PACIFIC RIM PROPERTY RESEARCH JOURNAL
2024, Vol.29, No.1, 57-83

De-mystifying the Housing Overhang problem in Malaysia

Hui Hon Chung[1]
Edinburgh Business School, Heriot Watt University
Email: h.hui@hw.ac.uk
ORCID: 0000-0003-0738-6535

Abstract

The problem of completed unsold housing, or overhang, has been regularly highlighted in the media in Malaysia. There was a particularly sharp increase in overhang after 2015. A large proportion of these unsold houses were in the mid- and high-price segments, suggesting that affordability could be the issue. Some media analysts pointed out that poor housing location could also explain the volume of overhang. Thus, a formal inquiry into this concern is timely and necessary. In this study, we collected and analysed a panel data of 12 states or federal territories in the country over the period 2008-2022. The main novelty of the research is that it constitutes the first panel data analysis of the overhang issue in Malaysia. Our findings from panel data models indicated that overhang was a persistent problem with past overhang contributing to the accumulation of overhang in the current period. More importantly, we presented strong evidence that deterioration in affordability contributed significantly to overhang. However, there was weak evidence supporting the view that housing location affected overhang. These results survived a series of robustness checks.

Introduction

Background of study

The stock-flow framework of housing is frequently used to study the behaviour of housing markets (Steiner, 2010). Within this framework, vacancy rate is included to capture the notion that markets do not always instantaneously clear (Gu, 2008). Struyk (1988) argued that high-income economies had lower vacancy rates than in developing economies. Data provided by OECD[2] indicate that vacancy rates in some OECD countries fall within the range of 9-16%.

Glaesser et al (2017) defined the vacancy rate as the share in housing stock of (i) completed units which are unsold by developers and (ii) unoccupied purchased units. Based on this definition, shifts in the vacancy rate can be due to either fluctuations in the number of unoccupied purchased units or changes in the number of unsold completed housing. The latter could reflect either a temporary oversupply in the market or a natural tendency in which some completed housing stock may never be sold due to stigmatisation (Huri et al, 2024)[3].

[1] The author would like to thank the two anonymous referees who provided helpful comments to improve the paper. Any further errors would remain the fault of the author.
[2] Data can be downloaded via the link: https://www.oecd.org/els/family/HM1-1-Housing-stock-and-construction.pdf
[3] This idea can be expressed as natural overhang rate, introduced by Professor Ting Kien Hwa. While research is still nascent, a lecture on this topic is accessible via the link: https://www.facebook.com/ReadingDotMy/videos/135-natural-overhang-rate-characteristics-prof-ting/579734200063109/?locale=hi_IN

In some countries, unsold and vacant housing has become a major issue. In China for instance, the problem of urbanisation and excessive housing development have resulted in vacant housing problems (Rogoff and Yang, 2020) and the emergence of 'ghost cities' (Sorace and Hurst, 2015, Gan et al, 2019, Ma, 2022). Housing bubbles leading to oversupply, were also observed in Spain (Esteban and Altuzarra, 2016), Korea (Nam et al, 2016, Yoo and Yoon, 2021) and some countries in the Mediterranean region (Gentile and Hoekstra, 2019). This problem also represents inefficiency and misallocation of resources (Durst and Ward, 2015). While housing markets usually had some vacant stock, too much vacancy can lead to economic problems (Newman et al, 2019). Thus, efforts must be made to understand how and why vacancy rates fluctuate.

Monkkonen (2019) argued that vacant housing in Mexico was mainly due to the nature of housing financing system in the country. Meanwhile, the large proportion of unsold housing in Korea was due to poor economic conditions (Lee et al, 2017), demographic factors (Yoo and Kwon, 2019), housing characteristics including location (Yoo and Yoon, 2021). Vacant housing in China was found to be negatively correlated with GDP, house prices and population size (Tan et al, 2020).

In Malaysia, unsold housing issues have been regularly highlighted by the media.[4] However, national statistics on housing vacancy rates in Malaysia is sparse[5]. The only available data on housing vacancy is found in the official 2020 census, which recorded the housing vacancy rate at 19.4% (Kunasekaran, 2023). Due to this data paucity, the closest proxy to vacancy is housing overhang – defined as the number of completed houses that remain unsold at least nine months after the official launch. This is the standard definition used by the National Property Information Center (NAPIC), the authority on housing data in the country.

Problem statement

The extent of housing overhang in Malaysia tends to be cyclical with period of expansion alternating with years of contraction (see Figure 1)[6]. In the aftermath of the 1997-98 Asian Financial Crisis and the unwinding of the mid-1990s real estate bubble, it took almost a decade to reduce the overhang. After 2010, this pattern of decline was reversed. The housing market saw consistent increases in housing overhang again, with a particularly sharp increase being experienced after 2015. Patterns of overhang in some economically significant states/federal territories (Figure 2) seemed to be contributing to the overhang at the country level.

Trends in overhang were closely connected to the demand-supply dynamics of the market. On the supply side, housing stock had been growing consistently with no dramatic changes over the last 20 years (see Figure 3). In this period, there were demographic changes due to urbanisation as more households moving to the cities. This event created new demand for housing. However, the presence of persistent housing overhang after 2015 showed that not all

[4] For instance, see: https://www.nst.com.my/business/2022/12/861264/experts-take-minimising-malaysias-property-overhang
[5] A commentary in a local real estate forum highlights this gap in the real estate data: https://www.penangpropertytalk.com/2019/03/residential-vacancy-rates-property-overhang-and-market-supply/
[6] Figures and tables are found at the end of this manuscript

new houses were taken up by buyers. This concern might reflect market inefficiencies[7]. Hence, it is crucial to understand the factors that were driving these trends in overhang.

Figure 4 shows that the bulk of housing supply are sustainably concentrated in the low and mid-priced categories in 2011. However, there was a greater concentration of housing construction in the mid- to high-priced housing categories over the decade. Figure 5 shows that there was a sustained decline in proportion of unsold housing in the lower priced segments, and an increasing proportion of unsold housing in the mid- to high-priced segments. This suggests that affordability could possibly be linked to overhang. It was pointed by a media report that overhang could also be due to poor housing location[8].

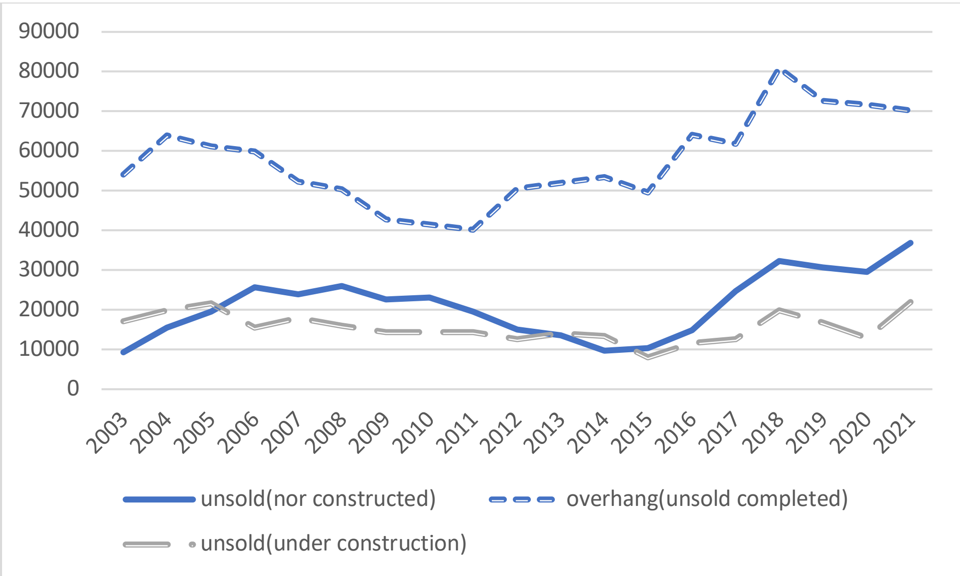

Fig. 1: Housing overhang and other categories of unsold housing (units)
Source: NAPIC, author's calculations

[7] This concern was raised by many real estate market commentators. See https://www.iproperty.com.my/news/overhang-residential-property-market-62492
[8] See report here: https://www.nst.com.my/property/2023/10/969373/demystifying-issues-behind-malaysias-overhang-unsold-residential-properties

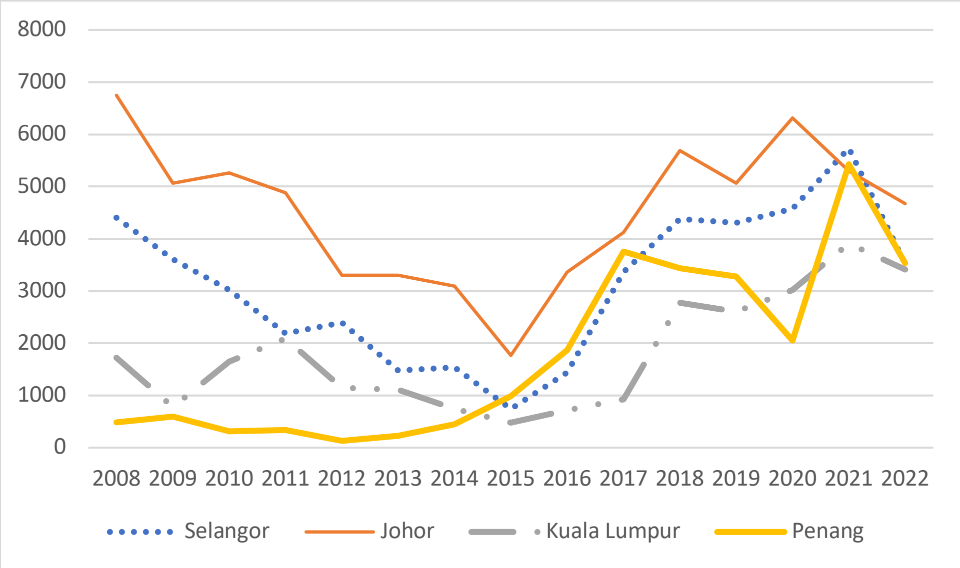

Fig. 2: Overhang in selected states/federal territory (units)
Source: NAPIC, author's calculations

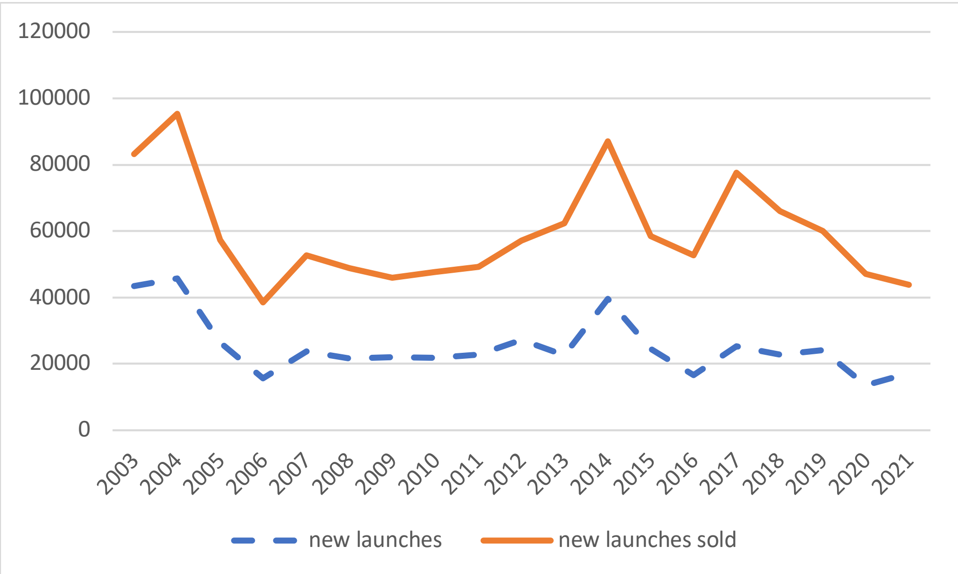

Fig 3: Newly completed houses and numbers sold
Source: NAPIC, author's calculations

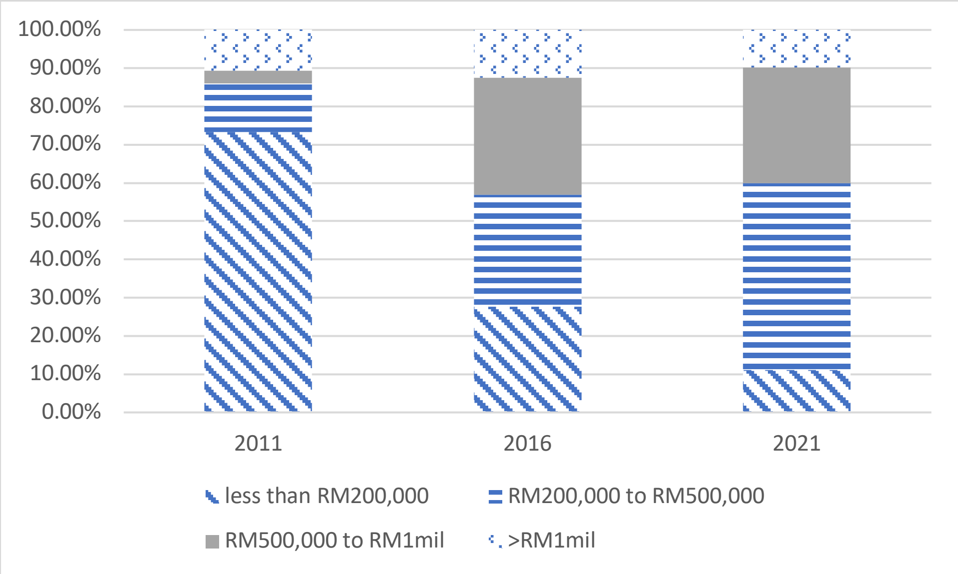

Fig. 4: % distribution of new housing launches according to price range (2011 vs 2016 vs 2021)
Source: NAPIC, author calculations, author's calculations

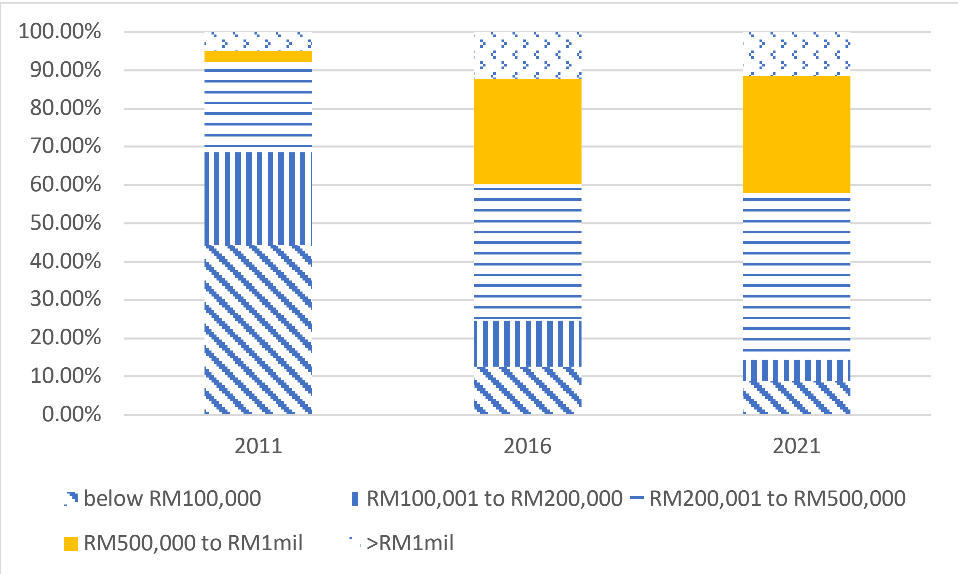

Fig 5: % distribution of housing overhang according to price range (2011 vs 2016 vs 2021)
Source: NAPIC, author calculations

There are numerous studies in the literature on overhang in Malaysia. In an exploratory study using secondary data from NAPIC, KRI (2022) assessed the spatial pattern of overhang data for Greater Kuala Lumpur area. There is no causality or correlation analysis of potential factors affecting overhang at the country level. Tey and Kassim (2022) analysed overhang problems in the state of Johor using primary data from a survey within a causality/correlation-type of statistical framework. They found price to be a significant determinant of overhang, while location factor was only weakly significant in affecting overhang. Yip et al (2020) attempted a time series study at the national level where housing oversupply was hypothesised to be affected by house price, economic growth and affordability. Studies by the

Central Bank of Malaysia (BNM, 2017a, 2017b, Cheah et al, 2017) mentioned the importance of affordability in the overhang debate but did not deploy a statistical causation/correlation approach. Aris et al (2020) is an innovative study on the spatial pattern of overhang in the state of Selangor but did not venture to propose causal factors affecting overhang.

Some gaps can be identified in the literature above. Firstly, it is possible to extend the work of Tey and Kassim (2022) by analysing causes of overhang at the national level. Second, it is also possible to extend the work of Yip et al (2020) that used time series data. In this regard, there is sufficient information from NAPIC to construct a meaningful panel dataset of overhang and its potential determinants covering the entire country. Since panel/longitudinal data covers both spatial and temporal dimensions, it could give a fresh perspective of the overhang problem. Third, as the literature review section will explain further, overhang could exhibit persistence because past overhang could contribute significantly to the current volume of overhang. The extent to which this dimension is significant, was not captured in these earlier studies.

To contribute to these gaps in literature, we collected a panel data of 12 states/federal territories in the country over the period 2008-2022 with secondary data from NAPIC. From a comprehensive review of the literature, we hypothesised that overhang as a share of housing stock, was linked to affordability and location, while controlling for other state-specific and macroeconomic factors. Using a variety of dynamic panel data regression models, our findings indicated that overhang was a persistent problem with past overhang contributing to the accumulation of more overhang in the current period. More importantly, we presented strong evidence that deterioration in affordability contributed significantly to overhang. However, there is weak evidence supporting the view that housing location affected overhang.

In terms of novelty, our study extends the current literature in two directions. Our first contribution is in the data. It is the first comprehensive panel data study of housing overhang for Malaysia, and only the second study that used panel data analysis in the literature after Yoo and Yoon (2021). The structure of a panel data adds a richer understanding to the overhang issue because both the spatial and temporal dimensions of the housing market are examined. Second, we provide empirical estimates of factors affecting overhang in a statistical model with short-run dynamics to emphasise the persistent characteristics of overhang. The study also adds to the body of knowledge on housing vacancy drivers (Newman et al, 2019, Monkkonen, 2019) by presenting analysis of a similar problem from the perspective of a developing economy in the ASEAN region.

We organise the paper in the following manner. We first explore the literature, then move on to describe the methodology, the variables and data sources. Next, we present our data analysis, key findings and discussions followed by some policy suggestions before concluding the paper.

Literature review

Vacancy, unsold housing and overhang – some technical considerations
As mentioned in the background section of the paper, this study adopts the definition of housing vacancy of Glaesser et al (2017). Vacancy rate is defined as the share in housing stock of (i) completed units which are unsold by developers and (ii) unoccupied purchased units. Thus, changes in vacancy rate can be due to fluctuations in unoccupied purchased units or/and unsold completed housing. The latter could be due to either a temporary oversupply in the market or a natural tendency for completed housing stock to never be sold due to stigmatisation (Huri et al, 2024).

In Malaysia, data on vacancy rate is sparse. However, there is rich official statistics on component (i) of the vacancy definition in Glaesser et al (2017). This component represents unsold housing stock by developers and is proxied by housing overhang. More descriptions of this concept are given in the methodology section.

Factors affecting vacancy and unsold housing – empirical literature
Monkkonen (2019) examined vacancy rates in Mexico. The author estimated Ordinary Least Squares (OLS) regression models, using cross section data for 100 large cities in Mexico. The dependent variable is measured by vacancy rates while the independent variables included housing finance and variables measuring city characteristics (e.g. include housing stock, population growth, population density, economic growth, murder rate, percentage of household members living abroad). The main findings of the study were: housing finance, murder rates and immigration out of the country correlate positively with vacancy rates.

Nam et al (2016), Lee et al (2017), Yoo and Kwon (2019) and Yoo and Yoon (2021) focus on the Korean housing market. Nam et al (2016) performed a correlation analysis between vacant housing and several factors including population composition, urbanisation rates, rate of urban decline, and building construction using district-level data. They found that vacant housing had the strongest correlations with excessive building construction and urban decline. Yoo and Yoon (2021) used a district-level, unbalanced panel data of over 60,000 observations, and regressed the proportion of unsold housing units in total housing stock on the availability of communal open spaces in a housing complex and numerous control variables. The authors concluded that availability of communal spaces in a housing complex affected unsold housing significantly.

Yoo and Kwon (2019) found that more housing construction increased vacancy rates, but how vacancy rates vary depends on the cluster they belong to. Finally, Lee et al (2017) developed a Vector Error Correction Model (VECM) with unsold new housing stock, transaction price, housing price, rental price, stock market index and interest rate, housing supply and loans as variables in a time series (monthly) dataset. Movements in unsold houses could be due to changes in transaction price, loan disbursements, interest rates and housing supply conditions. However, these relationships varied according to the time horizon.

Gu (2008) specified a vacancy model for Singapore, with vacancy rate defined as proportion of private vacant housing in total private housing stock. The vacancy rate was then regressed on its lagged values, log of real GDP, log of new private housing supply and change in housing occupancy rate. The author found that the coefficients on lagged vacancy rate and new private housing supply were statistically significant and positive, while the coefficients on log real GDP and occupancy rate were statistically significant and negative.

Saunders and Tulip (2019) estimated a model of Australian housing market, with vacancy rate as one of the equations in the model. The authors defined rental vacancy rate as the share of available rental properties in total number of rental properties. In their empirical specification, change in rental vacancy rate was regressed on a variable called excess completions, deviations of lagged vacancy rate from its steady state value and change in unemployment rate. All explanatory variables were found to be statistically significant and positive. However, there was no theoretical justification of the model structure.

In Wang and Immergluck (2019), the authors regressed log vacancy rates on its lagged value and a vector of explanatory variables grouped according to housing characteristics, neighbourhood location and socioeconomic characteristics. According to the main findings, lagged vacancy rate contributed positively to current vacancy rates, while higher unemployment tended to increase vacancy rate. However, location effects on vacancy rates were quite ambiguous. Newman et al (2019) developed models of vacancy rates for US cities, with the dependent variable specified as change in vacancy rates similar to Saunders and Tulip (2019). The main results showed that increased unemployment raised vacancy rates whereas population growth had the opposite effect.

Empirical literature for Malaysia – housing overhang problem

In Malaysia, the private sector is responsible for meeting the housing needs of society (KRI, 2015). The Sell-then-Build (STB) system is the predominant method of housing delivery (Yusof et al, 2010). Houses were built and sold before construction is completed. To finance the construction process, buyers pay a 10% downpayment of the house price. At various stages of the building process, buyers make progressive payments to developers.[9] However, there is an affordability issue due to rising prices and cost[10]. To increase home ownership, the government is targeting the construction of 500,000 units of affordable housing for the period of 2021-2025. Market analysts pointed out that the demand for affordable houses was still low due to mismatches between buyer expectations and developers' plans.[11] Thus, such affordable houses could be contributing to the housing overhang problem.[12] The disadvantages of owning such houses could deter demand.[13]

[9] More details here: https://www.iproperty.com.my/news/build-then-sell-bts-1090-housing-industry-74492
[10] Building material and compliance cost are rising: https://www.iproperty.com.my/news/how-rising-compliance-cost-is-impacting-housing-affordability-in-malaysia-55049
[11] See an example of report here: https://www.iproperty.com.my/news/why-are-so-few-people-buying-rumah-selangorku-rsku-homes-38796
[12] Report accessible here: https://www.iproperty.com.my/news/residential-property-overhang-malaysia-42611
[13] Owners may not be able to sell these houses easily. This was mentioned in the following report: https://www.propertyguru.com.my/property-guides/pros-and-cons-of-affordable-housing-schemes-in-malaysia-65373

Yip et al (2020) adopted a time series approach on aggregated country-level, in which housing glut (proxied by unsold housing) was regressed on indicators of economic growth, house price and affordability, respectively in three separate cointegrating equations. The authors' found that all the explanatory variables affected housing glut, albeit the extent of the effects were weak. Higher economics growth and affordability lowered the glut, while higher price increased glut.

KRI (2022) assessed the spatial pattern of overhang data for Greater Kuala Lumpur area using an exploratory analysis. However, there is no causality or correlation analysis of potential factors affecting overhang at the country level. Tey and Kassim (2022) analysed overhang problems in the state of Johor using primary data from a survey within a causality/correlation-type of statistical framework. The findings included the following: house price was a significant determinant of overhang, but the location factor was only weakly significant in affecting overhang. Karim et al (2017) surveyed perceptions of factors affecting overhang of double-storey terrace houses in Johor. According to their survey respondents, affordability was perceived to be a major contributor to the overhang. Numerous exploratory studies by the Central Bank of Malaysia (BNM, 2017a, 2017b, Cheah et al, 2017) cited the importance of affordability. Aris et al (2020) studied the spatial clustering pattern of overhang and affordability in the districts in the state of Selangor but did not venture to propose causal factors affecting overhang. The authors found an association between affordability and overhang in several regions.

Thanaraju et al (2019), Olanrewaju and Woon (2019) and Muhammad Zamri et al (2022) surveyed the housing preferences of buyers. According to a survey carried out by Muhammad Zamri et al (2022), financial capability to purchase a house, location and type of housing neighbourhood were factors considered by civil servants when making housing purchases. In other survey-based research by Thanaraju et al (2019) and Olanrewaju and Woon (2019), similar findings were reported. This implies that poor location and lack of affordability could lead to less demand and more unsold housing.

Since affordability was regularly mentioned in other studies on overhang, it is worthwhile to briefly comment on what affordability means. According to KRI (2015) affordability measures the cost of acquiring a house relative to an individual's payment capability. There are 3 approaches to measure affordability, namely the median multiple (ratio of median house price to median household income)[14], housing cost burden (proportion of housing cost in total household income) and residual income (proportion of leftover income after deducting housing cost).[15] Bank of Canada offers a detailed methodology to compute housing cost that takes account of interest rate on housing loans, size of loans and number of loan repayment periods.[16] In this methodology, higher interest rates and bigger loan size would lead to higher housing cost. Thus, the inclusion of variables related to housing cost (e.g. borrowing cost,

[14] Some studies inverted the ratio i.e. income to house price ratio. See Rangel et al (2023) for instance.
[15] Thresholds are provided in KRI (2015) indicating the level beyond which housing is considered unaffordable. For example, the housing cost burden must be lower than 30% for a house to be considered affordable.
[16] See methodology here: https://www.bankofcanada.ca/rates/indicators/capacity-and-inflation-pressures/real-estate-market-definitions/

home values, ease of financing) in an analysis, could lead to unnecessary duplication since these variables could be subsumed under affordability (see Yip et al, 2020).

The literature review leads to the following conclusions on housing overhang.

- Overhang is related to affordability (Cheah et al, 2017, Yip et al, 2020), demographic factors (Monkkonen, 2019), economic factors (Yoo and Kwon, 2019) and housing preferences of buyers (Yoo and Yoon, 2021), including location/distance effects (Wang and Immergluck, 2019, Tey and Kassim, 2022). In this study, the impact of affordability and location on overhang is examined. This segment justifies the inclusion of these variables into our analysis.
- Government policies may be important as well (KRI, 2015). Policies to construct more affordable housing can be counterproductive when houses do not meet buyers' preferences. We will not be examining the impact of government policies here. This could be undertaken as a future research.

Methodology, Description of Variables and Data

Methodology: Model

As mentioned in the introduction, we intend to analyse the impact of location and housing affordability on housing overhang. To implement this objective, we present a model of the following form:

$$\left(\frac{OH}{HS}\right)_{i,t} = \alpha + \sum_{j=1}^{n} \theta_j \left(\frac{OH}{HS}\right)_{i,t-j} + \beta_1 X_{i,t} + \beta_2 Afford_{i,t} + \beta_3 Location_{i,t} + e_{i,t} \quad (1)$$

Where:

OH = Housing overhang; HS = Housing Stock, Afford = affordability, proxied by the house price-to-per capita GDP ratio; Location = average travel distance, weighted by housing stock, X = a vector of control variables, including absorption rate, unemployment rate, population density and housing starts (explanation of variables made in the next section).

The dependent variable in (1) is the overhang rate. This model had a dynamic, autoregressive structure with lagged overhang to capture the short-run behaviour of overhang. The model preference was motivated by similar practices in the literature such as Gu (2008), Saunders and Tulip (2019) and Wang and Immergluck (2019) who used lagged dependent variable in their models. However, unlike Saunders and Tulip (2019) who specified variables in first differences, we chose to use variables in levels instead. This was because first-differenced variables produced models that fit the data very poorly. We maintained the variables in their original form and did not log-transform them. More discussions of the model would be found in the next section. Affordability (Cheah et al, 2017, Yip et al, 2020) and location (Tey and Kassim, 2022) are key factors identified in the literature. We added a vector of control variables that capture various demographic (Monkkonen, 2019) and economic factors (Yoo and Kwon, 2019).

Methodology: Estimation procedures

The sample has a panel data structure (data explained in the next section). As such, there are several estimation issues to be considered[17]. First, there may be time-invariant characteristics in the cross-section members (i.e. the different states/federal territories in Malaysia) in our data that cannot be observed. Using a pooled regression approach while leaving these unobservable characteristics out, would render the model estimates to be inconsistent if the unobservable characteristics are correlated with the other explanatory variables (Wooldridge, 2016). To capture these effects efficiently and consistently, the Fixed Effect (FE) model/estimators could be used, instead of the conventional pooled regression model/estimators. In contrast, if the unobservable factors could be uncorrelated to the explanatory variables. In this case the Random Effect (RE) model/estimators would be more consistent and efficient. On the other hand, another viewpoint would be to leave out both effects and assume they do not exist, resulting in a pooled regression model. In summary, we must first determine the choice of models between RE, FE and pooled regression model[18].

To decide between pooled regression vs RE, we first performed the Lagrange Multiplier (LM) tests such as the Breusch-Pagan (1980), Honda (1985), King and Wu (1997), Gourieroux, Holly and Monfort (1982) and the Standardised tests of Honda (1991) and King and Wu (Moulton and Randolph, 1989) respectively. If these tests favour the use of RE, then a choice must be made between RE vs FE. Here, we apply the test of Hausman (1978) to help make a decision.

Once model structure is identified, we estimated the baseline model using two stage least squares/instrumental variables (TSLS/IV) approach. This approach is more robust to potential endogeneity problems[19]. The instruments used were lagged values (lagged by one period) of the dependent and explanatory variables. For robustness checks, we also estimated the models using the General Method of Moments (GMM) of Arellano and Bond (1991).

Description of Variables and Data sources

This study attempts to examine factors affecting housing overhang in Malaysia. From the comprehensive review of the literature, and given data limitations, we hypothesise that the rate of housing overhang is affected by housing affordability and location characteristics, while controlling for other factors that might affect overhang such as national economic conditions, demographic factors and other regional and spatial characteristics. The dataset is a balanced panel data of 14 states/federal territories in Malaysia covering the period 2008-22. After performing an initial data assessment, we found that the states of Sabah and Sarawak were unsuitable for inclusion in our sample because they posed substantial outliers. Thus, the final data only included 12 states/federal territories over 2008-22.

Our dependent variable was the rate of housing overhang, proxied by total housing overhang as a proportion of total housing stock. By definition, overhang represents the number of houses that are completed but remain unsold at least nine months after the official launch, and this is the standard definition used by the National Property Information Center (NAPIC),

[17] This notion emerges from the nature of the panel data which had both cross-section and time series components. Pooled data involved aggregating both components into a large dataset.
[18] We did not test for stationarity, since our sample only covered 15 years. For such a short time period, the test of the order of integration would not be reliable (Wooldridge, 2016).
[19] Econometric textbooks would explain this issue in detail, see Gujarati and Porter (2009).

the authority on housing data in the country. Data for overhang and housing stock were obtained from NAPIC.

The main explanatory variables were indicators of affordability and location, whereas the other control variables were housing starts, population density, absorption rate and unemployment rate. Data for housing starts was obtained from NAPIC. Absorption rate is defined as the number of new houses sold within 3 quarters of launching as a proportion of total new launches and the data was also obtained from NAPIC. Population density and unemployment rate data were sourced from Department of Statistics, Malaysia (DOSM).

Affordability was measured by the price-to-income ratio. This indicator was motivated by the 'Median Multiple' approach pioneered by the United Nations Center for Human Settlement (UNCHS) and World Bank (KRI, 2015). Theoretically, an increase in the price-to-income ratio would indicate a decline in affordability. We hypothesise that an increase in price-to-income ratio would increase overhang. In this study, price was measured by the Malaysian House Price Index (MHPI) while income was measured by GDP per capita. MHPI data was sourced from NAPIC while data for GDP per capita was obtained from World Bank. While past studies use median income as proxy for income, we were unable to find such data at the state level. Hence, the MHPI and per capita GDP were the closest proxies we could find[20]. Both income and price were measured in nominal terms.

According to the literature (Wang and Immergluck, 2019, Tey and Kassim, 2022)[21], housing location was an important consideration when purchasing a home. Houses that were built in undesirable locations would unlikely be sold. Wang and Immergluck (2019) studied the US market, hence their methodology is unsuitable here. Meanwhile, Tey and Kassim (2022) obtained location effects on overhang from survey and not secondary data. Thus, we need to develop an appropriate location indicator using the collected secondary data. This indicator is the average travel distance and is calculated in the following manner:

- We identified the districts within a particular state/federal territory and then calculated the distance between the district seat (capital of the district) to the nearest major town. A major town was defined as one where the population exceeded 300,000. The districts within a state/federal territory were identified based on the overhang data provided by the Property Market Status reports from NAPIC. The distance indicator between the district seat and the nearest major town was obtained from Google and denominated in kilometres. We only use the distance for road travel as some states in the country do not have public transport networks. This calculation exercise was repeated for each district in the state/federal territory.
- Next, we averaged the distance indicator across all districts within the state/federal territory to obtain the average travel distance for the entire state/federal territory. The calculation of average distance was weighted by the proportion of housing stock in each

[20] There is data on household income from the Household Income Survey conducted once every two years by Department of Statistics Malaysia. However, the survey has data gaps as the data is not available annually.
[21] Analysts also concur with the importance of location. See the following source: https://www.iproperty.com.my/news/allowing-developers-more-flexibility-will-give-home-buyers-what-they-actually-want-ctr-33162

district within the state/federal territory. This exercise was repeated for every state/federal territory in our data.

- This indicator could be interpreted directly as a proxy for travel time and convenience, and thus captured locational advantages. If a state/federal territory was more developed with a higher degree of urban agglomeration with closer proximity to commercial centres and job opportunities, the average travel distance would shorten and houses in such locations would be theoretically more ideal, ceteris paribus. Examples include Penang, Melaka and Kuala Lumpur. If follows that there would unlikely to be serious overhang problems in these places. Conversely, if the state/federal territory was less developed and hence residents must endure longer travel times to commercial centres for work and recreation, then the distance indicator would be larger and hence houses in such locations would be less ideal, likely resulting in more serious incidences of overhang.
- Thus, we hypothesise that lower the average travel distance (increased locational advantage) reduces overhang.

Summary statistics, key results and analysis

Summary statistics

Table 1 displays the summary statistics of the key variables used in the study, namely the dependent variable and the main explanatory variables of affordability and location. Except for location which was measured in km, the other variables were proportions. The biggest overhang was almost 2% of the housing stock, with the average overhang rate slightly below 1% of housing stock. The travel distance in the sample varied greatly, with deviations from average hovering around 18km. This reflected the distinct geography of each state/federal territory. However, the fluctuations in affordability or housing overhang rate were rather small over the years and across various state/federal territory. Thus, the lack of variations in these variables would affect the fit of the empirical models, if these variables were specified in first differences.

Table 1: Summary statistics

Variables	Mean	Standard deviation	Minimum	Maximum
Overhang (OH)	0.0041	0.0027	0	0.0156
Average travel distance (Distance)	23.3860	17.6084	6.5895	63.7410
Affordability (Afford)	0.0148	0.0031	0.0102	0.0231

Key results and analysis

Regression analyses – baseline results

We first performed the Lagrange Multiplier (LM) tests such as the Breusch-Pagan (1980), Honda (1985), King and Wu (1997), Gourieroux, Holly and Monfort (1982) and the Standardised tests of Honda (1991) and King and Wu (Moulton and Randolph, 1989) respectively to identify the choice between pooled regression vs RE estimators. Four of the six test statistics were statistically significant, and rejected the null that there are no unobservable effects. These results favoured the use of the Random Effect (RE) model over the pooled regression model (Table 2a).

Table 2a: Pooled regression vs. Random Effect Model

Test for random effects	Test statistic
Breusch-Pagan	18.59 [0.00]
Honda	2.28 [0.01]
King and Wu	2.12 [0.02]
Standardised Honda	-1.03 [0.85]
Standardised King and Wu	-1.20 [0.88]
Gourieroux, Holly, Monfort	17.63 [0.00]

Values in squared parenthesis are p-values

In the literature for panel data analysis, it was common to estimate both the Random Effect (RE) and Fixed Effect (FE) models and then conduct the test of Hausman (1978)[22] to select the more appropriate of the two models. The results of the Hausman test suggest that the RE model would be more appropriate as a basis for drawing conclusions about the factors driving overhang (see Column (1), Table 2b).

Table 2b: Baseline results for Random Effect (RE) and Dynamic Panel Data (DPD) Models

Independent variables 1/	Models					
	(1) RE	VIF(1)	(2) DPD I	VIF(2)	(3) DPD II	VIF(3)
$(OH/HS)_{t-1}$	0.74*** (8.92)	1.13	0.56*** (30.4)	1.17	0.63*** (15.36)	5.82
Afford	0.17*** (3.62)	1.03	0.03 (0.65)	1.45	0.12*** (6.14)	3.43
Location	-0.00 (-0.96)	1.11	0.00 (0.94)	1.26	0.00 (0.31)	3.26
Constant	-0.00 (-2.03)	n.a.	n.a.	n.a.	n.a.	n.a.

[22] Hausman test was conducted with the base regression estimated using Ordinary Least Squares. However the main results displayed in Table 2b were estimated using instrumental variables method, as this procedure provided a more accurate estimation of the coefficients in the model, owing to the problems of potential endogeneity mentioned earlier.

R^2	0.54	n.a.	n.a.
Test of overall model significance: F-statistic	40.02***	n.a.	n.a.
No. of observations	156	156	156
Hausman test[2/]	0.81 [0.85]	n.a.	
Sargan test[3/]	n.a.	8.47 [0.49]	8.85 [0.45]

[1/] Numbers in parenthesis are standard errors. *** denotes statistical significance at 1%. Models estimated with White heteroskedastic-robust standard errors. Values in round parenthesis are the t-statistics.

[2/] Under the Hausman test, the null hypothesis states that the unobserved random effects are uncorrelated to the explanatory variables. This condition is important as the RE estimator is efficient and consistent under this null hypothesis. If the null is not rejected, we select RE model over the FE model. The square parenthesis is the p-value of the test-statistic.

[3/] The Sargan test provides a way to check the validity of chosen instrumental variables used to estimate the dynamic panel data models. Under the null hypothesis, the chosen instruments are valid. The square parenthesis is the p-value of the test-statistic.

The model in Column (1) Table 2b was estimated using two stage least squares/instrumental variables (TSLS/IV) approach. Notably, all models had been tested for multicollinearity problems, with the variance inflation factors (VIFs) reported in Table 2b being relatively low[23].

Adopting a general-to-specific model building procedure, we initially included not only affordability and location as explanatory variables, but also other control variables such as unemployment rate, population density, housing starts and up to 2 lags of the dependent variable.[24] Since these additional explanatory variables and the second lag of the dependent variable were statistically insignificant, they had been left out of the final results in Table 2b. Since we were keen to study the effects of affordability and overhang, these variables would remain in the model.

For thoroughness, we also tried estimated the model using General Method of Moments (GMM) of Arellano and Bond (1991). There are two variants of the model estimates, namely DPD I (column (2)) and DPD II (column (3)) in Table 2b[25]. The two dynamic panel data models yielded conflicting results, with DPD I reporting a non-significant effect of affordability on overhang whereas DPD II indicated that affordability effect was significant.

[23] The VIF are centered, since our model includes an intercept (see Groß, 2003).
[24] We limit the lagged dependent variable up to a maximum of 2 lags because the time period covered in our study was only very short (15 years). Allowing larger lags could lead to further loss in degree of freedom.
[25] The differences in the model lied in how the fixed effects (unobserved cross-section effects) were treated. In DPD I, the variables in the model were first-differenced prior to estimation. In DPD II, a method of orthogonalization was applied to remove the fixed effects. More details could be found in Arellano and Bond (1995).

Importantly, the analysis using panel data could suffer from a few issues, such as serial correlation, heteroskedasticity and correlation across the cross-section dimension of the data. To overcome these issues, we used White robust standard errors in all our previous analysis to compute the t-statistics for the coefficient estimates. This was to ensure that the hypothesis tests were not misleading.

Putting the results of the RE model[26] together with the results of the dynamic panel data models, we conclude the following:

- 2 of the 3 models favoured a statistically significant affordability effect. A deterioration of affordability would lead to increase in overhang.
- Overhang was a persistent problem, with past overhang significantly perpetuating the problem to the current period.
- Average travel distance did not seem a significant factor in explaining overhang problems. Location was not an important indicator for overhang.

Robustness checks

We performed several robustness checks on our results above. First, we re-calibrated our average travel distance indicator. Instead of defining major towns as one with population exceeding 300,000, we used the 100,000-population benchmark instead and re-calculated the average travel distance indicator using the same method explained in the Methodology, Description of Variables and Data section. Using this average distance indicator, we re-estimated the model. Second, we re-calibrated average travel distance in terms of a simple, arithmetic average instead of a weighted average used in the baseline. With this definition of average travel distance, we re-estimated the model also. The results of the re-estimation, reported in Tables 3a, 3b and 4a, 4b, were similar to the baseline results.

Table 3a: Pooled regression vs. Random Effect Model

Test for random effects	Test statistic
Breusch-Pagan	17.73 [0.00]
Honda	2.41 [0.01]
King and Wu	2.26 [0.02]
Standardised Honda	-0.89 [0.81]
Standardised King and Wu	-1.04 [0.85]
Gourieroux, Holly, Monfort	17.19 [0.00]

Values in squared parenthesis are p-values

[26] We had concluded earlier that between the RE and FE models, the former was the statistically more valid one.

Table 3b: Robustness checks with Distance measured by a weighted average of distances between a district seat and the closest town (population at least 100,000)

Independent variables [1/]	Models					
	(1) RE	VIF(1)	(2) DPD I	VIF(2)	(3) DPD II	VIF(3)
$(OH/HS)_{t-1}$	0.76*** (9.25)	1.09	0.55*** (19.38)	1.03	0.62*** (27.08)	2.07
Afford	0.17*** (3.37)	1.03	0.03 (0.70)	1.13	0.11*** (5.91)	2.03
Location	-0.00 (-0.42)	1.09	0.00 (1.56)	1.11	0.00 (0.78)	1.03
Constant	-0.00 (-2.03)	n.a.	n.a.	n.a.	n.a.	n.a.
R^2	0.54		n.a.		n.a.	
Test of overall model significance: F-statistic	37.63***	n.a.	n.a.		n.a.	
No. of observations	156		156		156	
Hausman test[2/]	0.79 [0.85]		n.a.			
Sargan test[3/]	n.a.		7.21 [0.61]		8.34 [0.50]	

[1/] Numbers in parenthesis are standard errors. *** denotes statistical significance at 1%. Models estimated with White heteroskedastic-robust standard errors. Values in round parenthesis are the t-statistics.

[2/] Under the Hausman test, the null hypothesis states that the unobserved random effects are uncorrelated to the explanatory variables. This condition is important as the RE estimator is efficient and consistent under this null hypothesis. If the null is not rejected, we select RE model over the FE model. The square parenthesis is the p-value of the test-statistic.

[3/] The Sargan test provides a way to check the validity of chosen instrumental variables used to estimate the dynamic panel data models. Under the null hypothesis, the chosen instruments are valid. The square parenthesis is the p-value of the test-statistic.

Table 4a: Pooled regression vs. Random Effect Model

Test for random effects	Test statistic
Breusch-Pagan	17.70 [0.00]
Honda	2.39 [0.01]
King and Wu	2.24 [0.02]
Standardised Honda	-0.91 [0.82]

Standardised King and Wu	-1.06 [0.85]
Gourieroux, Holly, Monfort	17.12 [0.00]

Values in squared parenthesis are p-values

Table 4b: Robustness checks with Distance measured by a simple average of distances between a district seat and the closest town

Independent variables [1/]	Models	
	(1) RE	VIF (1)
$(OH/HS)_{t-1}$	0.76*** (9.73)	1.09
Afford	0.15** (2.25)	1.02
Location	-0.00 (-0.96)	1.11
Constant	-0.00 (-1.05)	n.a.
R^2	0.56	
Test of overall model significance: F-statistic	36.22***	n.a.
No. of observations	156 .	

[1/] Numbers in parenthesis are standard errors. *** denotes statistical significance at 1%. Models estimated with White heteroskedastic-robust standard errors. Values in round parenthesis are the t-statistics

Third, we re-estimated the model using the dependent variable 'overhang share in housing starts', rather than 'overhang share in housing stock' as in the baseline. Fourth, the model was also re-estimated with 'Affordability' re-defined in terms of 2 separate explanatory variables, namely 'Price' and 'Income'. The outcomes of the re-estimation exercises (Tables 5a, 5b and 6a, 6b) were also similar to the baseline. Particularly, we highlighted that higher price led to higher overhang while higher income lowered overhang[27].

The robustness checks confirm the validity of the 3 results reported in the findings.

Table 5a: Pooled regression vs. Random Effect Model

Test for random effects	Test statistic
Breusch-Pagan	15.03 [0.00]

[27] Results for DPD I were discarded as the model had very high VIFs and hence susceptible to multicollinearity. In this case, instead of using overhang rate, we used the volume of overhang, and moved housing starts (proxy for housing supply) to the right-hand side of the equation as control variable. This specification yielded better diagnostics.

Honda	2.26 [0.01]
King and Wu	2.13 [0.02]
Standardised Honda	-1.06 [0.86]
Standardised King and Wu	-1.20 [0.88]
Gourieroux, Holly, Monfort	14.64 [0.00]

Values in squared parenthesis are p-values

Table 5b: Robustness checks with Overhang measured as a share of Housing Starts

Independent variables [1/]	Models					
	(1) RE	VIF(1)	(2) DPD I	VIF(2)	(3) DPD II	VIF(3)
(OH/Housing Starts) $_{t-1}$	0.68*** (6.99)	1.14	0.40*** (16.97)	2.83	0.44*** (11.48)	2.39
Afford	14.71*** (5.28)	1.01	5.88*** (4.95)	6.28	10.72*** (11.17)	3.37
Location	-0.00 (-0.93)	1.14	0.04 (3.33)	5.00	0.00 (0.22)	1.77
Constant	-0.15 (-3.44)	n.a.	n.a.	n.a.	n.a.	n.a.
R^2	0.54		n.a.		n.a.	
Test of overall model significance: F-statistic	36.91***	n.a.	n.a.		n.a.	
No. of observations	156		156		156	
Hausman test[2/]	2.30 [0.51]		n.a.			
Sargan test[3/]	n.a.		9.68 [0.38]		8.84 [0.45]	

[1/] Numbers in parenthesis are standard errors. *** denotes statistical significance at 1%. Models estimated with White heteroskedastic-robust standard errors. Values in round parenthesis are the t-statistics

[2/] Under the Hausman test, the null hypothesis states that the unobserved random effects are uncorrelated to the explanatory variables. This condition is important as the RE estimator is efficient and consistent under this null hypothesis. If the null is not rejected, we select RE model over the FE model. The square parenthesis is the p-value of the test-statistic.

[3/] The Sargan test provides a way to check the validity of chosen instrumental variables used to estimate the dynamic panel data models. Under the null hypothesis, the chosen instruments are valid. The square parenthesis is the p-value of the test-statistic.

Table 6a: Pooled regression vs. Random Effect Model

Test for random effects	Test statistic
Breusch-Pagan	25.37 [0.00]
Honda	2.11 [0.02]
King and Wu	1.92 [0.03]
Standardised Honda	-0.99 [0.84]
Standardised King and Wu	-1.20 [0.88]
Gourieroux, Holly, Monfort	22.32 [0.00]

Values in squared parenthesis are p-values

Table 6b: Robustness checks with Affordability measured by 2 explanatory variables, namely Price and Income

Independent variables 1/	Models			
	(1) RE	VIF (1)	(2) DPD II	VIF(2)
$(OH)_{t-1}$	0.80*** (10.74)	3.54	0.62*** (7.93)	10.41
Price	12.48*** (2.96)	2.67	7.39*** (5.64)	4.13
Income	-0.74*** (-2.76)	2.54	-0.08** (2.98)	3.57
Housing Starts	0.02 (0.90)	4.46	0.007 (0.02)	11.18
Location	-2.59 (-0.91)	1.60	173.24 (0.87)	1.20
Constant	6077.85 (2.47)	n.a.	n.a.	n.a.
R^2	0.80		n.a.	
Test of overall model significance: F-statistic	89.15***	n.a.	n.a.	
No. of observations	156 .		156	

Sargan test[2/]	n.a.	6.96 [0.43]

[1/] Numbers in parenthesis are standard errors. *** denotes statistical significance at 1%. Models estimated with White heteroskedastic-robust standard errors. Values in round parenthesis are the t-statistics.

[2/] The Sargan test provides a way to check the validity of chosen instrumental variables used to estimate the dynamic panel data models. Under the null hypothesis, the chosen instruments are valid. The square parenthesis is the p-value of the test-statistic.

Discussion of results

We earlier hypothesised that overhang could result from lack of affordability and poor housing location. In this section, we discuss the results by comparing them with the literature.

In the study, the inclusion of a housing supply indicator on the right-hand side in equation (1) did not yield a statistically significant outcome (i.e. increase in housing supply does not affect overhang rate significantly). Hence, the housing supply indicator was excluded in the final regression specification. This result could mean that rising housing supply does not generates a disproportionately larger increase in overhang volume, thereby leaving the rate of overhang unchanged. The result contradicted Gu (2008) who reported a significant housing supply effect on vacant housing rate in Singapore. On the contrary, this finding was consistent with Monkkonen (2019), who did not detect any significant impact of housing stock on vacancy rate in Mexico.

The findings were consistent with the claims in the existing literature on the effects of affordability on overhang in Malaysia, namely BNM (2017a, 2017b), Cheah et al (2017), Karim et al (2017) and Yip et al (2020). To recap salient points in the literature, BNM (2017a, 2017b) were exploratory studies where overhang was attributed to affordability. Karim et al (2017) was survey-based research investigating the perceptions of factors affecting overhang of double-storey terrace houses in Johor. The responses to the survey implied that affordability was perceived to be a major contributor to the overhang in Johor. Yip et al (2020) was a study of factors affecting unsold housing using time series data. The authors found evidence where oversupply was significantly affecting unsold housing. While the present study uses different type of data, model and methodology, the significant impact of affordability on overhang continues to hold regardless of model specifications. Our findings also support Lee et al (2017) who found house prices significantly influencing unsold houses in Korea.

To revisit the facts of the Malaysian housing market, it was mentioned in the beginning of this paper that trends in overhang coincided with the observation that the market did not fully absorb new housing supply. This led to a sharp increase in overhang after 2015 (Figure 1). There was also an increase in the proportion of housing overhang in the mid- and high-priced housing segments (Figure 5). This raises the possibility that houses could not be sold because of their prices. Thus, affordability could be related to housing overhang. The main findings of the paper support this hypothesis. Rangel et al (2023) observed that affordability had been declining since 2010, and our data captures a similar picture. Moreover, housing was particularly unaffordable to individuals between the ages 20-24 and 60-64 across all housing types. (Rangel et al, 2019). Against this backdrop of declining affordability (depicted in the literature and our data), the fact that developers were launching more houses in the upper ranges of the market (Figure 4) is likely to lead to more overhang. This narrative fits our results closely.

It was reported in Thanaraju et al (2019), Olanrewaju and Woon (2019) and Muhammad Zamri et al (2022) that financial capability to buy a new home and location affected buying preferences. Hence, we can infer that affordability and location also affected overhang indirectly. However, this argument was not supported in the findings. These studies were based on surveys that targeted specific respondents. For instance, Muhammad Zamri et al (2022) surveyed civil servants while Thanaraju et al (2019) surveyed respondents living in the Greater Kuala Lumpur area. In contrast, there a similarity between our findings and the results in Tey and Kassim (2022), namely that location did not significantly affect overhang for the entire country.

Conclusion and policy recommendations

Overhang in the housing market has garnered attention by the media and government. There was a particularly sharp increase in overhang after 2015. A large proportion of these unsold houses were in the mid- and high-price segments, suggesting that affordability could be the issue. Some media analysts pointed out that poor housing location could also explain the volume of overhang. Thus, a formal inquiry into this concern is timely and necessary. Our study is a contribution to the literature in this regard. We created a panel dataset of 12 states/federal territories in the country during the period 2008-2022. Hypothesising that overhang is primarily linked to affordability, location, and several demographic factors, our findings from panel data models indicated that overhang was a persistent problem with past overhang contributing to the accumulation of more overhang in the current period. More importantly, we presented evidence that deterioration in affordability contributed significantly to the overhang. However, the view that housing location affected overhang was only weakly supported by the data. These results survived a series of robustness checks.

The most crucial practical implication of this study is to disprove claims that overhang had little to do with the construction of expensive houses.[28] Affordability (both price and income effects) had been instrumental in contributing to overhang. The next question to consider would be how overhang would interact with prices and the relevant market dynamics. Presently, there is insufficient data to support the estimation of a fully specified housing market model. The connection between overhang and affordability cannot be ignored. While overhang and vacancies in housing markets are normal, as Gu (2008) and Monkkonen (2019) observed, it would be helpful to understand the underlying factors affecting these trends. In this regard, a policy recommendation to the government would be to expedite efforts to make housing more affordable if they wished to reduce overhang. How this can done, however, is beyond the scope of this paper.

What our study also showed was the lack of robust link between average travel distance (which proxied for location) and overhang. States and federal territories with better

[28] Comment by a local analyst seem to imply that overhang came from other factors: https://www.iproperty.com.my/news/overhang-residential-property-market-62492

connectivity to commercial centres and strategic places of interest (e.g. job markets) did not necessarily have fewer unsold houses, after controlling for the size of the local housing market. While location has been cited as a key consideration for housing choices, its contribution to overhang may have been overstated.

A final remark on what this study implies for future research. Average travel distance was not significant in explaining housing overhang. This begets the question of what other housing characteristics may be important in helping to shape house buying decisions and the relationship with overhang. Second, more detailed studies of changes in land use regulations (e.g. gentrification) and its impact on housing development and overhang is warranted given the rise in mixed development projects (shopping malls with commercial complex and residential units in the same plot of land). Third, more micro-oriented studies of metropolitan areas surrounding major cities in the country could shed light on the drivers of overhang. In this regard, a study by Ho and Lim (2022) found that regions adjacent to low-cost housing areas suffered from lower house prices – to what extent do such location effects affect overhang would be a key question to answer. This endeavour requires more dedication to the collection of micro-data.

Conflict of interest statement

The authors declare that they do not have any conflicts of interest to disclose. The authors do not have any actual, or perceived personal interests which could affect the impartiality of the publication process.

References

Arellano, M., and S. R. Bond (1991). "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," Review of Economic Studies, 58, 277–297.

Arellano, M., and O. Bover (1995). "Another Look at the Instrumental Variables Estimation of Error-components Models," Journal of Econometrics, 68, 29–51

Aris, N.A.M., Fathi, M.S., Harun, A.N. and Mohamed, Z. (2020) 'Spatial Pattern of Affordability and Overhang in Selangor', IOP Conference Series: Earth and Environmental Science, 540

BNM (2017a) Quarterly Bulletin (3rd Quarter), Bank Negara Malaysia

BNM (2017b) Quarterly Bulletin (4th Quarter), Bank Negara Malaysia

Breusch, T.S. and 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

Cheah, S.L., Almeida, S., Shukri, M., and Lim, L.S. (2017), Imbalances in the Property Market, Staff Insights 2017/13 (Economics Department), Bank Negara Malaysia

Durst N and Ward P (2015) Lot vacancy and property abandonment: Colonias and informal subdivisions in Texas. International Journal of Housing Policy 15(4): 377–399

Esteban, M. and Altuzarra, A. (2016), Local Political Power and Housing Bubble in Spain, Journal of Regional Research, 35, 107-127

Gan, L., He, Q., Si, R. and Yi, D. (2019), Relocating or Redefined: A New Perspective on Urbanization in China, NBER Working Paper 26585

Gentili, M. and Hoekstra, J. (2019), Houses without people and people without houses: a cultural and institutional exploration of an Italian paradox, Housing Studies, 34(3), 425-447

Glaesser, E., Huang, W., Ma, Y., Shleifer, A. (2017) 'A Real Estate Boom with Chinese Characteristics', Journal of Economic Perspectives, 31(1), pp. 93-116

Groß, J. (2003) Variance Inflation Factors, R News, 3(1), pp. 13-15

Gourieroux, C., A. Holly, and A. Monfort (1982). "Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters," Econometrica 50, 63–80.

Gu, J. (2008) Singapore Housing Market: An Econometric Model and A Housing Affordability Index, Master of Social Sciences Dissertation, National University of Singapore

Gujarati and Porter (2009) Basic Econometrics, McGraw-Hill Irwin

Hassan, M.A., Abdullah, Y.A., Omar, D., Hakim Danial, M. (2021), Location Housing Affordability Index: Analysing the Relationships, Planning Malaysia, 19(1),41-52

Honda, Y. (1985) "Testing the Error Components Model with Non-normal Disturbances," Review of Economic Studies, 52, 681–690.

Honda, Y. (1991) "A Standardized Test For the Error Components Model with the Two-Way Layout," Economics Letters, 37, 125–128.

Ho, G.W.S. and Lim, H.H. (2022) 'Living next to poor housing: A Regression Analysis of House Prices in Greater Kuala Lumpur', Khazanah Research Institute Working Paper 1/22

Karim, N.S.A., Maimun, N.H.A., Mat Noor, N.A., Mohd Yusoff, N.S., Abdul Rahman, N.S. (2017), Oversupply Causes of Double Storey Houses in Johor Bahru, International Journal of Real Estate Studies, 11(3), 31-36

KRI (2015) *Making Housing Affordable*, Khazanah Research Institute (Kuala Lumpur)

KRI (2022) *Residential Settlements and Spatial Inequality*, Khazanah Research Institute (Kuala Lumpur)

Kunasekaran, T. (2023) 'Curbing Overhang and Vacancy: The Case for a Vacancy Tax in Malaysia', Khazanah Research Institute (KRI) Views 19/23, https://www.krinstitute.org/assets/contentMS/img/template/editor/KRI%20Views_Curbing%20Overhang%20and%20Vacancy_The%20Case%20for%20a%20Vacancy%20Tax%20in%20Malaysia.pdf

King, M. L. and P. X. Wu (1997). "Locally Optimal One Sided Tests for Multiparameter Hypotheses," Econometric Reviews, 16, 131–156.

Lee, Y., Lee, S. and Kim, J. (2017), Analysis of the Dynamic Relationship between Fluctuations in the Korean Housing Market and the Occurrence of Unsold New Housing Stocks, Sustainability, doi:10.3390/su9010107

Ma, B. (2022), How Evergrande Defaulted: Positive Feedback Mechanism Subject to the Debt-Fueled Model in Chinese Real Estate Market, Advances in Economics, Business and Management Research, 654, 424-429

Monkkonen, P. (2019), Empty houses across North America: Housing Finance and Mexico's vacancy crisis, Urban Studies, 56(10), 2075-2091

Moulton, B. R. and W. C. Randolph (1989). "Alternative Tests of the Error Components Model," Econometrica, 57, 685–93

Muhammad Zamri, N.E.M., Yaacob, M. and Mohd Suki, N. (2022), "Assessing housing preferences of young civil servants in Malaysia: do location, financial capability and neighbourhood really matter?", International Journal of Housing Markets and Analysis, Vol. 15 No. 3, pp. 579-591.

Nam, J., Han, J. and Lee, C. (2016) Factors Contributing to Residential Vacancy and Some Approaches to Management in Gyeonggi Province, Korea, Sustainability, doi:10.3390/su8040367

Newman, G., Lee, R.J., Gu, D., Park, Y., Saginor, J., Van Sandt, S., Li, W. (2019) Evaluating drivers of housing vacancy: a longitudinal analysis of large U.S. cities from 1960 to 2010

Olanrewaju, A. and Woon, T.C. (2019) 'An exploration of determinants of affordable housing choice', International Journal of Housing Markets and Analysis, 10(5), pp. 703-723. DOI: 10.1108/IJHMA-11-2016-0074

Patrick, R.H. (2020) Chapter 28: Durbin–Wu–Hausman Specification Tests Handbook of Financial Econometrics, Mathematics, Statistics, and Machine Learning, pp. 1075-1108

Rangel, G. J., Ng, J. W. J., Murugasu, T. T., & Poon, W. C. (2023) 'Measuring long-run housing affordability for Malaysian millennial households: a geospatial and income distribution analysis', International Journal of Housing Markets and Analysis. DOI: 10.1108/IJHMA-02-2023-0017

Rangel, G. J., Ng, J. W. J., Murugasu, T. T., & Poon, W. C. (2019) 'A Micro-Level View of Housing Affordability in Malaysia Using an Age Cohort-housing Type Analysis', 56(1), pp. 1-22

Rogoff, K. and Yang, Y. (2020), Peak China Housing, NBER Working Paper 27697

Saunders, T. and Tulip, P. (2019) A Model of the Australian Housing Market, Reserve Bank of Australia Research Discussion Paper, RDP 2019-01

Sorace, C. and Hurst, W. (2015), China's Phantom Urbanisation and the Pathology of Ghost Cities, Journal of Contemporary Asia, doi: 10.1080/00472336.2015.1115532

Steiner, E. (2010) 'Estimating a stock-flow model for the Swiss housing market', Swiss Journal of Economics and Statistics, 146, pp. 601-627

Struyk R (1988) Understanding high housing vacancy rates in a developing country: Jordan. The Journal of Developing Areas 22(3): 373–380

Tan, Z., Wei, D. and Yin, Z. (2020) Housing Vacancy Rate in Major Cities in China: Perspectives from Nighttime Light Data, Complexity, 2020 (ID 5104578), DOI: 10.1155/2020/5104578

Tey, C.S. and Kassim, R. (2022) 'The Factors for Residential Properties Overhang in Johor', Research in Management of Technology and Business, 3(1), pp. 809-824

Thanaraju, P., Mentaza Khan, P.A., Juhari, N.H., Sivanathan, S., Md Khair, N. (2019), Factors Affecting the Housing Preferences of Homebuyers in Kuala Lumpur, Planning Malaysia, 17(1), 138-148

Wooldridge, J.M. (2016) Introductory Econometrics: A Modern Approach. Adrian MI: South-Western Cengage Learning.

Yip, C.Y., Senadjki, A., Au Yong, H.N., Abdul Adzis, A. (2020), Mitigating housing glut: an application to the Malaysian housing market, International Journal of Housing Markets and Analysis, 14(3), 498-522

Yoo, H. and Kwon, Y. (2019), Different factors affecting vacant housing according to regional characterstics in South Korea, Sustainability, DOI: 10.3390/su11246913

Yoo, H. and Yoon, H. (2021) The Effect of Green Characteristics in Reducing the Inventory of Unsold Housing in New Residential Developments—A Case of Gyeonggi Province, in South Korea, Land, doi: 10.3390/land10040377

Yusoff, N.A., Mohd Shafiei, M.W., Yahya, S. and Ridzuan, M. (2010) 'Strategies to implement the "build then sell" housing delivery system in Malaysia', Habitat International, 34(1), pp. 53-58