CHAPTER 3

METHODOLOGY

Contents within this chapter consist of four sections. The estimated model used in this study is described first. Then additional assumptions are made so as to adjust the model to work-well with a Bangkok case study. Later what the expected signs are and how to interpret the estimated results are explained. Finally, data is identified and variables are discussed.

3.1 The Estimated Model

Although a lot of empirical research on housing market has involved the estimation of hedonic price function, a little information about the effect of housing characteristic on consumer's choice of residence has been conveyed.[1] As interpretation of the hedonic theory, by focusing the alternative on one housing characteristic at a time, it obscured a virtue of hedonic approach (its ability to treat housing characteristics simultaneously). Thus the standard hypotheses of urban economics, translated into statements about bid-rent function, seem most reasonable when interpreted ceteris paribus. Urban economists have explained that it is more reasonable to ask how consumers will react to these housing characteristics dealing with bid-rent curves rather than holding housing characteristics fixed and ask which type of consumer would be willing to pay for as in hedonic approach involved utility function. For a given consumer, utility need not vary monotonically as a function of an attribute. For example, one consumer may value a house close to the central business district more highly while another prefers to live in the suburbs whether or not his/her income increases.

[1] Read more detail about hedonic price theory in Ellickson. (1981). An Alternative Test of the Hedonic Theory of Housing Market. Journal of Urban Economics. 9, 56-79.

Ellickson (1981) has developed this alternative approach, which is more in tune with the view that houses are indivisible commodities and have a powerful role on residential choice. This approach connected to econometric estimation which is essentially within the framework of multinomial logit model in reverse. However, it provides advantages that are linked between the logit equation and hedonic theory involving bid-rent rather than utility function, permits the empirical results to be given an extremely clear interpretation residential choice.

Since this approach is easily confused with the model we will propose, it is worthwhile outlining the main ideas behind this use of logit. Assuming that every household in a particular housing market has tastes that can be described by a utility function

$$U_i(h_i, z_i), i \in I \tag{3.1}$$

where z_i is a *Z*-dimensional vector of private goods, h_i is a *H*-dimensional vector of housing attributes, and *I* indexes the set of households where $i \in I$. The utility function is assumed to be quasi-concave and twice continuously differentiable in z_i. The budget constraint for consumer *i* is

$$p_z z_i + p(h_i) + K_i(h_{i1}) = y_i \tag{3.2}$$

where p_z is a *Z*-dimensional vector of private good prices, $p(h_i)$ is the hedonic price function relating the price function of a dwelling to its characteristics, $K_i(h_{i1})$ is the transportation cost function for household *i* (assumed, for simplicity, to depend only on the first characteristic, which should be interpreted as the distance to the central business district)[2], and y_i is the income of consumer *i*.

It is useful to model the consumer's choice process in two stages. In the first stage the consumer maximizes his utility function (3.1) subject to the budget constraint (3.2) with h_i held fixed. The solution to this problem of constrained maximization can be represented by an indirect utility function

$$\phi_i\left[p_z, h_i, y_i - p(h_i)\right] \tag{3.3}$$

giving the maximum utility that the consumer can achieve at prices p_z if he is residing in a dwelling with characteristics h_i.

---

[2] In previous theory, we define *l* is the distance from home to the CBD, while by generally, distance is treated as a characteristic of housing (housing's location) in hedonic approach.

Now suppose a consumer of type *i* has indirect utility function in the form $\tilde{\phi}_i\left[h, p(h)\right]$, where we have suppressed the price vector p_z (assumed invariant throughout the metropolitan area) and the income parameter y_i (since households of the same type have the same income). To translate the model into a form suitable for econometric estimation, this deterministic indirect utility function is replaced by a stochastic indirect utility function:

$$\tilde{\phi}_i\left[h, p(h)\right] + \varepsilon_h \tag{3.4}$$

where ε_h is a random variable associated with a house of type *h*. The probability that a consumer of type *i* will choose a house of type *h* is then given by

$$p(h/i) = prob\left\{\tilde{\phi}_i\left[h, p(h)\right] + \varepsilon_h > \tilde{\phi}_i\left[h', p(h')\right] + \varepsilon_{h'}; h \neq h', h' \in H\right\} \tag{3.5}$$

The specific form of the discrete choice model is determined by assuming distribution of an error term. If the random variables $\varepsilon_h, h \in H$, are independently and identically distributed Weibull[3], as shown by $F(\varepsilon) = \exp[-\varepsilon - \exp(-\varepsilon)]$. Then the probability can be rewritten as

$$p(h/i) = \frac{\exp\left\{\tilde{\phi}_i\left[h, p(h)\right]\right\}}{\sum_{h' \in H} \exp\left\{\tilde{\phi}_i\left[h', p(h')\right]\right\}} \tag{3.6}$$

If the indirect utility functions are linear, we obtain

$$p(h/i) = \frac{\exp\left[\beta_i h + \gamma_i p(h)\right]}{\sum_{h' \in H} \exp\left[\beta_i h' + \gamma_i p(h')\right]} \tag{3.7}$$

a model whose parameters can be estimated via maximum likelihood.

What the approaches described above have in common is the attempt to find an analog within the context of hedonic theory of the demand functions central to conventional microeconomic theory. Nevertheless, the most natural way to interpret such models is in terms of a prediction of what sort of consumer is most likely to occupy a house with a specified set of characteristics. The house will be occupied by the consumer offering the highest bid price. Thus the traditional accessibility models

[3] McFadden (1974) proved that the multinomial logit model can result if and only if the errors are independent and have a type I extreme-value or Weibull distribution.

predicts that houses located far away from the central business district will be occupied by household with low marginal commuting costs and relatively high demand for housing space.

Thus for purposes of empirical estimation we set a stochastic bid price function of the form

$$V_i = \tilde{\psi}_i(h) + \varepsilon_i, i \in I \qquad (3.8)$$[4]

where we interpret $\tilde{\psi}_i$ as the bid price function of a representative household of type *i* and ε_i is a random disturbance term reflecting differences in income and tastes among households of the given type. To determine the probability that a given house will be occupied by a household of type *i*, the relevant variables are the maximum bids from consumers of each type:

$$V_i^* = \max_{i \in I} V_i = \tilde{\psi}_i(h) + \varepsilon_i^*, i \in I \qquad (3.9)$$

where $\varepsilon_i^* = \max_{i \in I} \varepsilon_i$ If the ε_i^* are identically and independently distributed Weibull, then we obtain a logit model. Note that the Weibull distribution is adopted on an ad hoc basis because it leads to a convenient estimator. In the present context the Weibull distribution emerges endogenously through the process of choosing the highest bids. Assume that the random variables $\varepsilon_i, i \in I$, are i.i.d. with a normal distribution. Thus ε_i^*, the largest value among the ε_i, will have the Weibull distribution. Then the probability that a house with characteristics *h* will be occupied by a household of type *i*,

$$p(i/h) = prob\left\{\tilde{\psi}_i(h) + \varepsilon_i^* > \tilde{\psi}_{i'}(h) + \varepsilon_{i'}^*; i' \neq i, i' \in I\right\} \qquad (3.10)$$

will take the form

$$p(i/h) = \frac{\exp\{\tilde{\psi}_i(h)\}}{\sum_{i' \in I} \exp\left\{\tilde{\psi}_i{}'(h)\right\}} \qquad (3.11)$$

Assuming that the bid price functions are linear we obtain

$$p(i/h) = \frac{\exp(\alpha_i h)}{\sum_{i' \in I} \exp(\alpha_{i'} h)} \qquad (3.12)$$

---

[4] Bid price is the solution to constrained maximization through indirect utility function. Read more details in Ellickson (1981).

The parameters of this model can be estimated through maximum likelihood.

A comparison of equation (3.7) and equation (3.12) should make clear firstly how these two model differ from each other. The dependent variable in (3.7) gives the probability that a consumer of type *i* will select a house of type *h* while the dependent variable in equation (3.12) gives the probability that a house with characteristics *t* will be occupied by a consumer of type *i*. The parameters estimated in (3.7) represent the coefficients of the indirect utility function for consumer *i* while in equation (3.12) the parameters are coefficients of the bid price function for consumer *i*. The advantage of the formulation represented by (3.7) is the ease with which it can be used to test various propositions in the urban economics literature regarding the effect of such factors as transportation system, accessibility, filtering, racial discrimination or jurisdictional fragmentation on the residential choices of households of different types.

Following the theoretical concept of LeRoy and Sonstelie (reviewed in chapter 2), the effect of transport innovation as an introduction of the alternative choice of transit on residential choice of different type of household has been developed by Gin and Sonstelie (1992). Within a framework of reversed multinomial logit model, housing characteristics involved only residential location, l where $l \in L$. Given that the independent variable can be continuous and/or discrete types, the unknown parameter can be performed by maximum likelihood estimation (MLE). Let us define the observed income group as n_{il}, which take a value of one if location l is occupied by income group i and zero otherwise. Assuming that all income groups are independent[5], the joint distribution of sample is given by the product of the individual density function. Therefore, the likelihood function for a general multinomial choice model is

$$\mathbb{L} = \prod_{l=1}^{L} \prod_{i=1}^{I} P_l(i) \qquad (3.13)^6$$

[5] Independent of irrelevant alternative assumption is described later.

[6] Redenote $p(i|l)$ by $P_l(i)$ which refer to probability that location l is occupied by income group i, as follows;

$$P_l(i) = \frac{\exp(\alpha_i l)}{\sum_{i' \in I} \exp(\alpha_{i'} l)}.$$

Substituting $P_l(i) = \frac{\exp(\alpha_i l)}{\sum_{i' \in I} \exp(\alpha_{i'} l)}$ [7] into equation (3.13) and taking logarithm of the likelihood function, we obtain

$$\ln \mathbb{L} = \sum_l \sum_i n_{il} \left\{ \exp(\alpha_i l) - \ln \sum_i \exp(\alpha_{i'} l) \right\} \tag{3.14}$$

Maximum likelihood procedure is enabling to deal with the sample data set that contains individual characteristics rather than aggregate sample. The iterative search process is undertaken to find the estimate that maximize the likelihood of attaining the true parameter. The maximum likelihood procedure yields the odds of alternative income groups occupying a location rather than the base group, which can be interpreted in terms of the predicted probability. More precisely, the estimation of multinomial logit based on maximum likelihood method yields the odds of outcome *i* versus outcome *j* as follows

$$\frac{P_l(i)}{P_l(j)} = \frac{\frac{\exp(\alpha_i l)}{\sum_{i \in 1}^{I} \exp(\alpha_i l)}}{\frac{\exp(\alpha_j l)}{\sum_{i \in 1}^{I} \exp(\alpha_{i'} l)}} = \frac{\exp(\alpha_i l)}{\exp(\alpha_j l)} \tag{3.15}$$

Taking logs on equation (3.15), we obtain linearity in the logit as follows

$$\ln \left[\frac{P_l(i)}{P_l(j)} \right] = (\alpha_i - \alpha_j) l \tag{3.16}$$

which allows us to interpret the estimates coefficient as the change in logit is expected to change by $(\alpha_i - \alpha_j)$ units, for every unit change in l other things remain constant. In particular, they are the log odds of being in income group *i* versus income group *j*. However, the log odds ratio can help us understand the result more clearly. To get the odds ratio, we just take exponential log on the estimated parameter in equation (3.16); $\exp(\alpha_i - \alpha_j) = \alpha_{ij}$. This means that one additional unit of explanatory variable, *l*, multiplies the odds of being in income group *i* rather than group *j* by α_{ij}. In other

---

[7] Recall in equation (3.12) which only housing location, *l*, is treated as explanatory variable.

words, the odd ratios indicate that a unit changes in *l* increases the probability of being in income group *i* instead of group j by α_{ij} times.

To illustrate how the alternative choice of transit affects residential location patterns, we simply extend equation (3.16). Therefore equation (3.17) will be our basic estimation equation.

$$\log\left(p_{it}/p_{jt}\right) = (\alpha_i - \alpha_j) + (\beta_i - \beta_j)dist_t + \gamma_i\left(difdist\right)x_{it} \qquad (3.17)$$

where p_{jt} is the probability that a member of group j will occupy location t and suppose group i has a higher income than group j

$dist_t$ (or l_t) is the Euclidean distance from residential location *t* to the CBD

$difdist$ (or $l_t - l_{it}^*$) is the difference in Euclidean distance and break-even distance of income group *i*

x_{it} is dummy variable taking value unity if Euclidean distance at location *t* is greater than break-even distance of income group *i* that resides on location *t*, and takes value zero otherwise, and

$(\alpha_i - \alpha_j)$, $(\beta_i - \beta_j)$, and γ_i are estimated parameters.

Now, note that group j is the reference group.

For explicitly understanding adoption of the original model adopted by Gin and Sonstelie (1992) which is employed for this study, it is worthwhile outlining our additional assumptions to be used throughout this study

1.) Assumption made on Central Business District (CBD)

According to O'Sullivan (2000)[8], monocentric city is defined as a city in which commercial and industrial activity is concentrated in the central core area and there exists only one central core area which is called Central Business District (CBD)[9]. In general, CBD is defined as the area over which high employment concentration, high economic activity, or high population concentration occurs.

---

[8] O'Sullivan, A. (2000). Urban Economics. The forth edition, Boston: Irwin & McGrew Hill.

[9] Pinto, S. (1994). has studied location pattern of office activity in Bangkok can be confirmed the monocentric city of Bangkok.

However, in this study, I will define CBD as the area with high employment concentration. Since this study is based on a role of the alternative transportation mode as rapid rail transit that plays on residential location patterns in Bangkok, an individual tradeoff between cost of housing and cost of commuting to CBD involves working trip (assume that all individuals are working in CBD).

According to Wisaweisuan (2001), Bangkok has no readily defined CBD however, for several urban studies, (Pinto; 1994, Dawcharoen; 1996) Silom has been designated as the CBD due to its distinctive characteristics such as a steady and high employment concentration, a high percentage of economic activity and a small amount of vacant land. Therefore this study employed the Silom area as the CBD, not only because of its characteristics, but also for the advantage of comparability with different period of study. To answer “does availability of the alternative transport mode, rapid rail transit, have an influence on residential location patterns in Bangkok”, and “how so”, I will compare the estimated result before and after the introduction of this new transport mode applying the data from two different periods, 1998 and 2004. Thus, keeping the same city feature will be better for this study. Moreover, considering the influence of the introduction of the rapid rail transit, its network system passes through the Silom area at Silom station thus assuming Silom as CBD is still sufficient for the period of post rapid rail transit being mode available. Following the study of Kasemsook (2003), CBD area is exactly defined as the areas of Khwang Surawong and Khwang Sathon surrounded by Surawong Road to the north, Rama IV Road to the east, Sathon Neau Road to the south, and Mahyesak and Surasak Roads to the west (Depicted as Figure 3.1).

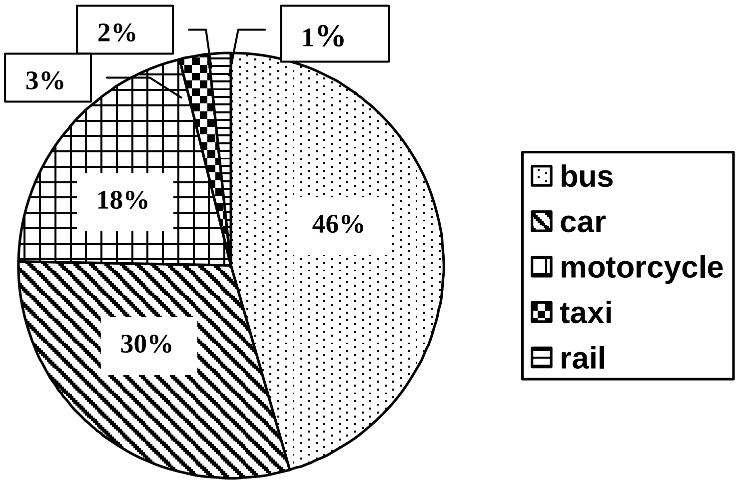
Figure 3.1

Central Business Area of Bangkok

Source: Bangkok Metropolitan Administration

To consider how transport facility innovation plays a role on residential location patterns in Bangkok, data from 1998 and 2004 are employed. In the period before the rapid rail became available, buses and automobiles were the main transport mode for people living in Bangkok in which buses were the original mode and automobiles were the alternative transport mode according to history of Bangkok transportation (Kamanamul; 2004)[10]. As depicted in Figure 3.2, buses and automobiles were the two most popular modes of transportation with 46 percent and 30 percent for buses and automobiles respectively, while the remaining 24 percent included all other modes such as motorcycles, taxis, and rail.[11]

Figure 3.2
Ridership Share for Transportation
Mode in Bangkok 1998



Source: OCMLT Final Report (1998)

For the period after the rapid rail transit became available, I employ data from 2004 (about five years after this new alternative mode was first introduced). Even though the two main modes were still buses and automobiles, the existence of the rapid rail transit made the ridership structure gradually change. The share of ridership of rapid rail increased overtime while the bus transit gradually declined. This conformed to the study of ICRA (2001) which stated that 75 percent of demand for rapid rail transit was found to come from previous bus users while the remaining 25 percent was from car users. However, the heavy traffic congestion in Bangkok

[10] Bus Transit was introduced for Bangkok in 1960 at first time, and then automobile was firstly introduced in Bangkok in 1965 (Kamanamul, 2004).

[11] Rail in this figure means SRT, the standard heavy commuter rail that has been concession for operation by State Railway of Thailand.

stimulated car commuters to seek alternative transportation, making it impossible to ignore the competition between car and rapid rail transit.

Therefore, in 2004, I will separately consider three pairs of competitive transport modes which are 1) bus transit, the initial transport mode, and automobile transit, the alternative mode, 2) bus transit, the initial transport mode, and rapid rail transit, the new alternative mode and 3) the initial automobile transit and the new alternative rapid rail transit. By assuming that the residents who previously commuted by bus transit is mutually exclusive from the residents who previously commuted by car transit.[12]

3.) Assumption made on individual household

I assume all households are identical except for their income since the initial assumption of this concept made on the income elasticity for housing demand relative to the elasticity for marginal cost of commuting. Moreover, dividing all residents by different incomes can also reflect different transport mode choices based on their income level. Different income plays a differing role on transport mode choice representing the value that each group places on their commuting time as a function of their income as an alternative forgone. But the value placed on commuting time might not be equal to their income. In this study, the value of commuting time is assumed to be a function of income and follow the same pattern for all groups. Value of time is assumed to be 0.49 of their average income in accordance with the study of Dissanayake (2005)[13] for 1998 and 2004 (keeping the value of commuting time constant overtime as in base year 1998, then income difference allow to play its role).

4.) Assumption made on the reference group for multinomial logit model

In the context of multinomial logit model, parameters are estimated in terms of log odd scale. It provides odd of probability that the alternative income group occupy a location rather than the base group. Thus the base group used for estimation should be assumed first. For this study, the reference group in the period before the rapid rail transit became available is assumed to be similar to Gin and Sonstelie that is the lowest income group who never choose the alternative car transit over the original

---

[12] The specific technical assumption is required in order to perform bimodal choice of transit model.

[13] Dissanayake, D., and Takayuki M. (2003). A Combined RP/SP Nested Logit Model of Vehicle Ownership, Mode Choice and Trip Chaining to Investigate Household Travel Behavior in Developing Countries (the case of Bangkok). The 82th Annual Meeting of TRB.

bus transit, will be used as the reference group as well. However, for the period after the rapid rail transit was introduced, the lowest income group still plays that role. The lowest income group is treated as the reference group for both parts of the estimation. Namely, when we firstly consider the residential location pattern that is affected by the coming of the new alternative transport mode, rapid rail transit, which bus transit is treated as the original mode, the lowest income group will never prefer rapid transit to bus transit since it generates much more monetary cost. Secondly, even though, I estimate the influence of the existence of rapid rail transit again, but now assuming the original mode is car transit and the new alternative mode is rapid rail transit, the reference group will still be the lowest income group as the higher income groups are much more sensitive to the alternative mode which can save their value commuting time (note that values that each household place on commuting time is assumed to be a proportion of their income).

3.2 <u>Hypothesis Setting and Expected Signs</u>

According to the theoretical concept, the expected sign of estimated coefficients and hypothesis setting are mentioned. Since estimated function is log-linear, its gradient is the percentage change in the difference in bid-rent gradient of the alternative income group relative to the reference group per unit change in distance. The gradient takes two values depending on whether the distance from l to the center is less than the break-even distance for group i and group j. Since two competing transport mode choices exist, the initial transit and the alternative transit, the break-even distance of group i will be closer to the center than of group j [14].

[14] Break-even distance of group i and group j (where $i > j$) are written as

$$l_i^* = \frac{f^A - f^I}{(c^I + w_i t^I) - (c^A + w_i t^A)} \quad \text{and} \quad l_j^* = \frac{f^A - f^I}{(c^I + w_j t^I) - (c^A + w_j t^a)}.$$

since $w_i > w_j$, thus the break-even distance of i will be closer to CBD than of j.

In the first zone where distances less than or equal to break-even distance of group i and j, both groups adopt the initial transit then dummy variable x is zero for both groups. According to analysis of Gin and Sonstelie (1992) the alternative mode means the transportation mode that is slower but cheaper while the alternative mode means the transport mode with faster speed but higher monetary cost. The difference in bid-rent gradient of group i relative to group j is $\beta_i - \beta_j$, which is the percentage change in difference in bid-rent gradient per unit change in distance, assuming the representative member of both groups commute by the initial transit since adopting the alternative mode is economical only for long traveling distance. Recalling equation (3.15) and (3.16), it is noticed that β_i represents bid-rent gradient of income group i and β_j represents bid-rent gradient of income group j if both adopt the initial transit. If $\beta_i - \beta_j < 0$, group i's bid-rent function is steeper than group j in this region and the frequency of group i relative to group j declines with distance. On the other hand, if $\beta_i - \beta_j > 0$, group j's bid-rent function is steeper than group i in this region and the frequency of group i relative to group j increases with distance. The theory of urban residential location outlined in Section 2.1 implies that β should be negative since bidding rent for a housing unit decreases as distance increases.

For the second zone, locations between the break-even distances of the two groups, i commute by the alternative transit and j commute by the initial transit, then dummy variable x is unity for group i while zero for group j. The difference in bid-rent gradient is $\beta_i + \gamma_i - \beta_j$, which is the percentage change in difference in bid-rent gradient per unit change in distance, assuming the representative member of group i commute by the alternative transit while representative member of group j still adopt the initial transit. Now $\beta_i + \gamma_i$ represents bid-rent gradient of income group i if the alternative transit is adopted, and β_j represents bid-rent gradient of income group j if they remain using the initial transit. Notice that γ_i should be positive since, beyond the break-even distance of group i, an individual adopted the alternative transit which reduces commuting cost as time saving then it flattens bid-rent slope of group i. So the frequency of group i relative to group j declines with distance if and only if $(\beta_i + \gamma_i) - \beta_j < 0$, which bid-rent slope of group i, $(\beta_i + \gamma_i)$, is still greater than of group j, β_j, when only group i adopts the alternative transit. Contrarily, if $(\beta_i + \gamma_i) - \beta_j > 0$, bid-rent slope of group i is less than of group j and the frequency of group i relative to group j increase with distance when group i adopt alternative transit and group j adopt initial transit.

Finally, in the third zone for distances beyond the break-even distance of the lower income group, then dummy variable x for both income group take value unity. The difference in bid-rent gradient is $(\beta_i + \gamma_i) - (\beta_j + \gamma_j)$, where $(\beta_i + \gamma_i)$ is the bid-rent gradient of group i and $(\beta_j + \gamma_j)$ is gradient of group j if both groups adopt the alternative transit. Thus the frequency of group i declines relative to group j if and only if $(\beta_i + \gamma_i) - (\beta_j + \gamma_j) < 0$, namely bid-rent slope of group i, $(\beta_i + \gamma_i)$, is still greater than of group j, $(\beta_j + \gamma_j)$ if both adopt alternative transit. Again, in contrast, if $(\beta_i + \gamma_i) - (\beta_j + \gamma_j) > 0$, bid-rent slope of group i is less than that of group j and the frequency of group i increases relative to group j with respect to distance when both groups adopt alternative transit.

The differences in bid-rent gradients are illustrated in Table 3.1. The upper half of Table 3.1 shows differences in bid-rent gradients in the first zone when both groups adopt the initial transit. And the bottom half of Table 3.1 shows differences in bid-rent gradients in the second zone, between break-even distance of the other higher income group and the reference group (the 1st group). Note that, the lowest income group is set as the reference category; it never chooses the alternative transit over the initial transit. Thus dummy variable x of the 1st group is always zero, which leads γ_1 to disappear. However the other lower income groups can adopt the alternative transit if it contributes less commuting cost relative to the initial transit. Therefore the differences in bid-rent gradients of the each pair of competitive income group (3rd-2nd, 4th-2nd, 4th-3rd, 5th-2nd, 5th-3rd, and 5th-4th) in the bottom half of Table 3.1 are obtained from the third zone.

The alternative transit can cause a fundamental change in residential pattern in the second and the third zone. Even if frequency of the higher income group can be declined in the first zone as $\beta_i - \beta_j < 0$, it may increase with distance in the second and the third zone if γ_i is sufficiently large. High value of γ_i leads to

$(\beta_i + \gamma_i) - \beta_j > 0$ or $(\beta_i + \gamma_i) - (\beta_j + \gamma_j) > 0$, which yields opposite sign from the difference in the first zone. This implies that the rich tend to live on more distant areas and commute by the alternative transit due to its time cost saving. Note that it assumes fixed and variable monetary costs of the alternative transit greater than of the initial transit while time cost of the alternative transit is less than that of the initial transit. It is obvious that the alternative transit is competing with the choice of the initial transit that the commuter must decide whether to choose monetary cost saving of the initial transit or time saving of the alternative transit. The key of the model is that the value that residents place on commuting time varies proportionally with income level. It may persuade the rich to prefer the alternative transit as time saving advantage. Thus the rich are more likely to choose the alternative transit than the poor for two reasons; 1) the alternative transit provides much more time cost saving as income increases, 2) the rich can afford to pay higher commuting costs of the alternative transit, in which monetary cost increased can be offset by its time cost saving.

Further most past studies supported that the effect of automobiles yielded decentralization of the rich. Due to time cost saving of automobile commuting, the rich moved to the fringe of the city which can occupy the bigger house with less housing cost than in the central areas. Thus the estimated coefficient for the first variable "*dist*" should be negative and for the second variable "*(difdist)x*" should be positive. That leads to a difference in bid-rent gradient, when all income groups commute by the initial transit, is negative, while the difference in bid-rent gradient, when only the high income groups commute by the alternative transit, is positive. However the effect of rapid rail transit is still vague. This facility induced bid-rent gradient was far shallower with distance since it reduced disutility of commuting distance and/or commuting time. Many studies conducted in American and Europe cities attempted to capture the effect of the rapid rail transit for regaining high-class people back to the old city center. However the key of regaining the rich back was that providing an alternative transit can reduce commuting costs enough for the poor to switch to this transport mode and move-out of the city center, which in turn would have them competing with the rich for residential areas on the peripheral side. Therefore the rich who lost his/her comparative advantage living on peripheral areas as housing cost on that location were raised by the competitor. Thus some rich move back downtown. However the empirical evidence showed that the effect of rapid rail transit was not powerful enough to return the rich back down town, since the poor were not willing to switch to rapid rail transit and live on outer areas. For Asian cities, the role of rapid rail transit was cited differently. This facility was treated as an instrument to induce people to reduce car dependency with improving mass transit system, it is hoped it will provide better choice of transit rather than private transit which generates much more social cost (for example road parking and congested traffic). However, for Bangkok case study, since the alternative rapid rail transit was not ubiquitous and lacks convenience in transport mode transferring, this facility may not be enough to induce mode switch and it will not obviously affect residential pattern change. Nevertheless, what the result is expected to be, not only this alternative can yielded time cost saving advantage as in automobile era, since less monetary cost of rapid rail transit (relative to of automobile) coupled with rising of gasoline cost overtime, the high income resident is more likely to respond to this alternative mode. Therefore, when considering competing rapid rail transit to any other initial transit (bus or automobile), the estimated coefficient for the first variable is expected to be negative and of the second variable is also expected to be positive as in the automobile era.

Table 3.1

Calculation of Differences in Bid-Rent Gradients

Initial Transport Mode	Initial Transport Mode: 1	2[15]	3	4
2	$\beta_2 - \beta_1$			
3	$\beta_3 - \beta_1$	$\beta_3 - \beta_2 = (\beta_3 - \beta_1) - (\beta_2 - \beta_1)$		
4	$\beta_4 - \beta_1$	$\beta_4 - \beta_2 = (\beta_4 - \beta_1) - (\beta_2 - \beta_1)$	$\beta_4 - \beta_3 = (\beta_4 - \beta_1) - (\beta_3 - \beta_1)$	
5	$\beta_5 - \beta_1$	$\beta_5 - \beta_2 = (\beta_5 - \beta_1) - (\beta_2 - \beta_1)$	$\beta_5 - \beta_3 = (\beta_5 - \beta_1) - (\beta_3 - \beta_1)$	$\beta_5 - \beta_4 = (\beta_5 - \beta_1) - (\beta_4 - \beta_1)$
Alternative Transport Mode	1[16]	2	3	4
2	$(\beta_2 + \gamma_2) - \beta_1 = (\beta_2 - \beta_1) + \gamma_2$			
3	$(\beta_3 + \gamma_3) - \beta_1 = (\beta_3 - \beta_1) + \gamma_3$	$(\beta_3 + \gamma_3) - (\beta_2 + \gamma_2) = (\beta_3 - \beta_2) + (\gamma_3 - \gamma_2)$		
4	$(\beta_4 + \gamma_4) - \beta_1 = (\beta_4 - \beta_1) + \gamma_4$	$(\beta_4 + \gamma_4) - (\beta_2 + \gamma_2) = (\beta_4 - \beta_2) + (\gamma_4 - \gamma_2)$	$(\beta_4 + \gamma_4) - (\beta_3 + \gamma_3) = (\beta_4 - \beta_3) + (\gamma_4 - \gamma_3)$	
5	$(\beta_5 + \gamma_5) - \beta_1 = (\beta_5 - \beta_1) + \gamma_5$	$(\beta_5 + \gamma_5) - (\beta_2 + \gamma_2) = (\beta_5 - \beta_2) + (\gamma_5 - \gamma_2)$	$(\beta_5 + \gamma_5) - (\beta_3 + \gamma_3) = (\beta_5 - \beta_3) + (\gamma_5 - \gamma_3)$	$(\beta_5 + \gamma_5) - (\beta_4 + \gamma_4) = (\beta_5 - \beta_4) + (\gamma_5 - \gamma_4)$

[15] Difference in bid-rent gradient, $\beta_i - \beta_j$, if $j \neq 1$ can recover the relevant coefficient to investigate how explanatory variables affect the odds, as report in table 3.1. However it exactly equal to coefficients when estimated by set any group *j* as based category.

[16] Note that γ_1 is equal to zero since the based group never choose the alternative transit over the initial transit, (recall in chapter 3 page 58).

3.3 Data Description

3.3.1 Source of Data

This analysis attempts to understand how the residential location pattern in Bangkok changed, in the period of pre- and post-rapid rail transit, as the effect of transport facility innovation. The data set used is the sampled of households in Bangkok from the Household Socio-Economic Survey (SES) conducted by the National Statistical Office (NSO) in 1998 and 2004. Sample households located in Bangkok and identified their income are used for estimation. It includes 1,445 sampled households in 1998 and 1,512 households in 2004. The record for each sample includes residential location in sub-district or Tombon or Khwang unit, household's transportation expenditure including fixed and variable cost which is defined later and average household's income. However, to avoid the effect of inflation, all income and expenditure of each household should be used in real terms. Furthermore in order to compare the final results of estimation in different period, not only are income level adjusted by consumer price index (CPI)[17] but transport costs are also. With 2004 as the base year, expenditure spent on bus transit must be adjusted by CPI for public transportation service (which is 85.76 in 1998), and expenditure spent on automobile transit must be weighted by CPI for transportation and communication (which is 80.97 in 1998). Each household was then assigned to one of the five income groups ranged from the lowest income group to the highest income group. Income classification is shown in Table 3.2.

Then each income group is grouped into 154 sub-districts (Khwang). However, as sampling survey method of the National Statistical Office (NSO) to collect households' data, it did not distribute surveys into all sub-district (Khwang) areas which made observation left just 96 and 93 sub-districts in 1998 and 2004 respectively. Nevertheless, by using multinomial logit with grouped data, we did not lose the number of individual observations thus it remain much more observations for estimating via maximum likelihood method with asymptotically unbiased.

---

[17] Let 2004 is the based year, CPI for all commodity used for income adjusting in 1998 is 91.778.

Table 3.2
Distribution of Sample among Income
Groups in 1998 and 2004

Group	Monthly Income	1998			2004		
		Average Hourly Income	Number of Samples	Percent	Average Hourly Income	Number of Samples	Percent
1	0 - 5,000	19.49	520	35.99	19.67	431	28.5
2	5,000 - 15,000	48.85	698	48.3	46.29	797	52.71
3	15,000 - 25,000	110.74	149	10.31	101.6	182	12.04
4	25,000 - 35,000	167.44	31	2.15	158.87	45	2.98
5	35,000 up	280.43	47	3.25	258.3	57	3.77
Total			1445	100		1512	100

Source: Socio-Economic Survey in 1998 and 2004, National Statistical office (NSO)

3.3.2 Variable Discussion

In essence, I estimate equation (3.17) by maximizing the log frequency weights of each income group in each residential area on the two distance variables in that equation.

The first of these two distance variables is the distance from their residential area to the CBD, $dist_t$. Euclidean distance measured from residential location to the CBD is used for this variable. The original point of each residential area is created by constructing the center (intersected point of the two diameters of the highest-lowest latitude and highest-lowest longitude square of that area) while the destination point for all residential grids is similarly constructed as the center point of CBD square area which is defined as Khwang Surawong and Khwang Silom surrounded by Surawong Road to the north, Rama IV Road to the east, Sathon Nuae Road to the south, and Mahyesak and Surasak Roads to the west.[18] Then distance from original to destination point is measured in term of kilometers, using Map Magic Program conducted by Thinknet. Co. Ltd 2003-2004.

---

[18] Refer to Kasamsook, A. (2003). Spatial and functional differentiation: A symbiotic and systematic relationship. University College London, UK.

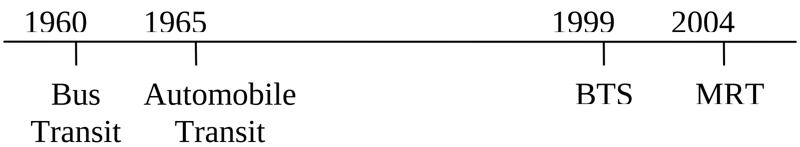
The second variable is the difference of Euclidean distance and break-even distance. Thus the break-even distance in 1998, case 1, case 2, and case 3 in 2004 will be calculated first. It is calculated as equation (2.7) as follows

$$bdist = l^* = \frac{f^A - f^I}{(c^I - c^A) + (wt^I - wt^A)}$$

where $bdist$ is break-even distance at which a resident switches from the initial transit to the alternative transit. Note that upper-script A denotes alternative transit and I denotes initial transit. The initial or alternative mode choice was determined by time table of coming of transit innovation in Bangkok.

Figure 3.3

Time Table of Bangkok Transport Innovation



Source: Kamanamul (2004)

In the period before rapid rail transit was introduced (1998) I consider two main competitive modes. Buses were treated as the initial transit while automobiles were treated as the alternative transit. Therefore each variable related to the break-even distance calculation of these two modes are discussed as follows;

- Fixed cost of bus commuting, f^b, we consider only on unrelated-distance fare, which is the first fixed charge fare that did not depend on distance of commuting. It does not include costs of access to bus stop, due to data limitations.[19]
- Variable cost of bus commuting, c^b, measured by related-distance cost by assuming all expenditure spent on buses and mini-buses relate to distance from residential area to workplace at a constant rate.

[19] Transport expenditure collected by NSO is defined expenditure on each transport mode of all member of a household which is not concerned integrated transport modes used for a trip.

- Fixed cost of automobile commuting, f^a, includes depreciation costs, parking charges, license fees and insurance costs, (while data on vehicle purchasing expenditure was collected within a month before interviewing, thus this data might not be included in fixed cost of car commuting). Depreciation cost is measured by expenditure on vehicle repairing and maintenance, including expenditure on tire and inner tube repairing and/or changing, grease and lubricating oil changing, overhaul and/or repairing, and washing. License fee and insurance costs include driver training, driving license, registration fees, and automobile insurance.
- Variable cost of automobile commuting, c^a, is gasoline expenditure which means regular gasoline, premium gasoline, unleaded premium gasoline, high speed diesel, low speed diesel, and liquefied petroleum gas.
- Time cost, wt^a, can be calculated by using inverse speed of automobile transit multiplied by proportion of mean income placed on commuting time, (0.49 × mean income).[20] In this research, inverse automobile's speed is assumed to be 0.032 hr/km for both sampled years. In addition, bus commuting time cost, wt^b, is calculated by using the value of time multiplied to inverse bus's speed, which is assumed to be 0.1 hr/km for both 1998 and 2004[21].

In the period after rapid rail transit was introduced (2004), break-even distance concerned three pairs of competitive mode choices. Since rapid rail network has not been extended throughout Bangkok yet, we divide residential areas into two parts; the areas with mass rapid lines passing through and without mass rapid lines passing through. Therefore the break-even distance is different for different areas when it involves rapid rail transit. MapMagic (Bangkok 2004-2005 ET) Program conducted by Thinknet, Co.Ltd is used to determine which sub-districts has rapid rail running through.

---

[20] Refer to Dissanayake, D., and Takayuki M. (2003). A Combined RP/SP Nested Logit Model of Vehicle Ownership, Mode Choice and Trip Chaining to Investigate Household Travel Behavior in Developing Countries (the case of Bangkok). The 82th Annual Meeting of TRB.

[21] Average car speed was 31.14 km/hr and average bus speed was 10 km/hr in 2004. Data source from Traffic and Transportation Department, BMA.

1.) bus versus automobile transit

2.) bus versus rapid rail transit

2.1) for grids with rapid rail line

2.2) for grids without rapid rail line

3.) automobile versus rapid rail transit

3.1) for grids with rapid rail line

3.2) for grids without rapid rail line

Therefore in the first case, the initial bus transit versus the alternative automobile, variables used to calculate the break-even distance are totally defined in the same way as in 1998 while, for the left, the break-even distance are set differently

Variable and fixed costs of bus and automobile transit in case 2 and case 3 in 2004 are still defined as in 1998. Instead of using mean variable and fixed costs of all sampled households, these commuting costs must be separated by areas in which a resident's is house located and whether it ahs or doesn't has rapid rail transit lines passing through firstly, then keep tract on calculating as in equation (2.7). However, variable and fixed cost of rapid rail transit for residents who are locate in areas with or without rapid rail lines must be added.

- For areas with rapid rail lines, fixed cost, f^s, is defined as first charge fare that does not depend on commuting distance which its average fixed cost equal to 12 Baht[22]

Calculating break-even distance in the residential areas without rapid transit lines, I attempted to take account of the fact that traveling by alternative transit (rapid rail transit) yielded more fixed commuting cost for residents living in areas without rapid rail lines pass through (bus for case 2.2, and automobile for case 3.2), since rapid rail transit must be served by the other feeding modes. Many residential areas were not served directly by rapid rail lines, thus break-even distances varied considerably in this sample, giving observations on how the availability of the rapid rail transit affected bid-rent, holding distance from the center constant.

[22] First charge fare was 10 Baht for BTS, and 14 Baht for MRT. Data sourced from the Bangkok Mass Transit System Public Company Limited and the Rapid rail transit Authority of Thailand.

- For areas without rapid rail lines, fixed cost must include first charge fare that does not depend on distance and all access cost, in this study means all expenditure spent on the other transport modes rather than rapid rail transit. It includes expenditure on taxis, Samlors, Tricycles, hired-motorcycles, boats, ferries, trains, and other local transport.
- Variable cost of rapid rail transit, c^s , for areas with and without rapid rail lines, is defined as expenditure spent on rapid rail transit after subtracting by its first fixed charge. It depends on in-line distance from the original station to the CBD station.
- Time cost, wt^s , is calculated by using inverse speed of rapid rail transit multiplied by proportion of mean income placed on commuting time, (0.49 × mean income).[23] Inverse rapid rail transit speed is assumed to be 0.029 hr/km

Note that fixed and variable costs of any transport mode must be measured in terms of cost per trip assuming that a household generally generates two trips per day (for home-to-workplace and workplace-to-home). Since commuting expenditure of each household was collected in term of expenditure per month, these costs must be divided by average worked-day per month first (which was 21.08 in 1998 and 21.99 in 2004). Thus fixed cost for each transit is measured in terms of Baht per trip and variable costs for each transit (both monetary and time costs) are measured in terms of Baht per kilometer. Furthermore, variable and fixed cost of each transit choices are extracted from transport expenditure data in Socio-economic Survey (SES) 1998 and 2004 at household level. However, it must be proper using that transport cost in terms of average per head level rather household level, thus these costs must be divided by each average transit user then computed in terms of average cost per head before calculating the break-even distance.

---

[23] Refer to Dissanayake, D., and Takayuki M. (2003). A Combined RP/SP Nested Logit Model of Vehicle Ownership, Mode Choice and Trip Chaining to Investigate Household Travel Behavior in Developing Countries (the case of Bangkok). The 82th Annual Meeting of TRB.

Since the break-even distance is defined as the distance that lower commuting time of the alternative mode can be offset with its higher variable, some commuters will switch from using initial transport mode to the alternative mode. Therefore switching transport mode will exist only if transport costs occurring when travel by initial mode, at least, equal to when travel by the alternative mode. Then we obtain the break-even distance as equation (2.7)

$$l^* = \frac{f^A - f^I}{(c^I - c^A) + (wt^I - wt^A)}$$

If fixed cost of the initial transit is less than that of the alternative transit, $f^I < f^A$, thus to satisfy positive break-even distance the first condition is $(c^I - c^A) + (wt^I - wt^A) \neq 0$, and the second is $(c^I - c^A) + (wt^I - wt^A) > 0$. Note that commuting by the initial transit is always slower than the alternative transit, $t^b > t^r$. It means that difference between time cost when commuting by initial mode and alternative mode must be positive (assuming that alternative has higher speed and takes less commuting time) and enough to offset negative terms of difference between variable cost of initial and alternative mode.

However, time saving varies accordingly with income proportion placed on commuting time. Thus it may generate negative break-even distance for the lower income group. The divisor can be negative and the numerator can be positive. It means that time saving of the alternative mode cannot be offset by its higher variable cost at any distances. We denote this break-even distance by infinity (∞). It can imply that commuter always uses the original mode for every distance. In practice, when break-even distance is infinity, difference between Euclidean distance and break-even distance is always negative which makes independent variable "$(difdist)x$" disappear as its dummy *x* is set to be zero.

Table B.1 presents parameter value used for calculating break-even distance when each income group chose between the initial bus transit and the alternative automobile transit in 1998. Correspondingly in 2004, Table B.2-B.4 presents parameter value used for calculating break-even distance for bus versus automobile transit (case 1 in table B.2), bus versus rapid rail transit for grids with and without rapid rail line (case 2 in table B.3.1 and B.3.2), automobile versus rapid rail transit for grids with and without rapid rail line (case 3 in table B.4.1 and B.4.2) respectively. Therefore the break-even distance is 1.658, 0.397, 0.234, and 0.129 for the 2nd, 3rd, 4th, and 5th respectively in 1998. In 2004, the break-even distance is calculated if resident switch from bus transit to automobile. As shown in Table B.2, the break-even distance is infinity for the 2nd and 3.981, 0.970, and 0.419 for the 3rd, 4th, and 5th income group. Considering any initial modes versus rapid rail transit, break-even distance for each income group is calculated separately for areas with and without rapid rail lines. Hence, the break-even distance of bus versus rapid rail transit is 8.499, 3.012, 1.805, and 1.065 kilometer for the 2nd, 3rd, 4th, and 5th respectively for area with rapid rail line, and infinity for the 2nd income group and 26.887, 7.266, and 3.205 for the 3rd, 4th, and 5th income group in areas without rapid rail lines. For automobile versus rapid rail transit, the break-even distances for the 2nd income to the 5th income group are 1.407, 1.379, 1.351, and 1.305 within areas with rapid rail lines. In areas without rapid rail lines assuming that commuting by rapid rail is circuitous and yield more fixed cost to reach the nearest station, therefore the break-even distances are infinity for all income groups.

Following the concept of the model how the residential pattern among different income groups change as advantage of the alternative transit, the advantage of availability of alternative transit is taken into the model only if the break-even distance is shorter than the Euclidean distance. It means that residents who are located in the area characterized by Euclidean distance have a chance to switch to the alternative transit rather using the initial transit as its net gain from time cost saving. Therefore presence of the second independent variable depends on its dummy variable x which equals unity if difference is positive and zero otherwise. Denoted this variable by "$(difdist)x$".

We already discussed about explanatory variables, $dist$ and, $(difdist)x$, now consider dependent variables used in the model. The basic estimated equation for this study (recalls equation 3.14) is shown as follows;

$$\log\left(p_{it}/p_{jt}\right) = (\alpha_i - \alpha_j) + (\beta_i - \beta_j)dist_t + \gamma_i\left(difdist\right)x_{it}$$

From equation (3.14), $\alpha_i - \alpha_j$, $\beta_i - \beta_j$, and γ are the estimated variables, while dependent variable involves the probability that a member of group i occupies location t which is actually unobserved. Following multinomial logit approach, actual choice denoted by 1, 2, 3, 4, and 5 which represents the outcome that location t is occupied by the 2^{nd}, 3^{rd}, 4^{th}, and the 5^{th} income group respectively, must be entered to the model as dependent choices.