall house prices has put pressure on lower-income groups. According to the 16th Annual Demographia International Housing Affordability Survey: 2020, all five leading housing markets in Australia are severely unaffordable (Median multiple above 5). Untargeted housing demand subsidies and asymmetric tax arrangements favoring residential housing investors have aggravated the housing affordability crisis (Blunden, 2016). A submarket analysis of cheaper versus expensive houses based on house prices may give us more insights into the housing affordability issues in Australia.

There are a handful of studies examining submarkets in the real estate market. These studies segment the housing market by the level of price (e.g., expensive and cheap locations within a city), and their examinations are limited to the price dynamics (see, for example, Bangura & Lee, 2020; Wilson et al., 2011). However, they do not identify the submarkets within each location. It is not easy to categorize submarkets based on prices within spatial units (postcodes in this study) because transaction-level data for many countries and cities are not publicly available (Peng, 2016). Therefore, we collected data for expensive and cheaper houses sold in each quarter at the postcode level in Brisbane, Australia, to simulate the quarterly housing returns for submarkets based on various house price cohorts.

Brisbane is one of the most popular interstate migration destinations among the major cities in Australia. This is evident from the fact that Brisbane has been the only city in Australia with net positive interstate migrants⁴ from 2010 to 2020 continuously, with a 7% quarterly compound growth rate. When it comes to the level of housing price, Brisbane was named the second most affordable housing market among the major cities in Australia (Cox & Pavletich, 2020). However, as Brisbane is expected to experience the highest housing price growth rate (20%) of all major cities in the country from 2019 to 2022, according to BIS Oxford Economics Institute, the housing affordability of the city may be under threat for the foreseeable future. On the other hand, the Brisbane housing market has not yet received much scholarly attention compared to the Melbourne and Sydney housing markets (Hatzvi & Otto, 2008; Hulse & Yates, 2017; Klimova & Lee, 2014; Randolph & Tice, 2013). These facts make Brisbane the ideal city for the empirical examination of housing risk and returns.

The contributions of this study are threefold. Firstly, we contribute to the real estate risk-return modeling literature. We introduce dynamic spatial dependence and heterogeneity factors into the subregional residential house risk-return modeling. Secondly, our study contributes to the residents and the real estate investors. In particular, the study offers a better understanding of the relationship between the risk and the returns on the residential house investments in cheaper and expensive submarkets to improve the effectiveness of investment decisions. Furthermore, investors can explicitly identify abnormal returns due to location-specific factors and

³ <https://www2.asx.com.au/about/market-statistics/historical-market-statistics>

⁴ <https://www.abs.gov.au/statistics/people/population/regional-internal-migration-estimates-provisional/latest-release#key-statistics>

unobservable common factors. Thirdly, our findings provide valuable information for the policymakers on what may affect the affordability of houses. This helps policymakers taking appropriate actions to help the economically vulnerable people in society.

Our empirical results show that conventional panel estimators that do not account for spatial dependence and heterogeneity overestimate the economic risk factors and underestimate the idiosyncratic risks. The AMG estimation yields more sensible risk factor parameters and shows that idiosyncratic risk factors determine the returns of residential housing at more disaggregated levels. The risk and return of residential houses vary across the postcodes and submarkets. The dynamic process, evolved through time-variant unobservable common factors, positively affects the housing excess returns in all the submarkets, possibly due to speculative market forces. On the other hand, the postal code-specific factors result in an abnormal loss.

Moreover, we find higher risk-adjusted returns in the affordable submarkets than in the expensive submarkets. The excess returns of the most affordable submarket are less exposed to risk factors as well. The land size, distance to CBD, and time on the market positively affect the excess returns of the houses in the second quartile. Crime rate and rental yield negatively affect the excess returns in all the submarkets. Better performances in the affordable submarkets may increase the potential demand for houses in that submarket, disadvantaging the first home buyers and lower-income groups.

The significant differences in risk-return relationships across locations and submarkets have important practical implications for government policies and investment decisions. The return predictions assuming homogenous risks across different locations may mislead investors. The local government policies and investment decisions may also be less productive when countrywide averages are considered. Therefore, we encourage policymakers and investors to view the housing market as a collection of spatial units and submarkets. The government may change the tax system to address housing affordability issues. For example, a capital gains tax on the affordable housing submarket may reduce the pressure on housing prices due to speculation. In addition, the government may rationalize the demand-side subsidies to benefit the most disadvantaged buyers (e.g., by subsidizing owner-occupied houses).

The remainder of the paper is structured as follows. Section 2 reviews the literature on risk-return modeling, spatial modeling, spatial dependence, submarket analysis, and methodological gaps. Section 3 explains the empirical framework and econometric techniques employed. Section 4 reports the results and discusses the findings. Finally, Section 5 concludes the study.

Literature Review

Risk and Return Modelling

Rational investors expect an extra return for each additional risk factor. The CAPM (Sharpe, 1964) is a widely used risk-return relationship model that assumes a

positive relationship between market portfolio return and expected returns on an asset. The CAPM is criticized for using a single factor and for the user's difficulty in choosing a better proxy for the market portfolio return. These CAPM limitations led to the development of multi-factor models such as APT (Ross, 1976) and the five-factor model (Fama & French, 2015). Previous studies have thoroughly examined the risks and returns of publicly traded real estate securities such as real estate investment trusts (REITs) and real estate-related stocks (Bond et al., 2003; Giacomini et al., 2015). By employing multi-factor models, those studies have found that equity market factors and idiosyncratic risks significantly affect public real estate returns. However, less attention has been paid to the private residential housing market.

Case et al. (2011) analyze housing risk-returns in 151 metropolitan statistical areas (MSA) in the USA. In their model, they specify standard risk factors such as market, size (SMB), and momentum as the determinants of the city-specific indexed housing returns. The market factor is the only significant determinant of housing returns according to them. Ho et al. (2015) find that changes in macroeconomic and real estate variables do not better explain residential returns compared to retail and office returns. Since housing markets are relatively information inefficient, their returns are not as predictable as stocks, bonds, and public real estate returns Chun et al., 2004). Hence, mapping the housing asset class to their specific factor exposures is essential when implementing housing risk-return models in practice.

The expected returns can be significantly different from realized returns in residential housing markets (Shilling, 2003). The reason for this may be the idiosyncratic risk factors presented in the error of a risk-return model. Previous studies have estimated the overall idiosyncratic risk factor of real estate investments by estimating the residuals of the CAPM equation, but without deeply exploring the variety of factors that drive idiosyncratic risk (Bali et al., 2005; Cannon et al., 2006; Case et al., 2011). Therefore, modeling idiosyncratic risk factors such as the marketing period, liquidity, spatial dependence, and neighborhood characteristics of the private residential housing asset class will significantly contribute to the literature.

Spatial Modelling

The returns from housing investment generated within a specific neighborhood are inseparable from local factors such as schools, crime rates, demography, and geography (Goetzmann & Spiegel, 1997). Location-specific information creates differences among residential house prices (Kiefer, 2011; Koulizos, 2011). Therefore, investigating cross-market differences in house prices is a growing interest in regional and housing studies (Han, 2013; Peng, 2016). Spatial differences provide good opportunities to diversify investment portfolios. Milcheva et al. (2020) find that properties with high geographic dispersion significantly outperform the market compared to concentrated property holdings. Simlai (2018) employs a spatial regression to model the spatial dependency of housing prices along with hedonic factors in a cross-sectional study. The author finds a significant volatility interdependence among the individual house prices because of

geographical proximity. Hence, studying risk and return relationships at more disaggregated levels may have significant implications for investments and policy decisions.

A few studies have attempted to investigate the housing risks and returns at disaggregated levels (see for examples, Cannon et al., 2006; Goetzmann & Spiegel, 1997; Han, 2013). However, they have not accounted for spatial dependence and heterogeneity across spatial units (Oikarinen et al., 2018). Our study is closely related to the study of Case et al. (2011) that investigates the housing returns of metropolitan areas in the USA context using indexed data. In our study, we analyze postcode level housing returns with transactional level data in Australia. We also examine four housing submarkets based on price quartiles in each postcode and consider spatial dependence and heterogeneity.

Spatial dependence arises due to outcomes in one area being affected by 1) outcomes in nearby areas (spatial lags of dependent variables), 2) covariations of other variables in nearby areas (spatial lags of independent variables), and 3) errors from nearby areas (spatial autoregressive error). Elhorst (2012) provides a comprehensive review of models, methods, and inferences relating to the above spatial interactions. In spatial models, interactions between cross-sectional units are typically modeled by a distance weight matrix (Lam & Souza, 2020). Since computing the distance weight matrix for a larger sample is difficult, some studies have used a GMM approach with instrumental variables (IV) to treat spatial dependence and endogeneity (see for example, Kapoor et al. (2007)). However, finding instruments that are correlated with the regressors and not correlated with the unobserved factors (error) would be difficult (De Hoyos & Sarafidis, 2006).

Moreover, unobserved common factors lead to correlations in the disturbances across spatial units and to correlations between the disturbances and the regressors (Eberhardt and Bond, 2009; Pesaran, 2006, 2007). These unobservable common factors may represent social norms, neighborhood effects, herd behavior, and other behavioral factors (De Hoyos & Sarafidis, 2006). The AMG estimation procedure introduced by Eberhardt and Bond (2009) well accounts for these unobservable common factors by employing a time dummy variable system called common dynamic process. Our study focuses on the housing returns from various geographical locations and accounts for spatial dependence using a common dynamic process evolved from unobservable common factors. We also allow spatial heterogeneity by employing a mean group estimator.

Speculation, Price Spill-over, and Spatial Dependence

It is well established in the housing literature that house prices driven by speculation and non-fundamentals are transmitted between neighboring and more distant regions. Costello et al. (2011) examine the house price spill-over from one state to another driven by speculative activities in Australia. They find that the state of New South Wales is more sensitive to speculative spill-overs from other states. Nneji et al. (2015) also analyze whether speculative bubbles in one region lead to bubbles in other regions, using USA census divisions data. They show that speculative bubble spill-over is multi-directional. The literature suggests that house price spill-over

is driven by speculation or non-rational and non-fundamental factors (DeFusco, Ding, Ferreira & Gyourko, 2018). Possible reasons for the spill-over of speculative forces are the herding effect, extensive news coverage of growing real estate markets, and the migrating equity transfer argument (Nneji et al., 2015).

In relation to the housing market, spatial dependence is also known as spatial autocorrelation and is represented as the spatial correlation between a given house price and the prices of houses adjacent to it (Lee & Robinson, 2016; Pijnenburg, 2017). Spatial correlation across property transactions is higher in a rising market than in a falling market (Hyun & Milcheva, 2018). Pijnenburg (2017) also defines the spatial dependence of house prices as the ripple effect. Thus, the spatial dependence of housing returns demonstrates the housing price spill-over effect. We can model the spatial dependence using unobservable common factors that affect cross-sectional units (Lee & Robinson, 2016). Unobservable common factors in a risk-return model represent the price spill-over effect that can arise due to a speculative force at a given point in time.

Submarket Analysis

Understanding the housing submarket dynamics has become increasingly important in conducting market and policy analyses (Costello et al., 2019, Xiao, Webster & Orford, 2016). However, the literature (see for example, Watkins, 2001; Xiao et al., 2016) highlights the following limitations in identifying housing submarkets. First, there is no consensus as to how submarkets should be identified in practice. Second, the urban area under examination varies from study to study. Third, the market conditions vary across studies. Related to these limitations, Costello et al. (2019) argue that structural dimensions become irrelevant in the long run and spatial dimensions evolve with time. In particular, the spatial dimensions are affected by factors like intra-urban migrations, new housing supply, and the tastes and preferences of the market.

In the past studies, the housing market was subdivided based on spatial determinants and property characteristics (including structural dimensions). For the former, the market was subdivided according to the location of the properties, and the geographic areas/spatial units were recognized in terms of determinants like suburbs (see for example, Costello et al., 2019), streets (see for example, Xiao et al., 2016), and postcodes (see for example, Watkins, 2001). For the latter, the factors used in recognizing the submarkets included the types of properties such as apartments, detached houses, offices (see for example, Melser & Hill, 2018; Young, 2008) and price level (see for example, Bangura & Lee, 2020). Related to this, Watkins (2001) argues that the housing market should be subdivided by using both spatial and structural dimensions simultaneously. One aspect of subdivision that is not well examined is subdividing the market by the level of affordability. To address this, we categorized the housing market in each spatial unit into four price cohorts according to house prices and identified four submarkets ranging from most affordable to most expensive. Our submarkets evolve with time and can be used in conjunction with any spatial context or physical dimension.

Methodological Gaps

Previous studies have widely used time series models to examine the risk and returns of real estate assets (see, for example, Bond et al., 2003; Melser & Hill, 2019; Shilling, 2003). However, their results were potentially limited by the endogeneity problem due to spatial differences (Goetzmann & Spiegel, 1997; Wright & Yanotti, 2019). Advanced panel estimation techniques overcome such limitations and produce more efficient coefficients (Oikarinen et al., 2018). Ho et al. (2015) employ a pooled-panel least squares model and index data to investigate the risks and returns of real estate assets. Their study covers a variety of real estate sectors and many cities globally. They do not consider any cointegrations or spatial dependence in their multi-factor model. In the context of cross-sectional housing risk-return analysis, spatial dependence is particularly salient, given the connectedness of housing prices through price spill-over. Therefore, we employ a recently developed panel data technique called AMG estimation that is robust to the impact of correlation across panel units and flexible with the standard assumption of cross-sectional independence.

Oikarinen et al. (2018) use AMG estimation to examine the USA house price dynamics. They emphasize the importance of taking spatial variations into account for the assessment of housing risks. However, a minimal number of studies in housing have addressed slope heterogeneity and spatial dependence across the cross-sections. Oikarinen et al. (2018) were the first to address those issues by applying the AMG technique to house price dynamics. To the best of our knowledge, no study has employed AMG to investigate the risk-return relationship of residential housing.

A common dynamic process that accounts for spatial dependence is introduced in the AMG estimation. This variable evolves from the year dummy coefficients of a pooled regression in the first difference. It represents the level equivalent mean unobserved common factors across all cross-sections (Eberhardt & Bond, 2009). Oikarinen et al. (2018) assume that the common dynamic process in AMG estimation represents both fundamental and non-fundamental unobservable factors.

There are also studies that suffer from analyst and transactor⁵ errors as they employ many transformation techniques to overcome the smoothing bias of indexed data (Melser & Hill, 2019; Pedersen et al., 2014). Those studies have not accounted for abnormal performances because transformed data do not reflect market activities (Linneman, 1986). The market value of real estate assets is only known when properties are sold. To obtain robust results, it is necessary to address these inherent statistical issues in performance data analysis of non-traded asset classes such as private residential housing (Farrelly & Stevenson, 2019). Our study employs a panel data approach with transaction-level data to analyze residential housing asset risk and returns.

⁵ Some sellers may sell houses too cheaply and some buyers may pay too much for their homes relative to the market expectations (Linneman, 1986).

Methodology

Empirical Framework

The literature suggests that economic and housing-specific risk factors primarily determine the risk-return relationship of housing assets. Following the arbitrage pricing theory in Finance, this study employs the multi-factor Eq. (1) to account for the above risk factors.

$$(r_i - r_f)_t = \beta_{1i}(r_m - r_f)_t + \beta_{2i}Crs_{spread}_t + \beta_{3i}RY_{it} + \beta_{4i}TOM_{it} + \beta_{5i}Dist_{it} + \beta_{6i}LA_{it} + \beta_{7i}Crime_{it} + \mu_{it} \quad (1)$$

where excess return $(r_i - r_f)$ is the dependent variable, i is the respective postcode, and t is the quarterly period. r is the quarterly log-returns of houses considering the median house price change between two quarters. r_f is the three months quarterly Treasury bill rate for each period. $\beta_{1i} - \beta_{7i}$ are the postcode-specific risk factor loadings of the observable risk factors. Equity market premium $(r_m - r_f)$ and Crs_{spread} are the vectors of observable systematic risk factors. RY , TOM , $Dist$, LA , and $Crime$ are the vectors of observable idiosyncratic risk factors. The error term, μ_{it} , represents unexplained idiosyncratic risk. All variables are inflationary adjusted.

r_m is the ASX ordinary share market index return for each period. Crs_{spread} is the housing credit spread (excess of housing lending rate to 10-year Australian government bond yield). RY is the rental yield, and TOM is time on the market (the difference between the first advertised date and the sale date). $Dist$ is the median distance to Central Business District (CBD), and LA is the median land size of houses. $Crime$ is the past (lagged by two quarters) crime rate. Time on the market, distance to CBD, and land size correspond to the same houses for which we calculated the housing returns. Rental yield accounts for the ability of the house to generate a fixed cash flow. Distance, land size, and crime rate represent the location risk factors. Time on the market is a proxy for liquidity risk. We observed a multicollinearity issue due to the higher correlation among the housing-specific risk factors (please see Table 6 in the appendix).

Therefore, we perform Principal Component Analysis (PCA) for the housing-specific risk factors to determine the number of correlated factors and reduce them into a set of fewer, uncorrelated components (Lettau & Pelger, 2020). PCA also helps to overcome the multicollinearity problem by developing a risk index with specific factor loadings. Eigenvalues suggest two components: PC1 and PC 2 that extract 57% to 62% of the variation in the total variables in each submarket and the overall market. We considered both principal components to derive two different housing risk indices: $HRI1$ and $HRI2$ because each index has different weights. The indices are concentrated on different factors. Index scores are the weighted sum of the standardized variables as given in Eqs. (2) and (3).

$$HRI1 = w1 * STDLA + w1 * STDDist + w1 * STDTOM \quad (2)$$

$$HRI2 = w2 * 1 STDCrime + w2 * 2 STDRY \quad (3)$$

where *HRI* is the housing risk index; *w* is the weight or factor loading of each housing risk factor obtained from PCA; *STD* means standardized variables; *HRI1* consists of factors with weights more than 0.4 suggested by rotated⁶ PC1; *HRI2* consists of the factors with weights more than 0.4, suggested by rotated PC2 (Sarstedt & Mooi, 2019). The weights of *HRI1* are concentrated on increases in land size, distance to CBD, and time on the market. The weights of *HRI2* are concentrated on the increases in crime rate and rental yield. Table 1 reports the loadings of the risk factors under each component and price cohort.

Then, we replace the housing-specific risk variables in Eq. (1) with derived risk indices as given below.

$$(r_i - r_f)_t = \beta_{1i}(r_m - r_f)_t + \beta_{2i}Crsprad_t + \varphi_{1i}HRI1_{it} + \varphi_{2i}HRI2_{it} + \mu_{it} \quad (4)$$

where φ_{1i} and φ_{2i} are the factor loadings of *HRI1* and *HRI2*, respectively.

Our Eq. (4) controls for systematic risks and encompasses observable idiosyncratic risk factors. Systematic factors affect the return on investments in different ways. Coefficient of market risk premium ($r_m - r_f$) represents the sensitivity of housing asset returns to equity market performances. The smaller or negative coefficient implies that housing investment provides an opportunity to diversify from equity investments. Real estate credit spread is important to understanding the space market (Sivitanides et al., 2003). A higher credit spread reduces the housing demand and leads to lower returns (David, 2013).

A house price is comprised of two components: the structure and plot of land (Davis and Heathcote, 2007). Land leverage theory (the ratio of land value to total property value) suggests that a large part of house price change is attributable to the land price change (Baker & Filbeck, 2013; Davis & Heathcote, 2007). There is strong empirical evidence for a relationship between land size and its value (see, for example, Thorsnes and McMillen, 1998; Colwell & Sirmans, 1978, 1993; Ecker & Isakson, 2005). While the relationship may not always be linear,⁷ studies generally support a positive relationship between the land size and its value. In relation to the impact of the unit price of the land on its overall value, there could be some other factors like location affecting it (Karl & Gareth, 2005). However, according to Lin and Evans (2000), there exists a positive relationship between land price and size after controlling for location and infrastructure services. Following these findings, our study selects land size as one of the housing risk (*HRI1*) factors. We indicate that ignoring location is a limitation in determining the value of the land.

The distance to the CBD is also important in determining property prices (Boarnet, 1994). Monocentric models show that rents and property values decline as the distance to the CBD increases (Meen, 2016). However, when the distance to CBD is longer, the respective land size is larger, and hence land leverage may positively impact the excess returns of houses away from the CBD. On the other hand, the land

⁶ We employed orthogonal varimax rotation to rotate the factors in each principal component.

⁷ There may be a cost associated with combining small lots into one large lot or dividing a large lot into smaller lots. These may cause the shape of the relationship to be concave or convex.

Table 1 Weights assigned for each factor based on the PCA

Variable	Q1		Q2		Q3		Q4		Overall	
	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2
Rental yield	0.156	0.422	0.022	0.520	0.010	0.449	0.082	0.511	0.105	0.409
Distance to CBD	0.600	0.129	0.599	0.202	0.614	0.197	0.615	0.136	0.616	0.131
Time on Market	0.421	0.246	0.387	0.2913	0.429	0.2615	0.370	0.304	0.402	0.334
Land size	0.658	-0.288	0.695	-0.262	0.653	-0.359	0.685	-0.238	0.664	-0.291
Crime rate	-0.070	0.814	-0.085	0.732	-0.058	0.750	-0.100	0.756	-0.082	0.787

Housing-specific risk factors were converted into two uncorrelated indices based on the principal component analysis. Eigenvalues suggest two principal components: PC1 and PC 2 that extract 57% to 62% of the variation in the total variables in each submarket and the overall market. HRI1 consists of factors with weights more than 0.4 suggested by PC1. HRI2 consists of factors with weights more than 0.4 suggested by PC2 (Sarstedt & Mooi, 2019). Distance to CBD, land size, and time on the market have significant weights on HRI1, respectively. The crime rate and rental yield have significant weights on HRI2, respectively

and house prices close to the CBD have increased due to high supply inelasticity (Davis & Heathcote, 2007), and hence the impact of land leverage on the excess returns is mixed. This study employs distance and land size variables as proxies for location risk (Song & Sohn, 2007).

Since real estate assets are thinly traded over long holding periods with a significant transaction cost, housing assets are illiquid. The degree of illiquidity is typically measured by the time on the market (TOM) (Anglin et al., 2003; Domian et al., 2015; Lin & Liu, 2008; Lin & Vandell, 2007). TOM also measures the marketing period risk that is specific to real estate properties. There is a positive relationship between TOM and the housing excess returns because real estate agents retain the properties in the stock only if there is an associated benefit (Anglin et al., 2003; Cheng et al., 2008).

The rental yield is a fixed income and a key indicator of housing market equilibrium because it directly affects housing prices (Ayuso & Restoy, 2006, 2007). On average, rental yield is around 40% of total housing returns (Melser & Hill, 2019). The rental yield is derived by dividing the gross rental income by house price. Rental yield is also considered the discounting factor for cashflows generated by residential housing. Therefore, a higher rental yield is usually associated with declining housing prices and excess returns (Davis et al., 2008; Kouwenberg & Zwinkels, 2014).

The crime rate negatively affects house prices. Homes in the higher crime areas are well known to sell at deep discounts or to take a longer time to sell (Klimova & Lee, 2014). In addition, spatially targeted crime prevention policies influence local house prices differently (Gibbons & Machin, 2008). However, the crime rate as a neighborhood factor has not been previously examined within a risk and return framework. Instead, there are studies that have considered schools, transport, employment, personal income, population, age groups, and ethnicities (Case & Shiller, 1990; Gibbons & Machin, 2008; Goetzmann & Spiegel, 1997). We employ the past (lagged by two periods) crime rate as a neighborhood risk factor because immediate crimes would not impact contemporaneous property returns (Klimova & Lee, 2014).

According to the multivariate analysis of variance (see Table 7 in the appendix), the distance, land size, time on the market, rental yield, and crime rate show substantial spatial heterogeneity that can influence the housing returns and risk. Due to the significant variations in those variables across locations, allowing risk factor heterogeneity in the estimation process is necessary. The mean group (MG) estimator allows heterogeneity across the cross-sections (Pesaran & Smith, 1995). However, the MG estimator does not account for spatial dependence. As housing markets are interconnected, taking spatial dependence into account when estimating the risk factor coefficients is also warranted.

The error term μ_{it} of the risk-return Eq. (4) represents the unexplained idiosyncratic risk. Since spatial dependence is an element of the error term (Costello et al., 2011; Eberhardt & Bond, 2009), our panel econometric framework should separate spatial dependence from the error term and cross-section specific factors. The spatial dependence variable is also assumed to be non-linear and non-stationary (Drake, 1995; Eberhardt & Teal, 2010; Pijnenburg, 2017). The unobserved

factors that create interdependence across postcodes may be correlated with the included regressors in Eq. (4). Traditional panel data approaches like mean group, fixed and random effect estimators will be biased and inconsistent in this context (De Hoyos & Sarafidis, 2006; Oikarinen et al., 2018). Therefore, we augment each regression with a common dynamic process to consider spatial dependence and estimate with mean group technique, as Eberhardt and Bond (2009) explain.

$$\mu_{it} = \alpha_i + \lambda'_{it} f_t + \varepsilon_{it} \quad (5)$$

$$\hat{\mu}_t^* = h(\bar{\lambda} f_t) \quad (6)$$

μ_{it} is the error term of Eq. (4), and it is decomposed into three parts. α_i represents the abnormal gain or loss attributable to postcode-specific fixed effects. f_t represents the unobservable time-variant common factors with postcode specific factor loadings λ'_{it} . ε_{it} is the unexplained error term independently and identically distributed. $\hat{\mu}_t^*$ is a function $h(\cdot)$ of unobservable common factor (f_t).

$$\begin{aligned} \Delta(r_i - r_f)_t &= \delta_1 \Delta(r_m - r_f)_t + \delta_2 \Delta Crspread_t + \delta_3 \Delta HRI1_{it} + \delta_4 \Delta HRI2_{it} + \sum_{t=2}^T c_t \Delta D_t + \varepsilon_{it} \\ &\Rightarrow \hat{c}_t \equiv \hat{\mu}_t^* \end{aligned} \quad (7)$$

Equation (7) represents a standard first differenced OLS regression with $(T - 1)$ year dummies in first differences to collect the year dummy coefficients (c_t). Estimated c_t ($\hat{c}_t$) is renamed as $\hat{\mu}_t^*$. Then, $\hat{\mu}_t^*$, evolving from the unobservable common factors, is augmented to 20 postcode-specific OLS regressions at levels as given below.

$$(r_i - r_f)_t = \alpha_i + \beta_{1i}(r_m - r_f)_t + \beta_{2i} Crspread_t + \varphi_{1i} HRI1_{it} + \varphi_{2i} HRI2_{it} + d_i \hat{\mu}_t^* + \varepsilon_{it} \quad (8)$$

where the postcode-specific coefficient d_i represents the implicit factor loading on the common dynamic process for period t , $\hat{\mu}_t^*$. Common dynamic process ($\hat{\mu}_t^*$), evolved from the unobservable common factors, represents the spatial dependence of housing returns. Eberhardt and Teal (2010) introduced the AMG estimation to study the production function and to derive the intercept that represents the Total Factor Productivity (TFP) parameter allowing for technological differences across different countries. They argue that once the elasticities of regressors differ across the cross-sections (heterogeneous), the intercept of the production function can no longer be interpreted as a common TFP-level estimate. As a solution, they augment a common dynamic process to capture the common TFP-level other than the country-specific TFP.

Similarly, the intercept as the abnormal return derived from asset pricing models in a panel data framework should be carefully analyzed because the idiosyncratic risk factors differ across the spatial units and correlate with the unobservable common factors. The AMG estimation method overcomes such shortcomings and allows us to model abnormal returns due to location-specific factors and countrywide speculative bubbles, or common shocks, separately.

The common dynamic process ($\hat{\mu}_t^*$) in a risk-return model represents an abnormal return due to time-varying unobservable common shocks. This shock is common to all postcodes at each point in time, but the response of each postcode to the shock is different. This is because there are distinctive price determinants within each postcode. For example, an abnormal return can be due to a speculative bubble in the general housing market; however, postcode-specific factors like amenities, liquidity, and demography determine how the market responds to such a bubble. For this reason, the AMG model employs different factor loadings (d_i) for the common dynamic process at different postcodes. If we do not account for unobservable common factors, they appear on the residuals. As a result, the residual of CAPM is no longer a better proxy for idiosyncratic risks when spatial dependence is present. AMG accounts for spatial dependence by augmenting a common dynamic process with the mean evolution of unobserved common factors across all postcodes.

Data and Sample

There exist several geographic segmentation systems in Australia. They include the postcode divisions as well as the Australian Statistical Geography Standards (ASGS) such as statistical area levels (SAs) for analyzing Census data. The ASGS is a social geography standard developed to reflect the location of people and communities. It is used for the release and analysis of statistics and other data. There are four SAs under the ASGS. SA 1 is the smallest geographic area for which census data are collected, and it has a population of 400 on average. SA 1 area definitions are based on the distribution of dwellings, not on the area size. The postcodes in Australia are approximated using one or more SA 1 spatial units.

The postcode system was primarily developed for the mail delivery service like the ZIP-code system in the USA. In the private sector, the postcode system is widely used for segmentation in the real estate market for marketing and lending activities.⁸ This is partly because the postcode system allows its users to identify the exact location of a property. Since postcode is a well-known and easily collected component of an address, many researchers and businesses use it to link data to a geographic area for spatial analysis. Following such research and industry practice, most of the real estate sector organizations (e.g., lenders, housing associations, agents, research companies etc.) and data providers use the postcode as a spatial unit of data analysis because houses can be easily identified by their addresses.

Given the well-established use of postcodes as a spatial unit in the Australian real estate market, collecting postcode level data is a reasonable and practical start to investigating the housing risk and return in Brisbane. Our sample data were collected from the Australian Urban Research Infrastructure Network (AURIN) database using the postcodes. We used population and percentage of 'separate houses' as criteria to choose the postcodes. The colored area of the following map (Fig. 1) shows the geographical area of the Greater Brisbane city covered for the study.

⁸ https://www.corelogic.com.au/sites/default/files/2018-01/cl_segmentation_flyer.pdf, and <https://www.nhffc.gov.au/what-we-do/property-price-caps/>

Fig. 1 Postcodes used for the study. The yellow area represents the postcodes close to Sunshine Coast CBD. The red area represents the postcodes close to Brisbane CBD. The blue area represents the postcodes near the Gold Coast CBD. Altogether we collected data from 20 postcodes. The Sunshine Coast, Brisbane, and the Gold Coast are situated in the Greater Brisbane city of Queensland

Our sample represents 34% of the population and 32% of the private dwellings in Greater Brisbane. The majority of Australian houses are separate houses (73%) and have three to four bedrooms. Therefore, we selected only the sales transactions of separate houses with three to four bedrooms in each postcode for each quarter period from January 2010 to December 2019. Altogether, we collected data associated with 93,869 transactions (117 houses, on average, per quarter in each postcode). In the analysis, we divided the quarterly transactions within each postcode into four quartiles as explained in the next subsection and calculated the median prices of these subsets of our sample. We then calculated the quarterly postcode-specific excess returns by subtracting the three-month Treasury bill rate from the log-transformed median price return. The main tests are conducted on each quartile dataset separately, and therefore, each regression analysis consists of 780 observations (39 quarters multiplied by 20 postcodes).

Defining the Submarkets

There are studies that have employed spatial and structural dimensions to define the housing submarkets (Costello et al., 2019, Watkins, 2001). In this study, as explained above, the postcodes are used to identify the spatial dimensions and geographic locale of houses in the data collection process. However, as one of the main purposes of this study is to examine the impact of spatial dependence and heterogeneity of affordable housing returns, we use the relative price level of houses (cheaper to expensive) in each postcode as the basis for segmenting the real estate market into submarkets. Our market segmentation extends the method used by studies like those of Bangura and Lee (2020) and Wilson et al. (2011), which divided the whole market into cheaper and expensive areas.

Figure 2 given below explains how we recognize the submarkets within each postcode. For each quarterly period, on average, there are 117 sample data items (i.e., data on prices of houses being purchased/sold) available for each postcode, and these are divided into four quartiles based on the price level. In the analysis, we define the lowest two quartiles (Q1 and Q2) as the affordable submarkets and the top two quartiles (Q3 and Q4) as the expensive submarkets. We then use the log-transformed median quartile prices for each postcode and for each period in the analysis.

Fig. 2 Submarket definitions

Fig. 3 The volatility of excess returns (exr). Time series plots given in Fig. 3 provide an initial review of the excess returns (exr) of each postcode with quarterly frequency. We see considerable temporal and cross-sectional variations in the excess return series of the selected postcodes and submarkets. Overall, the excess returns of the expensive submarkets (exr Q3 and exr Q4) are more volatile than the affordable submarkets (exr Q1 and exr Q2)

Time series plots given in Fig. 3 provide an initial review of the excess returns (exr) of each postcode with quarterly frequency. We see considerable temporal and cross-sectional variations in the excess return series of the selected postcodes and submarkets. Overall, the excess returns of the expensive submarkets (exr Q3 and exr Q4) are more volatile than the affordable submarkets (exr Q1 and exr Q2).

ASX ordinary shares index data were collected from Bloomberg. Three-month Treasury bill rates data were collected from the Reserve Bank of Australia. Housing lending rates and inflation data were collected from the Australian Bureau of Statistics. Australian residential housing index, S&P 500 index, and USA Treasury bill rates data were collected from Federal Reserve Economic Data (FRED). Weekly rent data were collected from the SQM research website. Time on the market, distance to CBD, and land size of houses were collected from the AURIN database. Crime rate data were obtained from the Queensland Police website. We used STATA software to analyze the data and estimate the model.

Results and Discussion

Our empirical analysis has four phases.

- (1) Estimation of the risk factor parameters of residential housing without considering spatial dependence and heterogeneity
- (2) Estimation of the risk factor parameters of residential housing only considering spatial heterogeneity
- (3) Estimation of the risk factor parameters of residential housing considering both spatial dependence and heterogeneity

Table 2 Variables and summary statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
r_{it} Q1	780	0.390%	5.237%	−32.996%	28.825%
r_{it} Q2	780	0.360%	4.424%	−17.934%	20.988%
r_{it} Q3	780	0.370%	5.870%	−32.231%	30.167%
r_{it} Q4	780	0.404%	9.280%	−55.207%	51.589%
r_{it} Overall	780	0.370%	4.857%	−23.507%	22.891%
Equity market premium	39	0.204%	6.063%	−15.191%	8.861%
Credit Spread	39	2.367%	0.709%	1.041%	3.818%
Time on the Market Q1	780	58	31	11	238
Time on the Market Q2	780	51	30	7	222
Time on the Market Q3	780	51	33	3	249
Time on the Market Q4	780	62	46	0	384
Time on Market Overall	780	53	25	12	188
Distance to CBD Q1	780	19.482	14.018	4.566	66.518
Distance to CBD Q2	780	19.492	14.059	3.349	66.447
Distance to CBD Q3	780	19.436	14.168	3.812	69.554
Distance to CBD Q4	780	19.595	13.972	3.079	71.288
Distance to CBD Overall	780	19.470	14.060	4.250	65.720
Land Size Q1	780	590	146	225	1389
Land Size Q2	780	632	152	260	2210
Land Size Q3	780	859	895	320	8752
Land Size Q4	780	2119	5371	400	75,570
Land Size Overall	780	701	320	320	5034
Rental Yield	780	4.712%	0.640%	2.985%	6.155%
Crime rate	780	1.252%	0.448%	0.471%	3.414%

All data are quarterly data. r_{it} is the residential housing returns of i^{th} postcode at each quarter t . Returns were calculated under each price quartile and with all the data. Q1 to Q4 are price quartiles or submarkets. Equity market premium is $(r_m - r_f)$ where r_m is the ASX ordinary share market index return for each period and r_f is the three months quarterly adjusted Treasury bill rate for each period. Credit spread and equity market premium are common systematic risk factors across the postcodes. Other variables are housing-specific risk factors. Time on the market is in days, distance to CBD is in Kilometers, and land size is squared meters. Time on the market, distance, and land size are also organized under each price quartile and postcode. The rental yield and crime rate of each postcode are common for all price quartiles

- (4) Estimation of the risk factor parameters of residential housing in price submarkets considering both the heterogeneous slope coefficients and spatial dependence

Table 2 shows summary statistics of the variables included in the proposed multi-factor model. Statistics are reported under overall and submarket levels. Credit spread and equity market premium are common systematic risk factors across the postcodes. Other variables are housing-specific risk factors.

Estimating the Risk-Return Model

Before estimating the risk factor coefficients, we tested Eq. (4) for cross-sectional dependence (CD) at each submarket with the Lagrange multiplier (LM) test (Breusch & Pagan, 1980), bias-adjusted LM test (Pesaran et al., 2008) and the cross-sectional dependence test (Pesaran, 2021). All the equations at different levels of affordability and overall sample are cross-sectionally dependent. Given the presence of cross-sectional dependence in the panel, the first-generation unit root tests and panel cointegration tests become invalid (Burdizzo & Sangiácomo, 2016; Persyn & Westerlund, 2008). Therefore, to analyze the stationarity of the variables, the cross-sectionally augmented Im-Pesaran-Shin (CIPS) panel unit root test (Pesaran, 2007) was employed. All the variables are stationary at the level. Though the variables are stationary at the level, there can be cointegrated relationships due to cross-sectional dependence. Hence, we employed the LM Bootstrap test of Westerlund and Edgerton (2007) as a second-generation panel cointegration test for all the equations applied. This test is the most appropriate because quarterly periods are greater than the cross-sections ($T > N$) in our study (Persyn & Westerlund, 2008). There is strong evidence of cointegration in all the models. These preliminary test results are given in the appendix (Tables 8, 9, 10).

Then, we estimated the risk-return Eq. (8) with data corresponding to the total sample. Under phase one, we estimated the equation with the fixed effect OLS (FE-OLS). FE-OLS estimation assumes homogenous risk factor coefficients across the postcodes. The p value corresponding to the slope homogeneity test (Pesaran and Yamagata, 2008) is significant at the 5% level, confirming that the coefficients are heterogeneous among the postcodes. Given the slope heterogeneity, we employed MG estimation at the second phase. MG estimation allows heterogeneous slopes but ignores spatial dependence. Therefore, we used the AMG estimation under the third phase of our analysis. AMG accounts for both slope heterogeneity and spatial dependence. Table 3 reports the results of all three estimations.

Spatial heterogeneity and dependence drive the idiosyncratic risk of residential housing assets because the response rates to various common shocks are different across the spatial units. Table 3 reveals that the idiosyncratic risk factors (*HRI1* and *HRI2*) and abnormal loss are underestimated, and the equity market factor is overestimated if we ignore spatial heterogeneity and dependence. For example, the coefficient of *HRI1* or *HRI2* in the FE-OLS column is not a significant determinant of housing excess returns, while the market factor is a significant determinant. Accounting for spatial heterogeneity with the MG estimation changed the results.

Table 3 Risk Factor coefficients with different panel estimators

Variable	FE	MG	AMG
Market beta (β_1)	0.050* (0.0297)	0.035 (0.0258)	0.001 (0.0284)
Credit Spread (β_2)	0.303 (0.2921)	0.700** (0.2058)	-0.057 (0.0284)
HRI1 (φ_1)	-0.005 (0.0035)	-0.028*** (0.0074)	0.004 (0.0088)
HRI2 (φ_2)	-0.005 (0.0042)	-0.010** (0.0050)	-0.034*** (0.0064)
Common Dynamic Process (d_i)			0.955*** (0.1175)
Abnormal loss/gain (α_i)	-0.015** (0.0071)	-0.044*** (0.0077)	-0.058*** (0.0119)
RMSE	0.0494	0.0465	0.0421
Slope homogeneity (p value)	0.011		

$$(r_i - r_f)_t = \alpha_i + \beta_{1i}(r_m - r_f)_t + \beta_{2i}Crspread_t + \varphi_{1i}HRI1_{it} + \varphi_{2i}HRI2_{it} + d_i\hat{\mu}_t^* + \varepsilon_{it}$$

where excess return $(r_i - r_f)$ is the dependent variable; i is the respective postcode, and t is the quarterly period. r is the quarterly log-returns of houses considering the median price change between two quarters. r_f is the three months quarterly Treasury bill rate for each period. Equity market premium $(r_m - r_f)$ and $Crspread$ are the vectors of observable systematic risk factors. β_{1i} , β_{2i} , φ_{1i} and φ_{2i} are the postcode-specific risk factor loadings of the observable risk factors. The postcode-specific coefficient d_i represents the implicit factor loading on the common dynamic process for period t , $\hat{\mu}_t^*$. Common dynamic process ($\hat{\mu}_t^*$), evolved from the unobservable common factors, accounts for the spatial dependence of housing returns. HRI stands for the housing risk index derived from principal component analysis. $HRI1 = w1$, 1 *STD*. $LA + w1$, 2 *STD*. $Dist + w1$, 3 *STD*. TOM ; $HRI2 = w21$ *STD*. $Crime + w22$ *STD*. RY . Respective weights (w) are given in Table 1. The weights of $HRI1$ are concentrated on the increases in land size, distance to CBD, and time on the market. The weights of $HRI2$ are concentrated on the increases in crime rate and rental yield

Though our number of cross-section dimensions (N) is less than the time dimensions (T), N is more than ten. Therefore, we employed the test proposed by Pesaran and Yamagata (2008) to test the homogeneity in large panels. The p value of the delta test statistic is significant at the 5% level and rejects the null hypothesis of homogenous slope parameters in the model, indicating that accounting for the spatial differences is necessary. Therefore, using estimators such as MG and AMG that allow for heterogeneous slopes is warranted when running this model. Root mean square error (RMSE) is smaller for the AMG estimation suggesting that it is the more efficient estimation

When we move from the FE-OLS to AMG, market factor and credit spread become insignificant while $HRI2$ becomes significant. Therefore, idiosyncratic risk factors explain the excess return variations of residential housing rather than economic factors. Allowing heterogeneity across the spatial units increases the abnormal loss represented by α_i . The coefficient on the common dynamic process, $\hat{\mu}_t^*$, under AMG is close to their theoretical value of unity and highly significant. Positive and significant common dynamic process confirms an abnormal gain due to the time-varying unobservable common shocks in the housing market. We specified models with housing-specific risk factors at a time and estimated them with all three methods. We found that time on the market, crime rate, and rental yield significantly affect housing excess returns. We also specified models with alternative risk premiums and estimated them with FE-OLS, MG, and AMG estimations. Results are the same and reported in Table 12

Individual postcode results are given in Table 11

Standard errors are given in the parentheses; ***, **, and * stand for significance at 1%, 5%, and 10% levels, respectively

There, coefficients of *HRI1*, *HRI2*, and credit spread are significant. However, ignoring spatial dependence still overestimates the economic factors such as credit spread. The AMG column shows that *HRI1* or *HRI2* is a significant determinant of housing excess returns, while market factor or credit spread is insignificant. Considering only spatial heterogeneity does not provide robust estimates. The more efficient root mean square error (RMSE) of AMG confirms that the AMG is better than the FE-OLS and MG estimation methods. The AMG estimation also provides an explicit estimate for the unobserved factors, and hence it is a better estimation method for analyzing housing risk factors and abnormal returns.

The equity market risk premium does not significantly explain the variations in housing excess returns. Therefore, equity investors have a greater potential to diversify their investments into residential housing assets. Previous studies with indexed data have found that equity market beta significantly determines the excess returns of real estate assets (Bond et al., 2003; Milcheva et al., 2020). However, in line with our findings, Case et al. (2011) show that the equity market factor cannot explain some metropolitan area-specific housing returns.

Case and Shiller (2003) state that declining interest rates increase property value due to supply inelasticity. Investments in real estate assets are exposed to refinancing risks, and hence expected real estate returns are sensitive to changes in the interest rates (Pedersen et al., 2014). The MG results indicated that credit spread has a significant positive impact on housing excess returns before we controlled for spatial dependence or the common factors. However, credit spread has no significant impact on housing excess returns according to the results of the most efficient AMG model that accounts for spatial dependence. The model will be misspecified if the impact of spatial dependence is not considered. The credit spread variable remains insignificant even when the analysis is conducted on individual postcodes (Table 11 in appendices). Aligning with our findings, Case and Shiller (1989) also confirm that information about real interest rates does not affect the price of single-family homes.

The standard asset pricing models do not account for unique risk factors such as illiquidity, indivisibility, high leverage, and the information inefficiency of residential real estate investments (Domian et al., 2015; Pedersen et al., 2014; Redmond & Cabbage, 1988). Our dynamic multi-factor panel model (Eq. 8) with AMG estimation shows that asset-specific factors determine housing excess returns in different postcodes and submarkets. Therefore, augmenting asset-specific risk factors to a standard model provides more insights into the risk-return relationship.

HRI2 significantly and negatively affects housing excess returns. An increase in crime rate decreases the excess returns according to the *HRI2* component. This negative relationship between the crime rate and the housing excess returns is consistent with the findings of Klimova and Lee (2014). Much of the literature has documented that the crime rate reduces property prices (Gibbons & Machin, 2008; Kiefer, 2011). Reducing property prices due to crime may lead to lower excess returns. The coefficient of *HRI2* also indicates that higher rental yield decreases housing excess returns. The role of rental yield as a discounting factor of real estate asset valuations supports the negative relationship between excess returns and rental yield (Davis et al., 2008; Kouwenberg & Zwinkels, 2014).

Our results suggest that residential housing returns are locally determined since the crime rate, rental yield, and postcode-specific abnormal returns are significant factors related to locale. Therefore, agreeing with Case et al. (2011), we document that housing and location-specific risk factors significantly determine residential housing returns. In addition, as shown in the result of the *Common Dynamic Process* variable, unobservable common factors such as speculative bubbles drive housing returns.

Allowing heterogeneity and spatial dependence across the postcodes increases the abnormal loss represented by α_i . However, the significant augmented common dynamic process ($\hat{\mu}_i^*$) represents an abnormal gain. The conventional interpretation of risk-return regression intercepts as abnormal return estimates break down once risk factor coefficients are allowed to differ across postcodes (Eberhardt & Teal, 2010). The significant and positive common dynamic process with a coefficient (d_i) close to unity shows that time-varying common shocks lead to abnormal gains. Commonality (cross-section correlation) may arise due to housing price bubbles across general housing markets. Since people can earn arbitrage profits during a speculative bubble (see for example, Zhang & Fan, 2019), time-varying abnormal gains can be observed. The positive relationship between the common dynamic process and excess returns also demonstrates the role of speculative forces in determining housing prices as Case et al. (2011) argued.

As the common dynamic process evolves through a time dummy variable process, the excess returns are cross-sectionally dependent at a given point of time and conditional on the previous period information set (Chudik et al., 2011). As a result, the common dynamic process represents a shock that affects the excess returns of residential housing at all the postcodes at a particular time. The common dynamic process is still significant after controlling the residential housing market premium and global equity market premium instead of the Australian equity market premium (see Tables 12 and 13 in the appendix). Therefore, we demonstrate that the common dynamic process in our model represents abnormal returns due to speculative bubbles rather than fundamental factors (Costello et al., 2011; DeFusco et al., 2018). Different factor loadings at each postcode show that local factors can change the degree of response to common speculative bubbles.

Forces affecting houses in one region create signals that eventually lead to systematic but non-uniform consequences on houses in other regions (Bourassa et al., 2007; Meen 1999). This process is called the price spill-over effect (Meen, 1999). Consistent with the housing literature (e.g., Meen, 1999; Pijnenburg, 2017), the spatial dependence of housing prices or returns represents the price spill-over effect. Speculation caused by these amplification mechanisms leads to higher house prices, disadvantaging lower-income groups and first home buyers (Case & Shiller, 2003). Therefore, house price spill-over motivated by speculative forces may generate more pressure on the housing affordability problem.

The residential property price index in Australia has grown rapidly at a compound growth rate of 4.8% from 2012 to 2020 despite the short downturn from March 2018 to June 2019.⁹ However, from 2012 to 2020, the economic growth of Australia has been slowing. Moreover, during the Covid-19 pandemic, the demand for regional

⁹ <https://www.abs.gov.au/statistics/economy/price-indexes-and-inflation/residential-property-price-index-es-eight-capital-cities/latest-release#key-statistics>

houses in Australia has increased due to sheer affordability, and the regional house prices have grown substantially (Lawless, 2021). Notwithstanding the pandemic, Australian housing prices have rapidly increased and recorded the highest value during the first quarter of 2021 (CoreLogic, 2021). Therefore, policymakers may carefully analyze the speculative forces and fundamental factors to identify the possibility of a housing bubble (Case & Shiller, 2003). Since housing prices are continually increasing over a long period, new policies may be required to control the housing affordability crisis in Australia. The following section analyzes the risks and returns relationships in submarkets based on house prices to provide more insights for policymakers and investors.

Submarket Analysis

Table 4 presents the average returns, risk, and risk-adjusted returns of residential housing assets in each postcode and submarket. Most affordable submarkets: Q1 and Q2 have lower risk and higher risk-adjusted returns than the expensive submarkets: Q3 and Q4. We statistically test the differences of means relating to each performance between affordable and expensive submarkets. The sample *t* statistics show that standard deviations and Sharpe ratios of affordable submarkets are significantly different from those of expensive submarkets. The difference between mean returns is not significant. Hence, we argue that the lower risk drives the higher risk-adjusted returns in more affordable submarkets. In line with our findings, Bangura & Lee (2020) show that house prices in affordable regions are less volatile than in expensive regions.

We next analyze the risk and return of each submarket using Eq. (8) described above. Results in Table 5 show that risk exposures are different among the submarkets. Neither market premium nor credit spread has an impact on the residential housing excess returns in any submarket, as shown in our previous results relating to the overall market. However, idiosyncratic risk factors have a significant role in determining the residential housing excess returns in all the submarkets. *HRI1* significantly affects only the excess returns of the affordable submarket (Q2). *HRI2* has a significant impact on the excess returns of all submarkets. The expensive submarket (Q3) is more exposed to housing-specific risk factors, as the coefficient of *HRI2* is larger. The most affordable submarket (Q1) has the least exposure to the risk factors (See Table 1 for the factor contributions of each *HRI* under each submarket).

Increases in the land size, distance to CBD, and time on the market improve the excess returns in the affordable submarket Q2. Houses away from the CBD have more land leverage, leading to higher returns (Baker & Filbeck, 2013; Davis & Heathcote, 2007). Recent regional house price growth in Australia also shows that demand for affordable houses away from the cities increases (CoreLogic, 2021). The positive relationship between the time on the market and affordable housing excess returns indicates that affordable houses away from CBD and with a larger land area take a longer time to sell. However, the return compensates for the longer time on the market (Anglin et al., 2003; Cheng et al., 2008).

Table 4 Average returns, risk and risk-adjusted returns at different submarkets

Postcode	Name	Average Return				Standard Deviation				Sharpe ratio			
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
4078	Forest Lake	0.06%	0.06%	0.21%	0.30%	11.50%	3.00%	2.86%	3.96%	0.00	-0.002	0.052	0.061
4301	Redbank	0.55%	0.20%	0.23%	0.17%	3.72%	3.40%	3.00%	6.20%	0.132	0.041	0.054	0.017
4503	Killgarrah	0.35%	0.34%	0.22%	0.36%	4.52%	4.21%	3.80%	5.44%	0.064	0.066	0.041	0.054
4500	Strathpine	0.25%	0.27%	0.09%	0.30%	4.94%	5.86%	8.02%	14.05%	0.038	0.036	0.004	0.017
4509	North Lakes	0.22%	0.14%	0.16%	0.34%	3.31%	3.01%	3.54%	4.36%	0.049	0.026	0.026	0.060
4208	Ormeau	0.26%	0.21%	0.20%	0.27%	5.54%	4.80%	7.53%	8.55%	0.036	0.031	0.018	0.025
4209	Coopers	0.09%	0.24%	0.27%	0.28%	2.39%	2.54%	4.54%	8.39%	0.011	0.009	0.046	0.027
4510	Calcutture	0.69%	0.56%	0.40%	0.98%	5.05%	5.24%	10.58%	16.20%	0.124	0.096	0.032	0.057
4211	Nerang	0.45%	0.50%	0.45%	0.56%	3.55%	2.91%	3.80%	6.49%	0.110	0.149	0.103	0.076
4570	Gympie	-0.16%	0.00%	0.02%	0.07%	6.27%	5.82%	8.60%	12.14%	-0.036	-0.010	-0.005	0.001
4053	Stuffed	0.56%	0.99%	0.60%	0.50%	3.11%	2.56%	3.01%	3.20%	0.162	0.204	0.179	0.138
4116	Clamvale	0.42%	0.39%	0.27%	0.30%	4.65%	4.27%	4.61%	5.24%	0.077	0.076	0.046	0.046
4152	Carina	0.57%	0.50%	0.43%	0.41%	5.29%	4.81%	4.16%	8.22%	0.095	0.092	0.089	0.043
4077	Indala	0.79%	0.58%	0.76%	0.70%	4.38%	5.53%	7.55%	12.65%	0.167	0.093	0.092	0.080
4214	Gold coast	0.65%	0.43%	0.51%	0.64%	5.51%	4.97%	6.93%	8.68%	0.107	0.074	0.065	0.067
4560	Monville	0.27%	0.25%	0.16%	0.05%	5.36%	6.42%	5.44%	15.50%	0.038	0.029	0.018	-0.001
4213	Austinsville	0.62%	0.58%	0.78%	0.51%	5.01%	3.79%	7.09%	7.39%	0.111	0.137	0.101	0.061
4551	Caloundra	0.32%	0.52%	0.65%	0.35%	3.80%	3.52%	3.54%	9.34%	0.068	0.130	0.166	0.030
4226	Robina	0.56%	0.61%	0.74%	0.81%	4.11%	4.95%	7.13%	9.90%	0.121	0.112	0.095	0.077
4109	Sunny bank	0.70%	0.42%	0.47%	0.37%	7.20%	4.85%	4.83%	7.33%	0.088	0.074	0.084	0.042
Average		0.41%	0.37%	0.38%	0.41%	4.96%	4.32%	5.53%	8.07%	0.078	0.076	0.065	0.047
t-statistic		0.1379				4.1023				-1.5963			

$Sharpe\ ratio = \frac{(r_t - r_f)}{\sigma}$. Where, r_t is the average quarterly return of the i^{th} postcode at time t , r_f is the risk-free rate (quarterly adjusted three months Treasury bill rate) at time t , σ is the standard deviation of the quarterly returns. Postcode is a spatial unit that has three to five suburban areas. This study covers 20 postcodes in Brisbane city of Australia. Those 20 postcodes represent 34% of the population and 32% of the private dwellings in Brisbane. Average return, Standard deviation, and Sharpe ratio are performance measures of the residential housing assets calculated for 39 quarterly periods. Q1-Q4 are the price quartiles of the house prices in each postcode. Q1 and Q2 are considered affordable submarkets. Q3 and Q4 are considered expensive submarkets. We statistically tested the differences of means relating to each performance between affordable and expensive submarkets. Performance measures of Q1 and Q2, as well as Q3 and Q4, were grouped together to test the differences of means. The sample t statistics show that standard deviations and Sharpe ratios of affordable submarkets are significantly different from those of expensive submarkets. The difference between mean returns is not significant. Standard deviations of affordable submarkets are lower than the expensive ones. The risk-adjusted returns (Sharpe ratio) of affordable submarkets are higher than the expensive ones

Table 5 Multi-factor risk-return model with AMG estimation under each submarket

Variable	Q1	Q2	Q3	Q4
Market beta (β_1)	0.001 (0.0339)	0.004 (0.0253)	-0.003 (0.0215)	-0.007 (0.0409)
Credit Spread (β_2)	-0.254 (0.2410)	-0.197 (0.1844)	-0.116 (0.2324)	0.104 (0.4243)
HRI1 (φ_1)	-0.002 (0.0080)	0.041*** (0.0111)	0.005 (0.0065)	0.045 (0.0981)
HRI2 (φ_2)	-0.012** (0.0042)	-0.026*** (0.0041)	-0.040*** (0.0066)	-0.016* (0.0106)
Common Dynamic Process (d_t)	0.974*** (0.1292)	0.960*** (0.0632)	0.894*** (0.0938)	1.026*** (0.1834)
Abnormal loss/gain (α_t)	-0.019** (0.0095)	-0.040*** (0.0090)	-0.059*** (0.0128)	-0.061 (0.0727)

$$(r_i - r_f)_t = \alpha_t + \beta_{1t}(r_m - r_f)_t + \beta_{2t}Crspread_t + \varphi_{1t}HRI1_{it} + \varphi_{2t}HRI2_{it} + d_t\hat{\mu}_t + \varepsilon_{it}$$

The above risk-return model was estimated for house price quartiles and the overall sample. Q1 to Q4 are price quartiles. Q1 is the most affordable submarket, and Q4 is the most expensive submarket. Dependent variable $(r_i - r_f)$ is housing excess return. The quarterly Australian three months Treasury bill rate is used as the risk-free rate to calculate the excess returns. Independent variables are market risk premium, credit spread, housing risk index one, index two, and common dynamic process. Market risk premium and credit spread are the same for each postcode and are considered as common fundamental factors for all postcodes. They have no significant impact on the excess returns

Housing-specific risk factor *HRI2* affects the excess returns of all the submarkets negatively. *HRI2* affects the excess returns of Quartile 2, an affordable submarket, positively. Quartile 1, the most affordable submarket, has the lowest risk exposure. Quartile 3, an expensive market, has the highest risk exposure. We estimated the model with the alternative market risk premiums under each submarket (Table 13). Results remain unchanged. The abnormal gains due to the time-varying common shocks are significant in all the submarkets. Abnormal loss due to the postcode-specific factors increases from Quartile 1 to 3, and it is absent in Quartile 4

Individual postcode results were not reported. Results are available on request

Standard errors are given in the parentheses; ***, **, and * stand for significance at 1%, 5%, and 10% levels, respectively

Increases in crime rate and rental yield reduce the excess returns in all submarkets. Abnormal losses due to postcode-specific factors are significant in all the submarkets except for Q4. The highest abnormal loss is in the Q3 submarket, and the smallest abnormal loss is in Q1. On the other hand, abnormal gains due to the time-varying common shocks are significant in all the submarkets, showing that unobservable factors such as speculative forces play a major role in all the submarkets.

Cannon et al. (2006) studied the relationship between housing return and risk across different price cohorts. Building on their concept, our study finds that risk and risk-adjusted returns are significantly different across the price submarkets. Risk factor loadings also vary from affordable submarkets to expensive submarkets. More affordable submarkets are favorable to investors as their exposure to risk factors is lower, and risk-adjusted returns are higher than the expensive submarkets.

Diagnostic Testing and Robustness of the Model

We substituted the S&P 500 index¹⁰ and Australian residential housing index returns for the market return in Eq. (8). The S&P 500 index represents the global market factor, and the residential housing market index represents the general housing market factor in Australia. The results are given in the appendix (Tables 12 and 13). They are similar to the primary results in Tables 3 and 5. We also specified the models by including a risk factor at a time. There we used the variables themselves instead of the index values. We found that rental yield, crime rate, and time on the market are significant factors¹¹ affecting housing excess returns. Then we added variables one by one to Eq. (8) and estimated them with AMG. There were no changes in our findings. When we kept adding housing-specific factors, economic factors became insignificant. Principal component analysis overcame the multicollinearity problems. We derived the covariance matrices of coefficients after estimation (Table 14). There was no significant covariance observed among the independent variables. We also investigated the density distributions for the residuals of the estimated Eq. (8) at all price levels. The plots indicate that the distribution of these parameter estimates is symmetric around their respective means and roughly Gaussian. Hence, no significant outliers drive our results.

There are spatial models that account for spatial correlations among the error components like the one discussed by Kapoor et al. (2007) and among the dependent variable lags like the one discussed by LeSage and Pace (2009). To check if spatial models can be applied, we conducted diagnostic tests. The results deny the applicability of spatial models to our sample suggesting that the known spatial weight matrix (distance weight matrix) is insufficient to account for the spatial dependence due to the unobservable common factors. AMG estimation accounts well for the unobservable factors by employing the common dynamic process ($\hat{\mu}_t^*$) as explained in Eq. (5)–(6). To further control the impact of spatial factors, we included a new variable consisting of an average distance between the relevant postcode and other postcodes in our sample (Table 15). The conclusion remained unchanged.

Conclusion

This study examined residential housing risk and return in different spatial units and submarkets in Greater Brisbane, Australia, using postcode and transaction-level housing data from 2010 to 2019. First, we estimated a multifactor risk-return model with Augmented Mean Group (AMG) estimation. The AMG estimation is a robust econometric technique that accounts for spatial heterogeneity and dependence. To

¹⁰ Global market premium was calculated by deducting the three months Treasury bill rate of USA from S&P 500 index returns. Respective housing excess returns were also calculated by deducting the USA Treasury bill rate.

¹¹ Rental yield, crime rate and time on the market have negative impacts on the housing excess returns. In the housing risk index 2 (HRI 2) crime rate has the largest weight.

the best of our knowledge, we are the first to apply the AMG technique in a risk-return model to estimate abnormal loss/gain due to the postcode-specific and unobservable common factors. Our analysis showed that the postcode-specific factors result in abnormal losses, while the unobservable common factors lead to abnormal gains. Furthermore, controlling for fundamental factors showed the role of speculative forces in determining housing excess returns. Our findings indicate that ignoring spatial dependence in a risk-return model overestimates the market risks and underestimates the idiosyncratic risks.

The results also reveal that macro-level systematic risk factors are not significant determinants of private residential housing returns. Therefore, rationalizing the idiosyncratic risk, we concluded that asset-specific risk factors such as time on the market, distance to CBD, land size, rental yield, and crime rate are priced in the private residential housing market in Brisbane.

Second, we analyzed the housing risks and returns in four price cohorts. The houses in more affordable submarkets were found to have lower return volatility and higher risk-adjusted returns. The most affordable submarket has the lowest risk exposure, and the expensive submarket, Q3, has the highest risk exposure. Accounting for spatial dependence and heterogeneity yields more accurate estimates for risk factor coefficients and abnormal returns in all the submarkets. Subsequently, investors and policymakers can make more effective decision-making. Instead of a single market, they may see the national housing market as a collection of submarkets and spatial units such as postcodes and suburbs. In particular, since our findings indicate that the excess return is determined largely by idiosyncratic risk factors, investors should analyze the housing market within each spatial unit. In addition, our results indicate that the affordable housing submarket has lower risk exposure than expensive submarkets. This result encourages investors to include affordable houses in their portfolios.

Furthermore, our sample covers three local government areas (LGAs) with their own city councils (i.e., Brisbane, the Gold Coast, and the Sunshine Coast). All these councils formulate their policies in line with the Southeast Queensland regional plan,¹² and providing affordable houses is one of the key issues for these councils. Our findings encourage these cities to work closely with each other to overcome such important issues. For example, one study¹³ indicated that a large number of individuals who work in the Gold Coast CBD are unable to live in the city area as the house prices are unaffordable. Movements of people from unaffordable to affordable areas may also increase the house prices in the affordable areas. The spatial dependence of housing returns (captured by unobservable common factors) among the three city council areas found in this study urges the importance of these councils cooperating with one another to improve the effectiveness of their housing policy initiatives regarding housing affordability. Future studies that will employ

¹² <https://cabinet.qld.gov.au/documents/2009/jul/seq%20regional%20plan%202009-31/attachments/seq%20regional%20plan%202009-31.pdf>

¹³ <https://cabinet.qld.gov.au/documents/2011/sep/northern%20gold%20coast%20strategy/Attachments/Att%201%20-%20strategy%202011%E2%80%932021.PDF>

AMG estimation may consider policy coordination among neighboring cities, states, and countries to manage cross-sectional dependencies.

Our analysis also showed that there is a significant spatial dependency of house price movements in the affordable submarkets. This indicates that the deterioration of housing affordability in one area may make houses in other areas unaffordable. Our results are consistent with those of many other studies on Australia that identified a spillover effect among house prices within the country (see for examples, Bangura & Lee, 2020; Costello et al., 2011; Wang & Croucher, 2021). Related to this, there are interstate spillover effects in Australia driven by speculative activities (see for example, Costello et al., 2011), with Sydney and Melbourne being the two major contributors to the house price spillovers according to Luo et al. (2007). As of 2020, the greater Brisbane area is relatively more affordable than the major cities of Sydney and Melbourne; however, its relative affordability means that Brisbane may become an attractive destination for interstate and foreign migrants. This spillover effect between the Australian states and the spatial dependency that we observed among the postcodes reveal an emerging housing affordability crisis in Greater Brisbane.

The significant common dynamic process implies that buyers may also penetrate the affordable market segment during speculative bubbles. When this happens, house prices in the affordable housing submarkets are likely to be unaffordable for disadvantaged buyers. Government and policymakers may need to introduce new capital gains taxes on the affordable and regional housing submarkets to reduce pressure from speculative forces. The government may also rationalize demand-side subsidies to control increasing house prices in affordable and regional submarkets.

Our study contributes to the literature in several ways. First, no previous study has concentrated on the return variation of residential housing by applying a panel estimator that permits heterogeneous slope coefficients and spatial dependence across different locations. Second, the submarket analysis provides insights for housing investors and policymakers regarding risk-return trade-offs in different price cohorts to take appropriate actions to maintain the affordability of houses. Thus, this study helps economically vulnerable people in society. Third, we are the first to discuss the abnormal returns due to time-varying unobservable common factors and location fixed effects. Finally, this is also the first study in Australia to investigate housing risk and returns in different postcodes and price submarkets. Therefore, our study can serve as a basis for developing more robust asset pricing models applicable to the Australian real estate market.

Due to the data constraint, our study sample is limited to 20 postcodes in Greater Brisbane. Future studies may replicate our model in other geographical locations using transaction-level data and analyze the impact of more variables such as transaction costs, governance, and environmental factors on the housing excess returns.

Appendix

Table 6 Correlation coefficients of variables in the model

Variable	Real excess returns	Real credit spread	Real market premium	Time on the market	Distance to CBD	Land area	Previous period crime rate	Real rental yield	Housing Risk Index 1	Housing Risk Index 2
Real excess returns	1									
Real credit spread	0.0261	1								
Real market premium	0.0652	0.1357**	1							
Time on the market	-0.1479**	0.1225**	-0.1055**	1						
Distance to CBD	-0.0213	-0.0001	-0.0013	0.4377**	1					
Land area	0.0382	-0.0297	0.0075	0.2475**	0.5133**	1				
Previous period crime rate	-0.0266	0.0361	-0.0144	0.2810**	0.1817**	-0.1104**	1			
Real rental yield	0.0309	0.5269*	0.1909**	0.0501	0.2118**	0.1055**	0.1884**	1		
Housing Risk Index 1	-0.0356	0.0223	-0.0289	0.6315**	0.8561**	0.8162**	0.1144**	0.1668**	1	
Housing Risk Index 2	-0.0087	0.2559**	0.07	0.2536**	0.2410**	-0.0459	0.9069**	0.5847**	0.1661**	1

There are significant correlations among the variables. Therefore, we employed a principal component analysis to overcome the multicollinearity issue

***, **, and * stand for significance at 1%, 5%, and 10% levels, respectively

Table 7 Test of equality of means and covariance matrices across postcodes

Test	F	p value
Wilks' lambda	253.97	0.000
Pillai's trace	48.26	0.000
Hotelling trace	8712.33	0.000
Roy's largest root	43,653.9	0.000
Modified LR chi ²	13.48	0.000

p-values of Wilks' lambda, Pillai's trace, Hotelling trace, and Roy's largest root tests are significant at 1% level suggesting the means of the variables are significantly different across the postcodes. The significance of Modified LR chi² shows that the covariances of the variables are different across the postcodes. Differences of menace and variances indicate a considerable potential heterogeneity of risk factor coefficients across the postcodes

Table 8 The bias-adjusted Lagrange Multiplier (LM) cross sectional dependency test

CSD test	Q1	Q2	Q3	Q4	Overall
LM	353.600 (0.000)	358.000 (0.000)	234.400 (0.016)	233.000 (0.018)	275.700 (0.000)
LM Adjusted	23.680 (0.000)	24.410 (0.000)	5.469 (0.000)	4.150 (0.000)	10.760 (0.000)
LM CD	7.757 (0.000)	12.010 (0.000)	4.833 (0.000)	2.154 (0.031)	6.193 (0.000)

(Note) $(r_i - r_p)_t = \beta \{ (r_m - r_p)_t + Crspread_t + HRI1_{it} + HRI2_{it} \}$

Q1- Q4 are housing submarkets based on the housing price quartiles. The dependent variable is real housing excess return. Independent variables are real market risk premium $(r_m - r_p)$, real credit spread $(Crspread_t)$, housing risk index1 (*HRI1*) and housing risk index 2 (*HRI2*). The cross-sectional dependence (CD) of the above equation at each submarket was tested with the Lagrange multiplier (LM) test (Breusch & Pagan, 1980), bias-adjusted LM test (Pesaran, Ullah, & Yamagata, 2008), and cross-sectional dependence test (Pesaran, 2021). All the models at different levels of affordability and overall sample are cross-sectionally dependent. We have 20 postcodes as cross-sections and 39 quarterly periods in our panel data set. *p*-values are given in the parentheses

Table 9 Stationarity tests

Variable	Test Statistic	Critical Value at 1%	Critical Value at 5%
Excess Return Q1	-6.303	-2.71	-2.85
Excess Return Q2	-6.337	-2.71	-2.85
Excess Return Q3	-6.389	-2.71	-2.85
Excess Return Q4	-6.341	-2.71	-2.85
Excess Return (Overall)	-6.411	-2.71	-2.85
HRI1Q1	-5.495	-2.71	-2.85
HRI2Q1	-4.439	-2.71	-2.85
HRI1Q2	-5.986	-2.71	-2.85
HRI2Q2	-4.356	-2.71	-2.85
HRI1Q3	-5.166	-2.71	-2.85
HRI2Q3	-4.402	-2.71	-2.85
HRI1Q4	-5.556	-2.71	-2.85
HRI2Q4	-4.369	-2.71	-2.85
HRI1 Overall	-4.754	-2.71	-2.85
HRI2 Overall	-4.439	-2.71	-2.85
Credit spread	-3.918	-1.98	-1.85
Equity Market premium ($r_m - r_f$)	-18.371	-1.98	-1.85

Due to the cross-sectional dependency, the cross-sectionally augmented Im-Pesaran-Shin (CIPS) panel unit root test (Pesaran, 2007) was employed to analyze the stationarity of housing excess returns and risk indices. Im-Pesaran-Shin (IPS) panel unit root test was employed for credit spread and equity market premium since they are common to each cross-section. All the variables are stationary at levels

Table 10 LM Bootstrap cointegration test

Dependent Variable	z values and robust p values			
	Gt	Ga	Pt	Pa
Excess Return Q1	-4.211 (0.000)	-19.709 (0.000)	-18.662 (0.000)	-20.467 (0.000)
Excess Return Q2	-4.067 (0.000)	-16.233 (0.000)	-17.334 (0.000)	-16.308 (0.000)
Excess Return Q3	-4.878 (0.000)	-20.521 (0.000)	-21.330 (0.000)	-20.166 (0.000)
Excess Return Q4	-4.701 (0.000)	-24.663 (0.000)	-20.994 (0.000)	-26.178 (0.000)
Excess Return Overall	-5.034 (0.000)	-20.908 (0.000)	-18.632 (0.000)	-20.600 (0.000)

(Note) $(r_i - r_f)_t = \beta \{ (r_m - r_f)_t + Crspread_t + HRI1_{it} + HRI2_{it} \}$

Though the variables in Eq. (4) are stationary, the common dynamic process is assumed non-stationary. Therefore, we employed the LM Bootstrap test of Westerlund and Edgerton (2007) as a second-generation panel cointegration test for all the models in each submarket. Since robust p -values of the group (Gt, Ga) and panel (Pt, Pa) mean tests are close to zero, the hypothesis of no cointegration is rejected for all submarkets. There is strong evidence of cointegration in all the models. p values are in the parentheses

Table 11 Postcode results for Eq. 8

Postcode	Market beta (β_1)	Credit Spread	<i>HRI1</i>	<i>HRI2</i>	Common Dynamic Process	Abnormal loss/gain (α_i)
4053	0.0680	-1.0448	-0.0045	-0.0025	0.6891**	-0.0166
4077	0.1757	0.5269	-0.0592	-0.0246	0.5396	-0.0620
4078	-0.0601	-1.0549	0.0290	-0.0079	0.7589**	0.0144
4109	0.0244	-1.0941	0.0286	-0.0474**	1.2328**	-0.0333
4116	0.0723	1.0412	-0.0539	-0.0043	-0.2538	-0.0518
4152	0.0102	-0.9624	0.0354	-0.0457*	0.9957**	-0.0456
4208	0.0630	-1.6990	0.0180	-0.0099	1.0724**	-0.0179
4209	0.0050	0.9430	-0.0143	-0.0173	0.4141	-0.0515*
4211	0.0547	1.6506	-0.0437*	-0.0484**	0.5439*	-0.1010**
4213	0.1160	0.8518	0.0115	-0.0463**	1.2010***	-0.1231**
4214	-0.0456	-0.3529	0.0045	-0.0088	0.5847	-0.0110
4226	-0.0296	0.3775	0.0505*	-0.0598**	1.364639**	-0.0040
4301	0.0417	-0.8710	0.0075	-0.0156	0.7066**	-0.0009
4500	0.1589	0.8898	-0.0914**	-0.1169**	0.8746	-0.1617*
4503	-0.0989	-0.6502	0.0336*	-0.0598**	1.6073***	-0.0582**
4509	-0.0670	0.3683	-0.0083	-0.0274	0.5600	-0.0573
4510	0.0818	-2.0773	0.0486**	-0.0239	2.0668***	-0.0823*
4551	-0.0945	-0.5573	0.0451	-0.0366**	1.1292**	-0.0405
4560	-0.0417	-0.1612	0.0291	-0.0076	1.3465**	-0.0635*
4570	-0.4240	2.7458	0.0070	-0.0678**	1.6734**	-0.1920**

$$(r_i - r_f)_t = \alpha_i + \beta_{1i}(r_m - r_f)_t + \beta_{2i}Crspread_t + \varphi_{1i}HRI1_{it} + \varphi_{2i}HRI2_{it} + d_i\hat{\mu}_t^* + \varepsilon_{it}$$

where the excess return $(r_i - r_f)_t$ is the dependent variable; i is the respective postcode, and t is the quarterly period. r is the quarterly log-returns of houses considering the median price change between two quarters. r_f is the three months quarterly Treasury bill rate for each period. Equity market premium $(r_m - r_f)$ and $Crspread$ are the vectors of observable systematic risk factors. β_{1i} , β_{2i} , φ_{1i} and φ_{2i} are the postcode-specific risk factor loadings of the observable risk factors. The postcode-specific coefficient represents the implicit factor loading on the common dynamic process for period t . The common dynamic process ($\hat{\mu}_t^*$) evolved from the unobservable common factors accounts for the spatial dependence of housing returns. *HRI1* stands for the housing risk index derived from principal component analysis. *HRI1* = $w1$, 1 *STD*. *LA* + $w1$, 2 *STD*. *Dist* + $w1$, 3 *STD*. *TOM*; *HRI2* = $w21$ *STD*. *Crime* + $w22$ *STD*. *RY*. Respective weights (w) are given in Table 1. The weights of *HRI1* are concentrated on the increases in land size, distance to CBD, and time on the market. The weights of *HRI2* are concentrated on the increases in crime rate and rental yield

Each postcode responds differently to the risk factors. Abnormal loss and gain are also different across the postcodes. Both *HRI1* and *HRI2* affect the excess returns in postcodes 4211, 4226, 4500, and 4503. These postcodes are away from the main CBD, Brisbane. Land size, distance to CBD, and time on the market are key risk factors for them

Table 12 Model specifications with alternative market premiums

Variable	Panel 1- Australian Residential housing index			Panel 2- S&P 500		
	FE	MG	AMG	FE	MG	AMG
Market beta (β_1)	0.552*** (0.0952)	0.468*** (0.1010)	-0.0282 (0.1127)	0.050* (0.0274)	0.049 (0.0325)	0.004 (0.0334)
Credit Spread (β_2)	0.575** (0.2881)	0.823*** (0.2195)	-0.136 (0.2692)	0.371 (0.2913)	0.757*** (0.1907)	-0.042 (0.2604)
HRI1 (φ_1)	-0.001 (0.0036)	-0.015* (0.0082)	0.007 (0.0090)	-0.006* (0.0035)	-0.030*** (0.0077)	0.003 (0.0092)
HRI2 (φ_2)	-0.007* (0.0041)	-0.012** (0.0049)	-0.033*** (0.0063)	-0.004 (0.0042)	-0.010** (0.0048)	-0.034*** (0.0062)
Common Dynamic Process (d_i)			0.958*** (0.1209)			0.950*** (0.1167)
Abnormal loss/gain (α_i)	-0.020** (0.0070)	-0.038*** (0.0083)	-0.051*** (0.0119)	-0.017** (0.0071)	-0.046*** (0.0081)	-0.057*** (0.0117)
RMSE	0.049	0.045	0.042	0.051	0.046	0.042

$$(r_i - r_f)_t = \alpha_i + \beta_{1i}(r_m - r_f)_t + \beta_{2i}Crspread_t + \varphi_{1i}HRI1_{it} + \varphi_{2i}HRI2_{it} + d_i\hat{\mu}_t^* + \varepsilon_{it}$$

Table 13 Estimated models with alternative market risk premiums under each submarket

Variable	Residential Housing				S&P 500			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Market beta (β_1)	0.007 (0.1170)	0.017 (0.0900)	0.059 (0.1126)	0.007 (0.1127)	-0.003 (0.0299)	0.004 (0.0215)	0.003 (0.0395)	-0.015 (0.0609)
Credit Spread (β_2)	-0.265 (0.2547)	-0.240 (0.2016)	-0.121 (0.2434)	0.131 (0.4142)	-0.241 (0.2669)	-0.214 (0.1718)	-0.146 (0.2317)	0.055 (0.4262)
$HRI1$ (φ_1)	-0.000 (0.0078)	0.038** (0.0113)	0.006 (0.0070)	0.064 (0.0926)	-0.001 (0.0081)	0.040*** (0.0112)	0.006 (0.0068)	0.070 (0.1009)
$HRI2$ (φ_2)	-0.011** (0.0047)	-0.024*** (0.0043)	-0.039*** (0.0064)	-0.017* (0.0104)	-0.012** (0.0046)	-0.025*** (0.0039)	-0.040*** (0.0066)	-0.016* (0.0105)
Common Dynamic Process (d_t)	0.972*** (0.1565)	0.940*** (0.0865)	0.869*** (0.1102)	1.028*** (0.1859)	0.982*** (0.1210)	0.953*** (0.617)	0.904*** (0.1004)	1.028*** (0.1815)
Abnormal loss/gain (α_t)	-0.018* (0.0096)	-0.038*** (0.0090)	-0.054*** (0.0133)	-0.060 (0.0673)	-0.020* (0.0110)	-0.037*** (0.0087)	-0.058*** (0.0128)	-0.057 (0.0710)

$$(r_t - r_{f,t}) = \alpha_t + \beta_1(r_m - r_{f,t}) + \beta_2Crs_{pread_t} + \varphi_1HRI1_{it} + \varphi_2HRI2_{it} + d_t\hat{\mu}_t^* + \varepsilon_{it}$$

We substituted the S&P 500 index and Australian residential housing index returns for the market return proxy represented by the ASX index in Eq. (8). Panel 1 reports the results of the model with residential housing market risk premium, and panel 2 reports the results of the equation with the global market risk premium. The results are similar to the results in Table 3 and Table 5. The global market premium was calculated by deducting the three months Treasury bill rate of USA from S&P 500 index returns. Respective housing excess returns were also calculated by deducting the USA Treasury bill rate. Standard errors are given in the parentheses; ***, **, * and * stand for significance at 1%, 5%, and 10% levels, respectively

Table 14 The covariance matrix of coefficients of Eq. (8)

	Market Premium	Credit Spread	HRI1	HRI2	Common Dynamic Process
Market Premium	0.001				
Credit Spread	-0.002	0.072			
HRI1	-0.000	-0.001	0.000		
HRI2	0.000	-0.001	0.000	0.000	
Common Dynamic Process	-0.001	-0.008	0.001	-0.000	0.014

$$(r_i - r_f)_t = \alpha_i + \beta_{1i}(r_m - r_f)_t + \beta_{2i}Crspread_t + \varphi_{1i}HRI1_{it} + \varphi_{2i}HRI2_{it} + d_i\hat{\mu}_t^* + \varepsilon_{it}$$

There are no significant covariances among the independent variables in our Eq. (8)

Table 15 Impact of the average distance of each postcode from other postcodes

Variable	Before Adjusting the Distance	After Adjusting the Distance
Market beta (β_1)	0.001 (0.0284)	-0.001 (0.0287)
Credit Spread (β_2)	-0.057 (0.0284)	-0.063 (0.2730)
HRI1 (φ_1)	0.004 (0.0088)	0.004 (0.0088)
HRI2 (φ_2)	-0.034*** (0.0064)	-0.034*** (0.0064)
Average Distance to each postcode (φ_3)		0.003 (0.0032)
Common Dynamic Process (d_i)	0.955*** (0.1175)	0.947*** (0.1186)
Abnormal loss/gain (α_i)	-0.058*** (0.0119)	-0.059*** (0.0121)

$$(r_i - r_f)_t = \alpha_i + \beta_{1i}(r_m - r_f)_t + \beta_{2i}Crspread_t + \varphi_{1i}HRI1_{it} + \varphi_{2i}HRI2_{it} + \varphi_3 avgdist + d_i\hat{\mu}_t^* + \varepsilon_{it}$$

where excess return $(r_i - r_f)$ is the dependent variable; i is the respective postcode, and t is the quarterly period. r is the quarterly log-returns of houses considering the median price change between two quarters. r_f is the three months quarterly Treasury bill rate for each period. Equity market premium $(r_m - r_f)$ and $Crspread$ are the vectors of observable systematic risk factors. β_{1i} , β_{2i} , φ_{1i} and φ_{2i} are the postcode-specific risk factor loadings of the observable risk factors. φ_3 is the coefficient of average distance ($avgdist$) variable. The postcode-specific coefficient d_i represents the implicit factor loading on the common dynamic process for period t , $\hat{\mu}_t^*$. Common dynamic process ($\hat{\mu}_t^*$), evolved from the unobservable common factors, accounts for the spatial dependence of housing returns. HRI stands for the housing risk index derived from principal component analysis. $HRI1 = w1, 1 \text{ STD. LA} + w1, 2 \text{ STD. Dist} + w1, 3 \text{ STD. TOM}$; $HRI2 = w21 \text{ STD. Crime} + w22 \text{ STD. RY}$. Respective weights (w) are given in Table 1. The weights of $HRI1$ are concentrated on the increases in land size, distance to CBD, and time on the market. The weights of $HRI2$ are concentrated on the increases in crime rate and rental yield

The “Before adjusting the distance” column is similar to the AMG results obtained in Table 3. We have added a new variable in the distance-adjusted column. That is the average distance to each postcode ($avgdist$). The average distance variable controls the variations due to the differences in distance. The results remain unchanged after accounting for the distance variations

Acknowledgments We acknowledge Mr. Mike Myers and the team from the National Affordable Housing Consortium (NAHC), Australia, for sharing the industry expertise. The authors would also like to thank the research workshop participants at Griffith University for their valuable comments at the following conferences: Griffith NAHC symposium, Griffith AFE research seminar, and Spatial data analysis symposium. We also provide our gratitude to Prof. Richard Chung for the initial contribution given to conceptualizing the study.

Availability of Data and Material All data comes from standard sources (Australian Urban Research Infrastructure Network, Australian Bureau of Statistics, Bloomberg, and Federal Reserve Economic Data). Due to the data rights, we have not deposited the data. The processed data sheet is available upon request.

Code Availability STATA Codes for data analysis are provided for replication purposes. They are available on request.

Author Contribution **P.S. Morawakage:** Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review and editing, Visualization. **G. Earl:** Resources and Supervision. **B. Liu:** Writing - review & editing, Validation, Supervision. **E. Roca:** Methodology, Validation, Writing - review and editing, Supervision. **A. Omura:** Validation, Writing - review and editing, Supervision.

Declarations

Conflicts of Interest/Competing Interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication All authors consent for publication.

References

- Anglin, P. M., Rutherford, R., & Springer, T. M. (2003). The trade-off between the selling price of residential properties and time-on-the-market: The impact of price setting. *The Journal of Real Estate Finance and Economics*, 26(1), 95–111. <https://doi.org/10.1023/A:1021526332732>
- Ayuso, J., & Restoy, F. (2006). House prices and rents: An equilibrium asset pricing approach. *Journal of Empirical Finance*, 13(3), 371–388. <https://doi.org/10.1016/j.jempfin.2005.10.004>
- Ayuso, J., & Restoy, F. (2007). House prices and rents in Spain: Does the discount factor matter? *Journal of Housing Economics*, 16(3–4), 291–308. <https://doi.org/10.1016/j.jhe.2007.08.002>
- Baker, H. K., & Filbeck, G. (2013). *Alternative investments: Instruments, performance, benchmarks, and strategies*. Wiley.
- Bali, T. G., Cakici, N., Yan, X., & Zhang, Z. (2005). Does idiosyncratic risk really matter? *The Journal of Finance*, 60(2), 905–929. <https://doi.org/10.1111/j.1540-6261.2005.00750.x>
- Bangura, M., & Lee, C. L. (2020). House price diffusion of housing submarkets in greater Sydney. *Housing Studies*, 35(6), 1110–1141. <https://doi.org/10.1080/02673037.2019.1648772>
- Beracha, E., & Skiba, H. (2013). Findings from a cross-sectional housing risk-factor model. *The Journal of Real Estate Finance and Economics*, 47(2), 289–309. <https://doi.org/10.1007/s11146-011-9360-x>
- Blunden, H. (2016). Discourses around negative gearing of investment properties in Australia. *Housing Studies*, 31(3), 340–357. <https://doi.org/10.1080/02673037.2015.1080820>
- Boarnet, M. G. (1994). The monocentric model and employment location. *Journal of Urban Economics*, 36(1), 79–97. <https://doi.org/10.1006/juec.1994.1027>

- Bond, S. A., Karolyi, G. A., & Sanders, A. B. (2003). International real estate returns: A multifactor, multicountry approach. *Real Estate Economics*, 31(3), 481–500. <https://doi.org/10.1111/1540-6229.00074>
- Bourassa, S. C., Cantoni, E., & Hoesli, M. (2007). Spatial dependence, housing submarkets, and house price prediction. *The Journal of Real Estate Finance and Economics*, 35(2), 143–160. <https://doi.org/10.1007/s11146-007-9036-8>
- Breusch, T. S., & Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies*, 47(1), 239–253. <https://doi.org/10.2307/2297111>
- Burdizzo, T., & Sangiácomo, M. (2016). Panel time series: Review of the methodological evolution. *The Stata Journal*, 16(2), 424–442.
- Cannon, S., Miller, N. G., & Pandher, G. S. (2006). Risk and return in the US housing market: A cross-sectional asset-pricing approach. *Real Estate Economics*, 34(4), 519–552. <https://doi.org/10.1111/j.1540-6229.2006.00177.x>
- Case, K., Coter, J., & Gabriel, S. (2011). Housing risk and return: Evidence from a housing asset-pricing model. *The Journal of Portfolio Management*, 37(5), 89–109. <https://doi.org/10.3905/jpm.2011.37.5.089>
- Case, K. E., & Shiller, R. J. (1989). The efficiency of the market for single family homes. *The American Economic Review*, 79(1), 125–137.
- Case, K. E., & Shiller, R. J. (1990). Forecasting prices and excess returns in the housing market. *Real Estate Economics*, 18(3), 253–273. <https://doi.org/10.1111/1540-6229.00521>
- Case, K. E., & Shiller, R. J. (2003). Is there a bubble in the housing market? *Brookings Papers on Economic Activity*, 2003(2), 299–362. <https://doi.org/10.1353/eca.2004.0004>
- Cheng, P., Lin, Z., & Liu, Y. (2008). A model of time-on-market and real estate price under sequential search with recall. *Real Estate Economics*, 36(4), 813–843. [10.1111/j.1540-6229.2008.00231.x](https://doi.org/10.1111/j.1540-6229.2008.00231.x)
- Chun, G. H., Sa-Aadu, J., & Shilling, J. D. (2004). The role of real estate in an institutional investor's portfolio revisited. *Journal of Real Estate Finance and Economics*, 29(3), 295–320. <https://doi.org/10.1023/B:REAL.0000036675.46796.21>
- Chudik, A., Pesaran, M. H., & Tosetti, E. (2011). Weak and strong cross-section dependence and estimation of large panels. *The Econometrics Journal*, 14(1), 45–90. <https://doi.org/10.1111/j.1368-423X.2010.00330.x>
- Colwell, P. F., & Sirmans, C. F. (1978). Area, time, centrality and the value of urban land. *Land Economics*, 54(4), 514–519. <https://doi.org/10.2307/3146175>
- Colwell, P. F., & Sirmans, C. F. (1993). A comment on zoning, returns to scale, and the value of undeveloped land. *The Review of Economics and Statistics*, 75(4), 783–786. <https://doi.org/10.2307/2109409>
- CoreLogic (2021). *Australian housing values reach a new record high as values continue to rise across every broad region of the country*. Retrieved from https://www.corelogic.com.au/sites/default/files/2021-02/CoreLogic_home_value_index_Feb_2021_1_1.pdf
- Costello, G., Fraser, P., & Groenewold, N. (2011). House prices, non-fundamental components and interstate spill-overs: The Australian experience. *Journal of Banking & Finance*, 35(3), 653–669. <https://doi.org/10.1016/j.jbankfin.2010.07.035>
- Costello, G., Leishman, C., Rowley, S., & Watkins, C. (2019). Drivers of spatial change in urban housing submarkets. *The Geographical Journal*, 185(4), 432–446. <https://doi.org/10.1111/geoj.12303>
- Cox, W., & Pavletich, H. (2020). *16th Annual demographia international housing affordability survey: 2020*. Retrieved from <http://www.demographia.com/dhi2020.pdf>
- David, N. B. (2013). Predicting housing prices according to expected future interest rate. *Applied Economics*, 45(21), 3044–3048. <https://doi.org/10.1080/00036846.2012.697124>
- Davis, M. A., & Heathcote, J. (2007). The price and quantity of residential land in the United States. *Journal of Monetary Economics*, 54(8), 2595–2620. <https://doi.org/10.1016/j.jmoneco.2007.06.023>
- Davis, M. A., Lehnert, A., & Martin, R. F. (2008). The rent-price ratio for the aggregate stock of owner-occupied housing. *Review of Income and Wealth*, 54 (2), 279–284. [10.1111/j.1475-4991.2008.00274.x](https://doi.org/10.1111/j.1475-4991.2008.00274.x)
- De Hoyos, R. E., & Sarafidis, V. (2006). Testing for cross-sectional dependence in panel-data models. *The Stata Journal*, 6(4), 482–496. <https://journals.sagepub.com/doi/pdf/10.1177/1536867X0600600403>

- DeFusco, A., Ding, W., Ferreira, F., & Gyourko, J. (2018). The role of price spill-overs in the American housing boom. *Journal of Urban Economics*, 108, 72–84. <https://doi.org/10.1016/j.jue.2018.10.001>
- Domian, D., Wolf, R., & Yang, H. F. (2015). An assessment of the risk and return of residential real estate. *Managerial Finance*, 41(6), 591–599. <https://doi.org/10.1108/MF-07-2013-0195>
- Drake, L. (1995). Testing for convergence between UK regional house prices. *Regional Studies*, 29(4), 357–366. <https://doi.org/10.1080/00343409512331349023>
- Eberhardt, M., & Bond, S. (2009). *Cross-section dependence in non-stationary panel models: A novel estimator*. Retrieved from https://mpira.ub.uni-muenchen.de/17692/1/MPRA_paper_17692.Pdf.
- Eberhardt, M., & Teal, F. (2010). *Productivity analysis in global manufacturing production*. Retrieved from https://ora.ox.ac.uk/objects/uuid:ea831625-9014-40ec-abc5-516ecfbd2118/download_file?safe_filename=paper515.pdf&file_format=application%2Fpdf&type_of_work=Working+paper
- Ecker, M. D., & Isakson, H. R. (2005). A unified convex–concave model of urban land values. *Regional Science and Urban Economics*, 35(3), 265–277. <https://doi.org/10.1016/j.regsciurbeco.2004.04.003>
- Elhorst, J. P. (2012). Dynamic spatial panels: Models, methods, and inferences. *Journal of Geographical Systems*, 14, 5–28. <https://doi.org/10.1007/s10109-011-0158-4>
- Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1–22. <https://doi.org/10.1016/j.jfineco.2014.10.010>
- Farrelly, K., & Stevenson, S. (2019). The risk and return of private equity real estate funds. *Global Finance Journal*, 42. <https://doi.org/10.1016/j.gfj.2019.04.005>
- Giacomini, E., Ling, D. C., & Naranjo, A. (2015). Leverage and returns: A cross-country analysis of public real estate markets. *The Journal of Real Estate Finance and Economics*, 51(2), 125–159. <https://doi.org/10.1007/s11146-014-9489-5>
- Gibbons, S., & Machin, S. (2008). Valuing school quality, better transport, and lower crime: Evidence from house prices. *Oxford Review of Economic Policy*, 24(1), 99–119. <https://doi.org/10.1093/oxrep/grn008>
- Goetzmann, W. N., & Spiegel, M. (1997). A spatial model of housing returns and neighborhood substitutability. *The Journal of Real Estate Finance and Economics*, 14(1–2), 11–31. <https://doi.org/10.1023/A:1007755932219>
- Han, L. (2013). Understanding the puzzling risk-return relationship for housing. *The Review of Financial Studies*, 26(4), 877–928.
- Hatzvi, E., & Otto, G. (2008). Prices, rents and rational speculative bubbles in the Sydney housing market. *Economic Record*, 84(267), 405–420. <https://doi.org/10.1111/j.1475-4932.2008.00484.x>
- Ho, D. K. H., Addae-Dapaah, K., & Glascock, J. L. (2015). International direct real estate risk premiums in a multi-factor estimation model. *The Journal of Real Estate Finance and Economics*, 51(1), 52–85. <https://doi.org/10.1007/s11146-014-9474-z>
- Hudson, C., Hudson, J., & Morley, B. (2018). Differing house price linkages across UK regions: A multi-dimensional recursive ripple model. *Urban Studies*, 55(8), 1636–1654. <https://doi.org/10.1177/0042098017700804>
- Hulse, K., & Yates, J. (2017). A private rental sector paradox: Unpacking the effects of urban restructuring on housing market dynamics. *Housing Studies*, 32(3), 253–270. <https://doi.org/10.1080/02673037.2016.1194378>
- Hyun, D., & Milcheva, S. (2018). Spatial dependence in apartment transaction prices during boom and bust. *Regional Science and Urban Economics*, 68(1), 36–45. <https://doi.org/10.1016/j.regsciurbeco.2017.11.001>
- Kapoor, M., Kelejian, H. H., & Prucha, I. R. (2007). Panel data models with spatially correlated error components. *Journal of Econometrics*, 140(1), 97–130. <https://doi.org/10.1016/j.jeconom.2006.09.004>
- Karl, G., & Gareth, T. (2005). Parcel size, location and commercial land values. *Journal of Real Estate Research*, 27(3), 343–354. <https://doi.org/10.1080/10835547.2005.12091159>
- Kiefer, H. (2011). The house price determination process: Rational expectations with a spatial context. *Journal of Housing Economics*, 20(4), 249–266. <https://doi.org/10.1016/j.jhe.2011.08.002>
- Klimova, A., & Lee, A. D. (2014). Does a nearby murder affect housing prices and rents? The case of Sydney. *Economic Record*, 90, 16–40. <https://doi.org/10.1111/1475-4932.12118>
- Koulizos, P. (2011). *Top Australian suburbs: A guide for investors and homebuyers*. Wiley.
- Kouwenberg, R., & Zwinkels, R. (2014). Forecasting the US housing market. *International Journal of Forecasting*, 30(3), 415–425. <https://doi.org/10.1016/j.ijforecast.2013.12.010>

- Lam, C., & Souza, P. C. (2020). Estimation and selection of spatial weight matrix in a spatial lag model. *Journal of Business & Economic Statistics*, 38(3), 693–710. <https://doi.org/10.1080/07350015.2019.1569526>
- Lawless, T. (2021). *How have housing markets performed across Australia's in-land rural towns*. Retrieved on February 20, 2021, from <https://www.corelogic.com.au/news/how-have-housing-markets-performed-across-australias-land-rural-towns>
- Lee, J., & Robinson, P. M. (2016). Series estimation under cross-sectional dependence. *Journal of Econometrics*, 190(1), 1–17. <https://doi.org/10.1016/j.jeconom.2015.08.001>
- LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.
- Lettau, M., & Pelger, M. (2020). Estimating latent asset-pricing factors. *Journal of Econometrics*, 218(1), 1–31. <https://doi.org/10.1016/j.jeconom.2019.08.012>
- Lin, T. C., & Evans, A. W. (2000). The relationship between the price of land and size of plot when plots are small. *Land Economics*, 76(3), 386–394.
- Lin, Z., & Liu, Y. (2008). Real estate returns and risk with heterogeneous investors. *Real Estate Economics*, 36(4), 753–776. <https://doi.org/10.1111/j.1540-6229.2008.00229.x>
- Lin, Z., & Vandell, K. D. (2007). Illiquidity and pricing biases in the real estate market. *Real Estate Economics*, 35(3), 291–330. <https://doi.org/10.1111/j.1540-6229.2007.00191.x>
- Linneman, P. (1986). An empirical test of the efficiency of the housing market. *Journal of Urban Economics*, 20(2), 140–154. [https://doi.org/10.1016/0094-1190\(86\)90003-3](https://doi.org/10.1016/0094-1190(86)90003-3)
- Luo, Z. Q., Liu, C., & Picken, D. (2007). Housing price diffusion pattern of Australia's state capital cities. *International Journal of Strategic Property Management*, 11(4), 227–242. <https://doi.org/10.3846/1648715X.2007.9637571>
- Lusht, K. M. (1988). The real estate pricing puzzle. *Real Estate Economics*, 16(2), 95–104. <https://doi.org/10.1111/1540-6229.00448>
- Meen, G. (1999). Regional house prices and the ripple effect: A new interpretation. *Housing Studies*, 14(6), 733–753. <https://doi.org/10.1080/02673039982524>
- Meen, G. (2016). Spatial housing economics: A survey. *Urban Studies*, 53(10), 1987–2003. <https://doi.org/10.1177/0042098016642962>
- Melser, D., & Hill, R. J. (2019). Residential real estate, risk, return and diversification: Some empirical evidence. *The Journal of Real Estate Finance and Economics*, 59(1), 111–146. <https://doi.org/10.1007/s11146-018-9668-x>
- Milcheva, S., Yildirim, Y., & Zhu, B. (2020). Distance to headquarter and real estate equity performance. *The Journal of Real Estate Finance and Economics*, Springer online publication. <https://doi.org/10.1007/s11146-018-9668-x>
- Nneji, O., Brooks, C., & Ward, C. W. (2015). Speculative bubble spill-overs across regional housing markets. *Land Economics*, 91(3), 516–535. <https://doi.org/10.3368/le.91.3.516>
- Oikarinen, E., Bourassa, S. C., Hoesli, M., & Engblom, J. (2018). US metropolitan house price dynamics. *Journal of Urban Economics*, 105, 54–69. <https://doi.org/10.1016/j.jue.2018.03.001>
- Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4), 967–1012. <https://doi.org/10.1111/j.1468-0262.2006.00692.x>
- Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, 22(2), 265–312. <https://doi.org/10.1002/jae.951>
- Pesaran, M. H. (2021). General diagnostic tests for cross-sectional dependence in panels. *Empirical Economics*, 60, 13–50. <https://doi.org/10.1007/s00181-020-01875-7>
- Pesaran, M. H., & Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, 68(1), 79–113. [https://doi.org/10.1016/0304-4076\(94\)01644-F](https://doi.org/10.1016/0304-4076(94)01644-F)
- Pesaran, M. H., Ullah, A., & Yamagata, T. (2008). A bias-adjusted LM test of error cross-section independence. *The Econometrics Journal*, 11(1), 105–127. <https://doi.org/10.1111/j.1368-423X.2007.00227.x>
- Pesaran, M. H., & Yamagata, T. (2008). Testing slope homogeneity in large panels. *Journal of Econometrics*, 142(1), 50–93. <https://doi.org/10.1016/j.jeconom.2007.05.010>
- Pedersen, N., Page, S., & He, F. (2014). Asset allocation: Risk models for alternative investments. *Financial Analysts Journal*, 70(3), 34–45. <https://doi.org/10.2469/faj.v70.n3.4>
- Peng, L. (2016). The risk and return of commercial real estate: A property level analysis. *Real Estate Economics*, 44(3), 555–583. <https://doi.org/10.1111/1540-6229.12111>
- Persyn, D., & Westerlund, J. (2008). Error-correction-based cointegration tests for panel data. *The Stata Journal*, 8(2), 232–241.

- Pijnenburg, K. (2017). The spatial dimension of US house prices. *Urban Studies*, 54(2), 466–481. <https://doi.org/10.1177/0042098015606595>
- Preqin. (2019). 2019 Preqin Global Real Estate Report. Preqin.
- Randolph, B., & Tice, A. (2013). Who lives in higher density housing? A study of spatially discontinuous housing sub-markets in Sydney and Melbourne. *Urban Studies*, 50(13), 2661–2681. <https://doi.org/10.1177/0042098013477701>
- Redmond, C. H., & Cabbage, F. W. (1988). Portfolio risk and returns from timber asset investments. *Land Economics*, 64(4), 325–337. <https://doi.org/10.2307/3146305>
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3), 341–360. [https://doi.org/10.1016/0022-0531\(76\)90046-6](https://doi.org/10.1016/0022-0531(76)90046-6)
- Sarstedt, M., & Mooi, E. (2019). Principal component and factor analysis. In M. Sarstedt & E. Mooi (Eds), A concise guide to market research: The process data and methods using IBM, SPSS statistics (pp. 259–293). Verlag GmbH Germany: Springer.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), 425–442. <https://doi.org/10.1111/j.1540-6261.1964.tb02865.x>
- Shilling, J. D. (2003). Is there a risk premium puzzle in real estate? *Real Estate Economics*, 31(4), 501–525. <https://doi.org/10.1046/j.1080-8620.2003.00075.x>
- Simlai, P. (2018). Spatial dependence, idiosyncratic risk, and the valuation of disaggregated housing data. *The Journal of Real Estate Finance and Economics*, 57(2), 192–230. [10.1007/s11146-017-9610-7](https://doi.org/10.1007/s11146-017-9610-7)
- Sivitanides, P. S., Torto, R. G., & Wheaton, W. C. (2003). Real estate market fundamentals and asset pricing. *The Journal of Portfolio Management*, 29(5), 45–53. <https://doi.org/10.3905/jpm.2003.319905>
- Song, Y., & Sohn, J. (2007). Valuing spatial accessibility to retailing: A case study of the single family housing market in Hillsboro, Oregon. *Journal of Retailing and Consumer Services*, 14(4), 279–288. <https://doi.org/10.1016/j.jretconser.2006.07.002>
- Thorsnes, P., & McMillen, D. P. (1998). Land value and parcel size: A semiparametric analysis. *The Journal of Real Estate Finance and Economics*, 17(3), 233–244. <https://doi.org/10.1023/A:1007772223239>
- Wang, J., & Croucher, J. S. (2021). Information linkages among national, NSW, VIC, and QLD real estate markets in Australia. *Accounting & Finance*, 61(2), 3207–3234. <https://doi.org/10.1111/acfi.12698>
- Watkins, C. A. (2001). The definition and identification of housing submarkets. *Environment and Planning A: Economy and Space*, 33, 2235(12), –2253. <https://doi.org/10.1068/a34162>
- Westerlund, J., & Edgerton, D. L. (2007). A panel bootstrap cointegration test. *Economics Letters*, 97(3), 185–190. <https://doi.org/10.1016/j.econlet.2007.03.003>
- Wilson, P., White, M., Dunse, N., Cheong, C., & Zurbrugg, R. (2011). Modelling price movements in housing micro markets: Identifying long-term components in local housing market dynamics. *Urban Studies*, 48(9), 1853–1874. <https://doi.org/10.1177/0042098010380960>
- Wright, D., & Yanotti, M. B. (2019). Home advantage: The preference for local residential real estate investment. *Pacific-Basin Finance Journal*, 57, 1–21. <https://doi.org/10.1016/j.pacfin.2019.06.014>
- Xiao, Y., Webster, C., & Orford, S. (2016). Can street segments indexed for accessibility form the basis for housing submarket delineation? *Housing Studies*, 31(7), 829–851. <https://doi.org/10.1080/02673037.2016.1150433>
- Young, M. S. (2008). Revisiting non-normal real estate return distributions by property type in the US. *The Journal of Real Estate Finance and Economics*, 36(2), 233–248.
- Zhang, D., & Fan, G.-Z. (2019). Regional spill-over and rising connectedness in China's urban housing prices. *Regional Studies*, 53(6), 861–873. <https://doi.org/10.1080/00343404.2018.1490011>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.