Forecasting Housing Choices Selection In Penang, Malaysia

Mohd Ali, N.H.S., Zainun, N.Y

Abstract— National Property Information Centre (NAPIC), Malaysia recorded that for the first half year 2018 there was a decline about 37.8% on the residential stock especially for new planned supply and an increment of 18.2% on the overhang units. However, in Penang, it was estimated the demand to reach a total of 46,740 units due to the growing number of populations, formation of new households and the replacement of existing houses. Crucial housing challenge in Malaysia, especially in Penang is majorly related to a mismatch in demand and supply for affordable housing. The goal of this research is to predict a housing demand in Penang based on 4 categories of housing which is consist of low-cost, low-medium cost, medium cost, and high-cost houses based on household formation. The research used Census Data 2010 from the Department of Statistic Malaysia to determine the headship rate and will use the Headship Rate Method to obtain household formation. A questionnaire was designed and approved by the expert before distributed to 400 households. The households were choose based on five districts in Penang and the respondents or household were divided according to 14 levels of age group which are; (1) 15-19 years old; (2) 20-24 years old; (3) 25-29 years old; (4) 30-34 years old; (5) 35-39 years old; (6) 40-44 years old; (7) 45-49 years old; (8) 50-54 years old; (9) 55-59 years old; (10) 60-64 years old; (11) 65-69 years old; (12) 70-74 years old; (13) 75-79 years old and (14) 80 and above. Multinomial Logit analysis was carried out to determine the choice of probability on house type selection and to produce a Choice Probabilities (CP), Model. The result analysis showed that Medium Cost housing was the most preferable type of house with the highest CP value of 0.4875. This prediction can assist local authorities, developers, consultants, contractors to plan which type of housing to construct based on demand in the future.

Index Terms— Household, headship rate, choices of probabilities, prediction.

1 INTRODUCTION

In Eleventh Malaysian Plan (2016-2020), Economic Planning Unit (2015) projected that the construction sector is expected to grow by 10.3% per annum during the plan period. This increment is due to continuous of civil engineering works and sub residential growth to meet housing demand, mainly from the middle-class citizen. Crucial housing challenge in Malaysia, especially in Penang is majorly related to a mismatch in demand and supply for affordable housing. Even though many initiatives had been done, not all the initiatives that had been taken by the government or developer to solve this issue was successful (Hong, 2016; Kamal et al., 2016). Till 31 July 2017, Ministry of Urban Wellbeing, Housing and Local Government had reported that about forty-eight (48) abandoned housing projects were in recovery plan (i.e. in the process of identifying able developers to rescue the projects) whereby 8,932 units were built and only 6,082 units, approximately 68% were able to be sold. With the rapid urbanization and the increment of the proportion of the population had resulted in a severe shortage of affordable housing. The first half of 2018, Malaysia Property Market Report reported that the Northern Region contributed 33,275 units of overhang and unsold residential property, which is the second-in-line after Central Region (NAPIC, 2018). Therefore, prediction on housing demand is crucial to know the actual number of houses and which type of houses to be built to match the demand.

.

- *Mohd Ali, N.H.S is currently pursuing masters degree program in civil engineering in Universiti Tun Hussein Onn Malaysia, MALAYSIA. E-mail: masterhelmie@yahoo.com*

2 POPULATIONS, HOUSEHOLD FORMATION AND HEADSHIP RATE METHOD

(Hong, 2016; Macdonald, 2011) stated that the rapid rate of rural dwellers migrating to urban centers has created an increasing demand for housing, particularly affordable low-income houses in many towns. In 1970, Penang state population was only 776,124 and continued to increase to 1,064,166 in 1991 with an annual growth rate of 1.53% between 1980 and 1991 (Munir et al., 2010). The Penang population increased quite significantly in nine years to 1,313,449 in 2000, representing a 2.37% annual growth rate between 1991 and 2000. In the subsequent ten years (2000 to 2010), the Penang population has been estimated to increase at 3.05% annually to 1,773,442 in 2010 and at 2.89% annually to 2,357,982 by 2020. The key drivers of household formation over time are population growth and the evolution of headship rates (Furlong, 2016).

Headship model will be using the time series data information taken from the census database collected from

Department of Statistic Malaysia to determine the headship rate. This census data for example age and the population will be act as a secondary time series data, and this research will be using a census data in 2010 since the Malaysia Housing Survey was carried out per 10 years only and current Census Data 2010 is considered as the latest data. New household formation later will be identified by calculating the headship rate based on census time series data on 2010. By definition, the percent change in household formation equals the sum of the percent growth in population and the percent change in headship rates which the key drivers of household formation over time are population growth and the evolution of headship rates (Furlong, 2016; Zeng et al., 2013). Headship rates used to help developer and planners to determine the housing needed by the household (Siniavskaia, 2018). Therefore, this study will apply headship rate multiply with the population as the new household formation for all Penang district.

.

3 RESEARCH FRAMEWORK

3.1 Penang population, selection of research area and sampling size

The research focused on Penang due to its high and rapidly urbanized area compared to other states in Malaysia. Based on the record, Penang population had increased from 0.9% in 2014 and expected growth to 1.3% in 2018. In the year 2018, Penang State estimated population reaching 1.76 million (Penang Institute, 2018).

The research was carried out in 5 districts in Penang include; 1) Penang-Barat Daya, 2) Penang-Timur Laut, 3) Penang-Prai Utara, 4) Penang-Prai Tengah and 5) Penang-Prai Selatan. The total of respondents is 400 persons include various ethnic in Penang which are Malay, Chinese, Indian and others. Generally, this survey was used the Sample Size Determination Table that has been prepared by Krejcie and Morgan (1970) to determine the sample size. Based on the total Penang population, a few sample sizes should be 384 samples but, in this research, the total of 400 samples has been prepared to distribute into five districts in Penang.

3.2 Headship rate and Indicators validation

Willekens and Van Imhoff (2015) suggested to obtain future households we will need the data on the future population and future headship rate. The headship rate method also provides the projections in useful detail (Holmans, 2012). By considering the headship rate, the household formation can be forecasted accurately based on their demand (Ibrahim, Zainun, & Rawan, 2017).

$$\text{Headship} = \frac{Head}{Population} \times 100 \qquad (1)$$

For example, a distribution of future heads of households by sex and age is needed for economic and social planning purposes and only the headship rate method can afford to provide such information (Iv, 1957). In this research, data in the year 2010 were used as the baseline and is considered as the latest census data. Using **Equation 1**, the headship rate in Penang for the year 2010 are shown in **Figure 3-1.**

Headship rate equation is as below;

$$HR_i = \frac{HH_i (H_1 + H_2 + H_3 \ldots\ldots H_i)}{P_i (P_1 + P_2 + P_3 \ldots\ldots\ldots P_i)} \qquad (2)$$

Where HR_i the number of household's head in the age group and P_i is the total population in the age group.

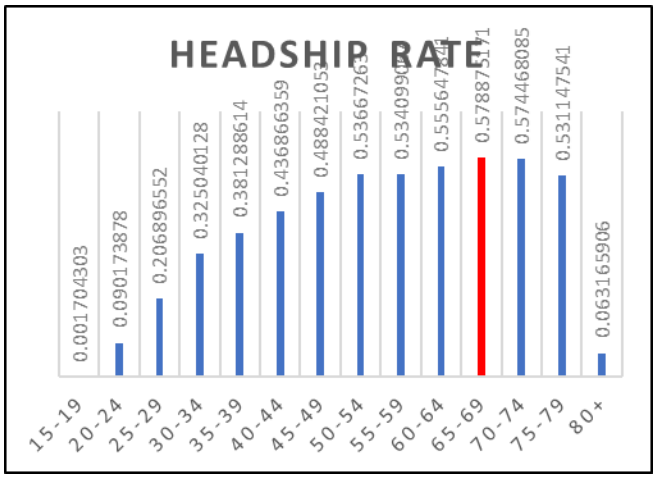

Fig 3.1: Age-specific headship rate in Penang (2010)

Based on the figure, the highest headship rate was from the age group of 65-69 years old. Meanwhile, the lowest headship rate was from age group 15-19 years old.

According to Alias et al., (2018), 10 indicators of housing demand is confirmed by expert housing players as well as economist in Malaysia includes; 1) number of person, 2) household number, 3) migrant status, 4) education attainment, 5) citizenship, 6) ethnic group, 7) marital status, 8) household income, 9) age of household head and 10) tenancy.

The ten indicators had been tested using corrected item correlation to indicate the validity using Cronbach's Alpha Table. The validation of the indicators shows the value range within 0.34 to 0.83. It is in line with Nunnally and Bernstein (1994), decides that the indicators used are valid if the corrected item-total correlation is greater than 0.25. Furthermore, the acceptable indicators are considering value alpha coefficient of 0.7 and the reliability test using Cronbach's Alpha indicate a value of 0.828. So that, the ten indicators were considered acceptable.

From these ten indicators, the result showed only two numbers of indicators were significant which are 1) Age of the head of family or Headship and 2) Household income, where both indicators obtained significant (sig) value less than 0.05.

3.3 Compatibility Analysis and Performance Model

Model suitability analysis of data is determined by goodness-of-fit and the relationship between dependent and independent variables in this research were tested using model fitting information and the result is shown in **Table 3.0** and **Table 3.1**

Table 3.0: Goodness-of-fit

	Chi-Square	*df*	*Sig.*
Pearson	283.936	474	1.000
Deviance	213.991	474	1.000

Table 3.1: Model fitting information

Model	*Model Fitting Criteria -2 Log Likelihood*	*Likelihood Ratio Tests*		
		Chi-Square	*df*	*Sig.*
Intercept Only	678.720			
Final	332.830	345.891	6	0

(Chan, 2011; Madhu, Ashok, & Balasubramanian, 2014) and (Alias et al., 2018) agreed that the model is determined according to the data used when the p-value (Sig.) obtained through the *goodness-of-fit* analysis is exceeds 0.05. Whereas for *model fitting information*, the required value of p is less than 0.05. Based on both Table 3.0 and Table 3.1, it can be stated that the relationship between dependent and independent data was supported (Bayaga, 2001) and used is in accordance with the model produced.

Table: 3.2: Classification Table

Observed	Predicted				
	Low	Low Medium	Medium	High	Percent Correct
Low	23	69	5	0	23.7%
Low Medium	17	76	24	0	65.0%
Medium	2	15	163	1	90.1%
High	0	0	4	1	20.0%
Overall Percentage	10.5%	40.0%	49.0%	0.5%	65.8%

To know the model's performance in predicting, the results obtained through the classification table (percent overall) are shown in **Table 3.2** The housing category is compared to the proportional by chance accuracy criterion that obtained through marginal percentages in each household category (El-habil, 2012).

Table 3.3: Case Processing Summary

Category of Houses	**N**	**Marginal Percentage (%)**
Low	97	24.3%
Low Medium	117	29.3%
Medium	181	45.3%
High	5	1.3%
TOTAL	400	100.0%

Based on the results obtained, the overall percentage in **Table 3.3** was greater than the proportional by chance accuracy criterion (Table 3.2) with a value of 65.8% > 43.53%. So, this model is expected to predict well for the all four categories of housing in Penang. The calculations for the proportional by chance accuracy criterion are shown below where 1.25 or a 25% improvement over the rate of accuracy achievable by chance alone as the benchmark for the MNL model (El-habil, 2012).

Marginal Percentage equation is as follow;

$$1.25*[(0.242^2+0.292^2+0.452^2+0.012^2)*100\%] = 43.53\% \quad (3)$$

3.4 Housing Choices Possibilities (CP) for all categories (low cost, low medium cost, medium cost and high cost) in Penang

The purpose of CP Model is to identify the choice of probabilities value by respondents based on their house selection for the four housing categories in Penang (Chan, 2011). Choice of probabilities for four housing categories in Penang is determined using Multinomial Logit (MNL) which are adopted via IBM SPSS 20 software. To run the MNL analysis, a total number of 400 samples from field housing survey questionnaire were keyed-in into the SPSS software. Over the MNL analysis, only the appropriate and significant variables are used in determining the CP value. By using the MNL analysis results, the equations of the CP model are shown below;

$$\log (Pi/Pr) = \beta 0 + \beta 1 + \beta 2 + \beta 3 + \ldots\ldots. \beta n \quad (4)$$

Where Pi is the input parameter, Pr is probabilities of the decision maker and $\beta 0$ is the value of intercept and $\beta 1$ until βn are the value of the significant variable. By replacing the variables involved in **Equation (3.3)**, the general equation for the housing category selection model (g_i) is shown in **Equation (3.4)**. The logistic equation (g_i) of the dependent variable having the n category can generally

be written as Equation

$$g_i = \left(\frac{P(Category,\ i_i)}{P(Category,\ i_j)}\right) = \beta_{io} + \beta_{i1}X_1 + \beta_{i2}X_2 + \beta_{i3}X_3 + \ldots\ldots \beta_{in}\,X_n \quad (5)$$

where;
β_{i0} = constant
β_{in} = Constant independent variable
X_1 = household income
X_2 = age

4 Result and Discussion

As a result, the data from housing survey or field questionnaire preference by 400 respondents in all district in Penang were analyzed. Multinomial Logit analysis was carried out to determine the choice of probability on house type selection and to produce a Choice Probabilities (CP) Model. MNL analysis was used to classify the most significant variable and to determine the choice of probabilities for the housing preference.

Table 4.0: Respondent feedback on the type of housing preferences and CP value.

Type of housing	Respondent	Choice of probability
Low cost (LC)	41	0.1025
Low medium cost (LMC)	162	0.4050
Medium cost (MC)	195	0.4875
High cost (HC)	2	0.0050
TOTAL	**400**	**1.000**

In **Table 4-0,** the result showed that Medium Cost housing was the most preferable type of house with the highest CP value of 0.4875.

4 Summary

Total housing demand in Penang for low cost, low medium cost, medium cost, and high-cost housing was obtained based on MNL analysis and CP Model. As a result, the data from housing survey or field questionnaire preference by 400 respondents in all district in Penang were analyzed. MNL analysis was used to classify the most significant variable and to determine the choice of probabilities for the housing preference. Medium Cost type housing is shown high probabilities which is at the value of 0.4875. Therefore, it is anticipated that the research could be use by related agencies such as developer or any other relevant government agencies as a basis in making their development planning for housing sector or industry especially in Malaysia towards the future sustainable development.

5 Acknowledgement

This research is fully supported by MOSTI. The authors fully acknowledged Ministry of Higher Education (MOHE) and Ministry of Science, Technology and Innovation for the approved fund for which makes this important research viable and effective.

6 References

[1] Alias, A. R., Zainun, N. Y., & Rahman, I. A. (2018). Headship Rate Projections for Housing Demand in Headship Rate Projections for Housing Demand in Johor ,.

[2] Bayaga, A. (2001). Multinomial Logistic Regression: Usage and Application in Risk Analysis. *Context*, *5*(2), 288–297.

[3] Chan, Y. H. (2011). Multinomial logistic regression. *Singapore Medical Journal*, *46*(6), 259–269. https://doi.org/10.1016/j.cose.2005.05.003

[4] Economic Planning Unit. (2015). *Eleventh Malaysian Plan (2016-2020)*. *Eleventh Malaysian Plan*. https://doi.org/10.1017/CBO9781107415324.004

[5] El-habil, A. M. (2012). An Application on Multinomial Logistic Regression Model. *Journal of Statistics and Operation Research*, *8*(2), 271–291. https://doi.org/10.18187/pjsor.v8i2.234

[6] Furlong, F. (2016). *Household Formation among Young Adults*.

[7] Holmans. (2012). *Household Projections in England : their history and uses*.

[8] Hong, T. T. (2016). Affordable Housing for First-Time Homebuyers : Issues and Implications from the Malaysian Experience, *5921*(March). https://doi.org/10.1080/14445921.2013.11104381

[9] Ibrahim, I. S., Zainun, N. Y., & Rawan, N. M. (2017). Head of Households in Terengganu, *3015*, 1–6. https://doi.org/10.1051/matecconf/201710303015

[10] Iv, C. (1957). Headship Rate Method, 31–40.

[11] Kamal, E. M., Hassan, H., & Osmadi, A. (2016). Factors Influencing the Housing Price: Developers' Perspective. *World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, *10*(5), 1656–1662. https://doi.org/scholar.waset.org/1999.10/10004541

[12] Krejcie, R. V, & Morgan, D. (1970). Determining Sample Size for Research Activities, (Education and Psychological measurement), 607–610.

[13] Macdonald, S. (2011). Supply and demand in the Penang housing market : Assessing affordability October 2011, (October).

[14] Madhu, B., Ashok, N. C., & Balasubramanian, S. (2014). A Multinomial Logistic Regression Analysis to Study the Influence of Residence and Socio-Economic Status on Breast Cancer Incidences in Southern Karnataka. *International Journal of Mathematics and Statistics*

Invention, *2*(5), 1–8.

[15] Munir, S., Azlan, O., Fauziah, M. T., Ishak, I., Hasnah, H., Noor, H. A., & Wan Mohd Muhiyuddin, W. I. (2010). The State of Penang , Malaysia. *The State of Penang, Malaysia: Self-Evaluation Report*, (201 0), 107–120.

[16] NAPIC. (2018). *Property Market Report: First Half 2018*.

[17] Penang Institute. (2018). *Penang Statistics Quarter 3, 2018*. Retrieved from https://penanginstitute.org/resources/data-centre/97-quarterly-penang-statistics/

[18] Siniavskaia. (2018). Geography of Declining Young Adult Household Formations. Retrieved from http://eyeonhousing.org/2018/07/geography-of-declining-young-adult-household-formations/

[19] Willekens, F., & van Imhoff, E. (2015). *Families and Households, Formal Demography of*. *International Encyclopedia of the Social & Behavioral Sciences: Second Edition* (Second Edi, Vol. 8). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.31016-9

[20] Zeng, Yi, Lan Li, Zhenglian Wang, Helin Huang, and J. N. (2013). Effects of Changes in Household Structure on Future Housing Demand in Hebei Province , China, (2), 85–111.