Dynamic Hedonic Analysis Using Time-Varying Coefficients:
Application to Dubai’s Housing Market

Hiroki Baba
(Center for Spatial Information Science, The University of Tokyo)

Hayato Nishi
(Graduate School of Engineering, The University of Tokyo)

Ashoka Mahabala Seetharamapura
(Emirates Real Estate Solutions, Dubai Land Department)

Chihiro Shimizu
(Center for Spatial Information Science, The University of Tokyo/
Nihon University)
Abstract

In the housing market analysis, capturing the temporal change in market structure is of importance. We aim at proposing housing price index estimation models employing time-varying coefficients and compare the models considering the lengths of the time periods. We firstly employ ordinary least squares for each time period: a separate hedonic model, and then, a time window is applied for a specific length of time period: a rolling hedonic model. We, thirdly, consider coefficient-wise stochastic innovation terms: a dynamic hedonic model, since random walks in the Kalman filter enable us to estimate time-varying coefficients that are robust against exogenous shocks. We study the Dubai’s housing market, where housing prices have been very volatile for political reasons. One of our findings indicates that the dynamic hedonic model shows the highest predictive power. Moreover, since the trends of the coefficients differ for each variable, observing them helps identify the factors that are more important for explaining housing price indexes. A stability analysis shows that the separate hedonic model is far more unstable than the others, whereas the dynamic hedonic model is the most stable method for all the variables. We consider that the dynamic hedonic model is applicable even when only a small number of samples have been obtained, because the method refers to all the samples for smoothing the coefficients.

Keywords: housing price index; hedonic model; Kalman filtering; Dubai; housing market.

JEL Classification: C43; E31; R31.
1 Introduction

In a housing market analysis, it is important to capture temporal change in the market structure. A failure to predict the market conditions can cause a financial crisis (Kaplan, Mitman, and Violante 2017). A great number of papers, thus, have constructed a housing price index (HPI) to improve the reliability and timeliness (Case and Shiller 1989; Hill and Scholz 2018; Shimizu et al. 2010; Shimizu, Nishimura, and Watanabe 2010; Diewert and Shimizu 2015, 2016).

One of the fundamental ways to construct an HPI is to employ a hedonic method, whose theoretical framework is provided by Rosen (1974). The hedonic method includes a market clearing function within the interaction between households’ bid functions and suppliers’ offer functions (see Chapter 2, Diewert et al. (2020) for details). The framework of a hedonic approach is applied extensively not only to HPIs (Goodman 1978), but also to measure the demand for housing (Shefer 1986) and the impact of neighborhood externalities on housing prices (Michaels and Smith 1990).

Hedonic methods have been developed to deal with changes in time series (Hill 2013). One of the basic methods dealing with time is adding time fixed dummies, but the idea of rolling time dummies has been getting more attention in recent years (Shimizu et al. 2010). Hill et al. (2020) explore the optimal length of the rolling window and the optimal linking method by minimizing the $X$ metric proposed by Diewert (2002, 2009). Nevertheless, those studies use samples within time periods of a fixed length to estimate the coefficients and regard the coefficients as non-stochastic variables, so that those models are not always robust against exogenous shocks.

We aim at proposing housing price index estimation models employing time-varying coefficients and compare the models considering the lengths of time periods. We firstly employ ordinary least squares for each time period, and secondly, a time window is applied for a specific length of time period. We, thirdly, consider coefficient-wise stochastic innovation terms, since random walks in the Kalman filter enable us to estimate time-varying coefficients that are robust against exogenous shocks. The idea of stochastic coefficients was proposed by Guirguis, Giannikos, and Anderson (2005) with the use of time-varying Kalman filtering. However, since they explore various methods from the perspective of macro finance, the micro-structure of the housing market has not been well
investigated using their framework.

Another novelty in the present paper is the study area: Dubai, one of the most prosperous real estate markets in the Arab countries. Despite the success of its real estate market, Dubai has sometimes experienced large volatility in its housing prices, due to geopolitical issues. It is true that Hepsen and Vatansever (2011, 2012) have previously tried to forecast HPIs in Dubai, but they focus on macroeconomic trends, not considering the micro-structure of the buildings and property owners. We thus think that validating our proposed method fits well in Dubai and well predicts the rise and fall in the HPIs.

2 Dubai’s housing market

Dubai is one of the most attractive cities in the world. Surrounded by desert, a large part of the land is artificially developed and significant tall buildings are agglomerated. The characteristic of the city is the high density of buildings in its central business district. In the field of economics, high density encourages knowledge spillover and drives growth of a city (Glaeser et al., 1992). While such a high density in Dubai is appreciated among experts and policy makers, the citizens in Dubai consider that the other areas also represent their cultural context (Alawadi and Benkraouda, 2019). For the middle-class population, considering not only the high density area but also the other areas enhances the value of the city (Alawadi, Khanal, and Almulla, 2018).

From a bird’s eye view, a developed land that seems to be a palm tree is called Palm Jumeirah and Palm Jebel Ali, where all the lands in the island are reclaimed and developed for the rich (Figure 1). The reason why they have developed the island is that the acquisition of an ocean view is even worth developing the reclaimed land. The tallest building in Dubai is called Burj Khalifah, with a height of 828.0 m and 206 stories. This height provides both buyers and sellers with a rent premium for a good view.

Knowing the characteristics of the property rights in Dubai helps us understand the nature of the housing market. Due to the influence of the British housing market system, Dubai’s property rights are categorized into freehold and non-freehold. Although in the

---

1The base information of the map is retrieved from OpenStreetMap (https://www.openstreetmap.org).

2British property law originally categorized property rights into freehold and leasehold. While in the freehold market, any buyers and sellers are able to transact as they like, in the leasehold market, either the individuals or the corporate bodies own the property right and they lease the right for a specific term.
freehold market, any foreign buyers and sellers are able to transact as they like, this area is clearly delineated. In the non-freehold market, only residents of the Gulf Cooperation Council are allowed to make property transactions. Since foreign capital is invested in the freehold market, innumerable skyscrapers have been built in the area (Bagaeen, 2007).

For further prominent successes in Dubai’s housing market, dealing with risks from geopolitical issues is of importance. For instance, diplomatic relations with adjacent countries would affect the stability of the financial status, leading housing price volatility.

Focusing on the freehold market, such exogenous shocks are amplified due to the specific market mechanism. Most properties in the freehold market start appearing on the market together with the beginning of development, and developers collect the expense of construction. Since the development lasts for two to three years, most buyers expect to purchase the properties for the purpose of investment. Actually, real estate prices in Dubai is severely influenced by the 2007–2008 financial crisis, causing discontinuation of many development projects since then.

Considering the uncertainty, investors carefully observe whether and when economic

---

3The Gulf Cooperation Council consists of the United Arab Emirates, Bahrain, Kuwait, Oman, Qatar, and Saudi Arabia.
growth stagnates. One of the risks in the market is housing oversupply. Since foreigners expected to invest properties, substantial housing properties were supplied. Moreover, housing market, not only in Dubai but in countries worldwide, is always facing exogenous shocks, such as COVID-19 crisis. Thus, predicting the trend of the housing market is important, for the sake of reducing the uncertainty in the market.

3 Method

3.1 Hedonic models and HPI

We start by explaining the hedonic approach theorized by Rosen (1974), one of the most popular models to estimate HPIs. The hedonic method deals with a price as a bundle of housing characteristics and estimates the value using regression analysis. When we obtain pooled data that includes comprehensive housing characteristics, it enables us to calculate quality-adjusted price indexes.

The estimation of HPIs with hedonic models has two major problems: omitted variables bias and structural change. The former problem occurs due to the difficulty in collecting all the variables required for the estimation of the functions as well as the presence of unobservable variables (Case and Quigley, 1991; Clapp, 2003). The latter involves the necessity of accommodating changes in the house price structure, since HPIs change over long periods of time (Case, Pollakowski, and Wachter, 1991; Clapp, Giaccotto, and Tiririglou, 1991; Clapp and Giaccotto, 1998).

Possible ways to avoid the problem of omitted variables bias include collecting other variables or using the repeat-sales method. The repeat-sales method estimates HPIs by calculating changes in property prices based on the same property within a specific time period (Case and Quigley, 1991; Case and Shiller, 1989). While this method is convenient because the fixed effects cancel out, the sample size for properties that are traded more than twice is likely to be small, especially in a thin market like Japan, which may cause a sample selection bias. In the case of Dubai’s data, the sample period is limited to 2007–2018, so that we decided to collect as many variables as possible, instead of applying the repeat-sales method.

Under such circumstances, the importance of estimating HPIs with high accuracy is
associated with changes in the market structure. To consider changes in the market structure, we introduce two models: the restricted hedonic model and the overlapping-period hedonic model, and then propose hedonic models with time-varying coefficients. The restricted hedonic model fixes the market structure and measures the HPIs with the baseline of the HPI in period \( t \). The overlapping-period hedonic model, proposed by Shimizu et al. (2010), captures the successive changes in the regression coefficients with a certain period length, which is called a window. Our proposed hedonic models also assume that the market structure changes each period, and moreover one of the models assumes that the regression coefficients have a stochastic nature.

Assuming a semi-log for the hedonic function, we describe the restricted hedonic model:

\[
\ln p_{it} = \sum_{k=1}^{K} \beta_k x_{kit} + \sum_{s=1}^{T} \delta_s d_s + \varepsilon_{it} \quad (1)
\]

where \( p_{it} \) is the transaction price for housing unit \( i \) in period \( t \), \( \beta_k \) are the coefficients in \( k \)th housing characteristic, \( x_{kit} \) is the housing attribute in \( k \)th housing characteristics and \( i \)th housing unit in period \( t \), \( \delta_s \) denotes the time dummy coefficients in period \( s \), \( d_s \) is an index that takes the constant value of 1 when \( s = 1 \), while for \( 2 \leq s \leq T \), it is a time dummy variable, taking the value of 1 when \( s = t \) and the value of 0 otherwise, and \( \varepsilon_{it} \) is the idiosyncratic error in housing unit \( i \) in period \( t \). This model assumes that the regression coefficient \( \beta_k \) is constant through the periods. Using ordinary least squares, \( \hat{p}_t \) is estimated as follows.

\[
\ln \hat{p}_t = \sum_{k=1}^{K} \hat{\beta}_k x_k + \hat{\delta}_1 + \hat{\delta}_t \quad (2)
\]

\[
\ln \hat{p}_1 = \sum_{k=1}^{K} \hat{\beta}_k x_k + \hat{\delta}_1 \quad (3)
\]

where \( \hat{\beta}_k \), \( \hat{\delta}_1 \), and \( \hat{\delta}_t \) are the estimated parameters. Subtracting equation 3 from equation 2, the HPI in period \( t \) is calculated by:

\[
\ln \left( \frac{\hat{p}_t}{\hat{p}_1} \right) = \hat{\delta}_t \quad (4)
\]

where the HPI in period \( t = 1 \) is used as the reference. In the same way, the change in the
HPI from $t - 1$ to $t$ is calculated by:

$$\ln\left( \frac{\hat{p}_t}{\hat{p}_{t-1}} \right) = \hat{\delta}_t - \hat{\delta}_{t-1}. \quad (5)$$

This indicates that the HPI in period $t$ is normalized by the HPI in period $t - 1$, so that equation (5) enables us to compare the HPIs between periods $t$ and $t - 1$. However, the restricted hedonic model has several problems. First, the HPIs are estimated using only the time dummy variables without any housing attributes. This causes an estimation bias if the housing characteristics have been changing during the periods. Second, autocorrelation in the time series data should be considered. For instance, seasonality and its changes over time may affect the HPIs.

Assuming structural changes in the market, the overlapping-period hedonic model defined by Shimizu et al. (2010) is given as follows:

$$\ln p_{it} = \sum_{k=1}^{K} \beta_k x_{kit} + \sum_{s=1}^{\tau} \delta_s d_s + \varepsilon_{it} \quad (6)$$

where $t = 1, 2, ..., \tau$ and $\tau$ indicates a specific period length (time window) satisfying $1 \leq \tau \leq T$.

The HPIs are estimated with successive time windows $[1, \tau], [2, \tau + 1], ..., [r, \tau + r - 1], ..., [T - \tau + 1, T]$. This successive manner of estimating the HPIs can take into account changes in the parameters. With the overlapping-period hedonic model, the period length $\tau$ is the length of the overlapping estimate period, so that the model is a type of restricted hedonic model with respect to a certain period length $\tau$. The overlapping-period hedonic models for all periods are estimated successively by shifting the period length $\tau$ one by one.

### 3.2 Proposed models

We now extend the hedonic model using time-varying coefficients. Here, we modify the overlapping-period hedonic model with a time window size $[t-\tau, t+\tau]$, which means that
the period length is \(2\tau + 1\). We regress the housing prices within the time window:

\[
\ln p_{it} = \sum_{k=1}^{K} \beta_k x_{kit} + \delta_t + \varepsilon_{it}, \quad \text{for } t \in [t-\tau, t+\tau].
\] (7)

The estimated HPIs can be considered as smoothing values in the designated time window. When \(\tau = 0\), we separately estimate the HPIs for each \(t = 1, 2, ..., T\). We call this the "Separate Hedonic Model" and our first proposed model as a benchmark of the HPI values.

We then designate the time window size as a fixed value. Ideally, a time window size should be determined by some criteria such as AIC, BIC, and so forth. However, since this analysis is the starting point for the time-varying coefficients approach, we set the window size to be \(\tau = 1\). We call this model the "Rolling Hedonic Model" and our second proposed model to compare with the separate hedonic model.

Although both the separate hedonic model and the rolling hedonic model deal with the coefficients of housing attributes as non-stochastic parameters, the coefficients can be redefined as stochastic parameters. The reason is that even though we consider the chronological changes in the HPIs, we still need to predict the coefficients \(\hat{\beta}_k\).

To smooth the coefficients, we employ another approach, where we deal with the dynamic changes in the coefficients. We specify a hedonic function with a random walk model. The hedonic function is slightly modified:

\[
\ln p_{it} = \sum_{k=1}^{K} \beta_{kt} x_{kit} + \delta_t + \varepsilon_{it}.
\] (8)

In the random walk model, since the coefficients in period \(t\) fluctuate from the previous period, we have

\[
\beta_{dt} = \beta_{dt-1} + v_{dt}
\] (9)

where \(\beta_{dt}\) is the \(d\)th value of the coefficient and \(v_{dt}\) is a stochastic innovation term that follows a Gaussian distribution.

We then estimate the models with the Maximum Likelihood method in the framework of Kalman filtering. This estimation procedure optimizes coefficient-wise the variance
(σ_d)^2, which is the variance of the noise \( v_{dt} \), which implies that the strength of the smoothing is automatically determined. Therefore, we can distinguish the dynamic coefficients and the nearly non-dynamic coefficients. We call this the “Dynamic Hedonic Model”, which is our third proposed model, and considers the stochastic variation between periods.

4 The data

To compare these methods, we use the housing sales data in Dubai from 2007 to 2018. Although the sample size fluctuates each year, every year has more than 1,000 samples. Therefore, we think that the estimation result obtained from the data can be stable even after the exclusion of all the outliers (Table 1).

This data includes building coordinates, floor area space, and buyers’ information such as nationality, gender, and age. Therefore, building fixed effects are controlled through the coordinates of the building. Taking advantage of the POIs from OpenStreetMap[^1] we have measured the distances from each building to the nearest grocery store, restaurant, park, police station, mosque, university, school, hospital, and coastal line. Since the shape of the roads in Dubai is complex and measures the distances in a small scale, we calculated the distances via the road network. Moreover, we assume that land use restrictions affect the housing prices, so that a residential land use dummy is included in our models.

[^1]: https://www.openstreetmap.org/
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>gen_dummy</td>
<td>Gender dummy of buyer</td>
</tr>
<tr>
<td>AGE_AT_CONTRACT_TIME</td>
<td>Buyer’s age at contract time</td>
</tr>
<tr>
<td>UNIT_SIZE_SQM</td>
<td>Floor space (m²)</td>
</tr>
<tr>
<td>RES</td>
<td>Land use regulation dummy (residential use)</td>
</tr>
<tr>
<td>grocery_dist_st</td>
<td>Street distance to the nearest grocery store</td>
</tr>
<tr>
<td>restaurant_dist_st</td>
<td>Street distance to the nearest restaurant</td>
</tr>
<tr>
<td>park_dist_st</td>
<td>Street distance to the nearest park</td>
</tr>
<tr>
<td>police_dist_st</td>
<td>Street distance to the nearest police station</td>
</tr>
<tr>
<td>mosque_dist_st</td>
<td>Street distance to the nearest mosque</td>
</tr>
<tr>
<td>univ_dist_st</td>
<td>Street distance to the nearest university</td>
</tr>
<tr>
<td>school_dist_st</td>
<td>Street distance to the nearest school</td>
</tr>
<tr>
<td>hospital_dist_st</td>
<td>Street distance to the nearest hospital</td>
</tr>
<tr>
<td>coastal_dist_st</td>
<td>Street distance to the nearest coastal line</td>
</tr>
</tbody>
</table>

The housing attributes used in this paper are shown in Table 2. For the preprocessing of the data, some variables have been trimmed; AGE_AT_CONTRACT_TIME lower than 18 and higher than 100, and UNIT_SIZE_SQM lower than 20 and higher than 400 are omitted, respectively. Moreover, all the variables are standardized. We then conducted a list-wise deletion of the data, which means the data for analysis does not include either NA or NULL values.

5 Result

The estimation results are illustrated in Figure 2. For simplicity, we only show the results for const, UNIT_SIZE_SQM, gen_dummy, and grocery_dist_st. The coefficients of const can be interpreted as annual standard prices. UNIT_SIZE_SQM is the most influential variable for prices, having the largest coefficient. gen_dummy and grocery_dist_st are taken as examples of dummy and continuous variables, respectively.

Direct inspection of figure 2 indicates that there are distinctive trends for all the variables. Observing UNIT_SIZE_SQM, we find a gradual increase in the coefficient, which means that the floor space of the properties becomes more important as time passes. Focusing on one of the property owners’ characteristics, gen_dummy, we see it decreases according to the time periods. This presumably means that gender makes no difference in the recent Dubai’s real estate market of Dubai.
We also find that the estimated coefficients of the dynamic hedonic model are similar to those of the separate hedonic model. In particular, the coefficients of the influential variables const and UNIT\_SIZE\_SQM are completely overlapped. This indicates that the dynamic hedonic model weakly smooths the coefficients. We presume that because the sample size each year is large enough to obtain reliable estimated values, the coefficients are not necessarily smoothed out.

We have also examined the stability of the estimation for a small sample size, using the following procedure. We pick 1% samples each year at random and estimate the coefficients. This procedure is repeated 30 times to test the stability of the estimations.

The results are shown in Figure 3. The estimation of the separate hedonic model is far more unstable than the others, so that we present the results obtained from both the rolling and the dynamic hedonic models in Figure 4. These figures show that both the rolling and dynamic hedonic models are more stable than the separate hedonic model. Comparing the dynamic hedonic model to the rolling hedonic model, the estimated coefficients of the influential variables, such as const and UNIT\_SIZE\_SQM, are approximately equal. However, the dynamic hedonic model is more stable in the estimation of the less influential variables, such as gen\_dummy and grocery\_dist\_st.

The possible reasons for the stability of the dynamic hedonic model are as follows. First, The dynamic hedonic model refers to all the data used to estimate and smooth the
coefficients. This means that the dynamic hedonic model can refer to a larger number of samples than the other methods. Second, the automatic control of strength of the smoothing contributes to the stability. A smoothing with an appropriate strength makes the coefficients temporally stable, while the rolling hedonic model is not as stable as the dynamic hedonic model due to the fixed length of the time window. Another finding is that temporal changes in the coefficients of $gen_{dummy}$ and $grocery_{dist}_{st’}$ in the dynamic hedonic model are small even if only a small number of samples are available.

6 Concluding remarks

Based on time-varying coefficients, we have compared three proposed methods for measuring HPIs.

The results indicate that the dynamic hedonic model shows the highest predictive power of the three and the HPI roughly follows the other two HPIs. However, since the sample size in this analysis is sufficient, the coefficients only yield sight differences between the three methods. Moreover, since the trends of the coefficients differ for each
variable, observing them helps identify the factors that are more important for explaining housing price indexes.

In a stability analysis using a small sample size, we confirmed that the separate hedonic model is far more unstable than the others. The dynamic hedonic model is the most stable method of the three, for all the variables. We consider that the dynamic hedonic model is applicable even to only a small number of samples, because the method refers to all the samples for smoothing the coefficients.

Although we have only assumed a random walk error in the dynamic hedonic model, it can be extended to linear (see Appendix) and non-linear trends. Further exploration is needed for the model structure of time-varying coefficients.

Acknowledgement

The authors are grateful to Dubai Land Department for data provision and helpful comments.
References


Appendix

In addition to a random walk coefficients model, we also attempt a linear trend coefficients model. The latter model considers a linear trend among the coefficients, so that the first difference of the $d$th value of a coefficient in period $t$ is equal to the one in period $t - 1$ with an error:

$$\beta_{dt} - \beta_{dt-1} = \beta_{dt-1} - \beta_{dt-2} + \nu_{dt}$$ (10)

where $\nu_{dt}$ is a stochastic innovation term that follows a Gaussian distribution. This estimation procedure also optimizes the coefficient-wise variance $(\sigma_d)^2$ of the variance noise $\nu_{dt}$.

Despite the development of the model, we would not highlight it, because the predictive power of the model is not superior to the dynamic hedonic model proposed in the main text. The figure below illustrates the $R$-squared between the two models, which shows the dynamic hedonic model is the better one to employ.

![Comparison between random walk and trend models](image)

Figure 5: Comparison between random walk and trend models