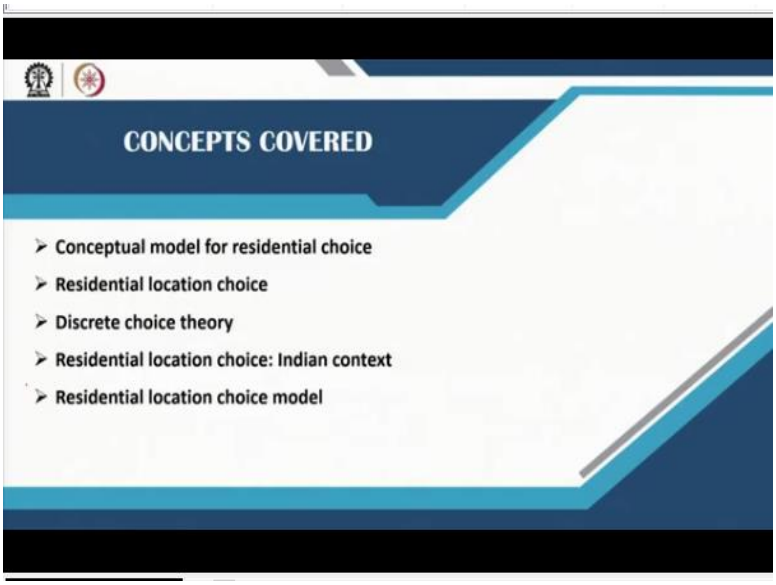
Urban Landuse and Transportation Planning
Prof. Debapratim Pandit
Department of Architecture and Regional Planning
Indian Institute of Technology-Kharagpur

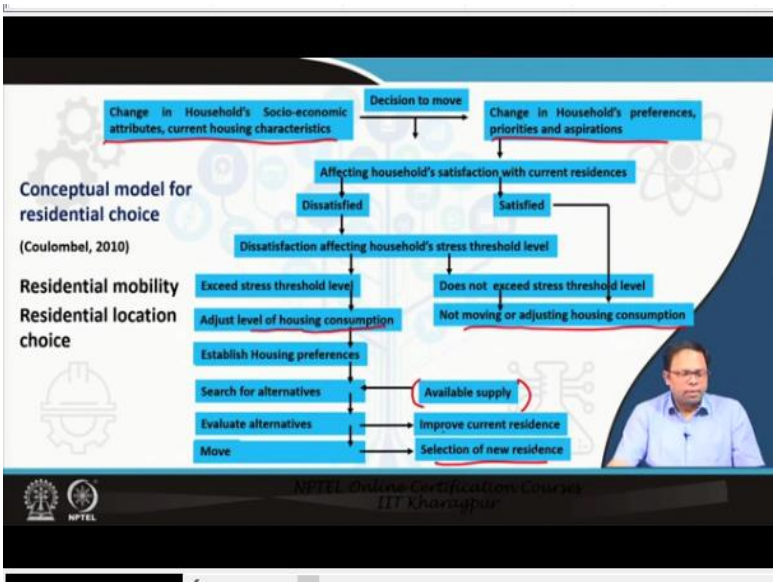
Lecture-29
Residential Mobility and Location Choice 2

(Refer Slide Time: 00:28)



The different concepts covered in this lecture are the conceptual model for residential location choice or residential choice, discrete choice theory, residential location choice in the Indian context, and finally the residential location choice model.

(Refer Slide Time: 00:48)



Conceptual model for residential location choice

Residential location choice is a fairly complex process, involving several decisions, like where to move, when to move, tenure choice, and so on. All these decisions are based on the household's constraints and opportunities, which adds another layer of complexity. To deal with this, many researchers have developed conceptual models and explained the process by breaking it into sequential decisions.

The basic conceptual model for residential location choice can be explained as a model, which includes both the process of intention to move and actual mobility. In other words, it explains when and where a person starts considering his intention to move and then to the final point, when he actually moves to a particular location. The given conceptual framework was developed by Wong.

Considering this model, the household first considers the decision to move when there is a change in the household's socio-economic attributes, current housing characteristics, or change in the household's preference, priorities, and aspirations. All these factors influence the household's level of satisfaction with the current dwelling. Based on this level of satisfaction only, the household decides to stay or move from the current location.

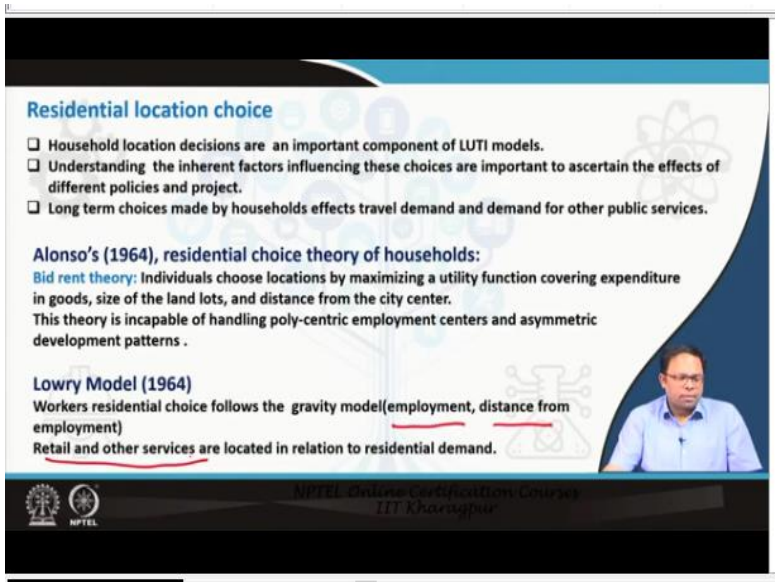
For example, there is a change in household needs due to an increase in household size, but the present housing location can take care of this need, so the household will not consider a move (or the household is satisfied). On the other hand, if the needs are not catered by the present dwelling, the household will get dissatisfied and may develop a desire to move. It is important to understand that dissatisfaction with the current dwelling does not always result in a decision to move.

Once the household gets dissatisfied with the current dwelling, the decision to move or search for a new dwelling depends on the stress threshold. In other words, if the dissatisfaction is beyond the threshold level, the household will starts looking for new dwellings or take another course of action like constructing an additional floor, purchasing a car. Whereas, if it is not beyond the threshold level, the household will not move. For example, there is a change in the workplace location of the household head which is quite far from the current dwelling. So, the household has to travel more. If the level of dissatisfaction is beyond the threshold level, he will move to a location near to his workplace, or he may purchase a car.

When the household finally decides to move, he establishes his housing preferences that means what kind of housing he would prefer. Once the preferences are established, he starts searching for the alternatives, and form a choice set from the pool of available housing i.e. housing supply. Then, he evaluates the alternatives and selects one dwelling to move.

Suppose, if the preferred housing is not available, he may terminate the search. Also, during the evaluation stage, he may abandon the thought of moving and improves the current dwelling. So, the process of residential location choice is very complex. It involves several decisions, and all these decisions need to be considered when modelling the residential location choice behaviour.

(Refer Slide Time: 05:26)



Residential location choice

When the household decides to move, the next decision is where to move. This residential location choice decision is a very important component of the land use and transportation model. The inherent factors influencing these choices are important to understand to ascertain the effects of different policies and projects. For example, how a new transport facility is going to affect the location choice decision of a household. Moreover, these choices also affect the travel demand and demand for other public services.

In 1964, Alonso developed a residential location choice model based on the concept of bid rent theory. He considered a monocentric city with employment opportunities. He argued

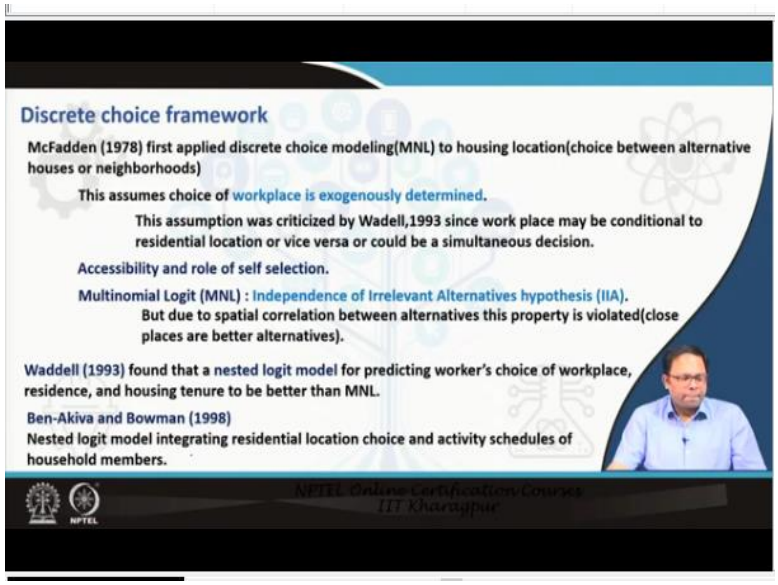
that individuals choose locations by maximizing a utility function, covering expenditure in goods, and the size of the land lots, and the distance from the city centre i.e., people will choose a location based on the distance from the city center and also the amount of money he can spend on the size of the land lot. The major criticism of Alonso's work was the assumption of a monocentric city. So, his residential location theory was incapable to explain the polycentric employment concepts and asymmetric developing pattern.

Parallel to the work of Alonso, Lowry developed a residential location choice model. This model considered residential location choice for the first time within the context of land use and transportation model. The land use component was in the form of the location choice of employees of industries. The worker's location choice followed a gravity model based on the distance of employment, nature of employment, and the type of employment.

The gravity model has two components. The first component is potential and the second component is impedence. In Lowry's model, potential is the total quantum of employment and distance is the impedance. So, based on these two factors, people will distribute themselves around the employment center.

Once the people distribute themselves, then retail and other kinds of service employments are created because of these new residences. Because wherever there are new residences, there is a demand for retail and other services. This demand will lead to the creation of new jobs, which eventually create a requirement for new residences. So, this iterative process was considered in Lowry's model.

(Refer Slide Time: 08:53)



Discrete choice framework

In 1978, McFadden first applied the discrete choice framework to model the choices between alternative houses or alternative neighborhoods. This model was based on the utility maximization principle.

In residential location choice modelling, the alternatives can be considered at the building level or neighbourhood level, generating a huge number of choices. In reality, people do not consider all the alternatives. They only consider a subset of these alternatives. For example, at a particular point in time, there is only a certain number of apartments available in the market, so a household who is looking for an apartment will consider these apartments/apartment complexes only. So, the discrete choice model is between choices of alternatives of different apartments/apartment complexes.

McFadden assumed the workplace location to be exogenous to the residential location choice model. It means that workplace choices were already known and used as exogenous input in the model i.e., choice of workplace is not determined and the existing location of the workplace is used to determine the residential location choice. This assumption was criticized by Waddell. He stated that workplace location may be conditional on residential location or vice versa, or could be a simultaneous decision. In other words, workplace location choice decisions are endogenous to this kind of modelling process.

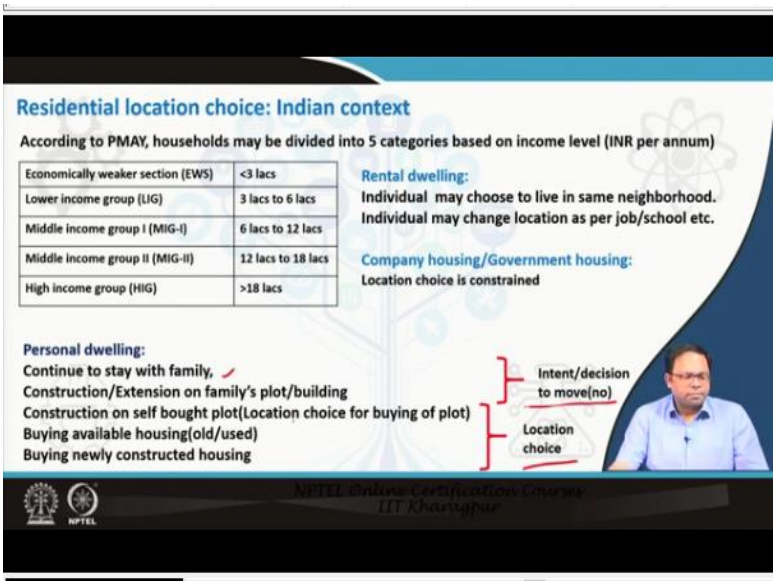
The residential location choice was analysed using a multinomial logit model (MNL). The advantage of using a multinomial logit model is that the probability of choosing two

alternatives does not depend on the presence of other alternatives. This is the independence of irrelevant alternatives (IIA) property of the MNL model which is beneficial in some sense in the overall modelling process of residential location. However, in this particular case, this property is violated because of the spatial correlation between alternatives. For example, location 'B', and 'C' are near to location 'A', and location 'D' is quite far from location 'A'. So, a person living in location 'A' wants to move to a nearby location 'B'. If he does not get a dwelling in this location, he may choose location 'C'. It means that these choices are influenced by each other. In other words, if one alternative is chosen, the other alternative will be affected. Therefore, the nearby alternatives will have a different set of probabilities.

The IIA property will be discussed in detail using the blue bus and red bus paradox in the mode choice modelling section.

In 1993, Waddell found that a nested logit model for predicting worker's choice of workplace, residence and housing tenure is better than a multinomial logit model. Ben-Akiva and Bowman also used the nested logit model by integrating residential location choice and activity schedules of household members.

(Refer Slide Time: 13:55)



Residential location choice: Indian context

The residential location choice behavior in the context of a developing country, like India, may differ from that in western countries due to differences in social, economic, and cultural aspects. So, one needs to have a certain understanding of the local context, to understand the residential location choice behaviour.

According to Pradhan Mantri Awas Yojana (PMAY), households may be divided into 5 categories based on the households' income level, which is in INR per annum. The income level of an EWS household is less than 3 lacs, lower-income groups is 3 to 6 lacs, middle-income group-I is 6 to 12 lacs, middle-income group-II is 12 lacs to 18 lacs, and higher-income group is greater than 18 lacs.

Each of these income groups can be considered as homogenous categories and different models can be developed for each of these categories. This categorization is important because the housing supply, housing preference, and the decision making process varies for each household category. If all of these household categories are modeled together, then probably the model will give insignificant results because the constituent variables varies among these categories. Alternatively, if the basic variables are the same for all the groups, then all categories can be modelled together. But in many cases, it is not the case.

There are different kinds of dwellings like rental, company housing, government housing, or personal dwelling. The motivation or reason for an individual to stay in each dwelling category also varies. For example, people who are staying in rental dwellings may choose to live in the same neighbourhood or may change location as per job, schools, etc. In other words, when the rental tenure gets over, people will try to choose a location that is nearby, and when they are changing their job or school for their kids, there is a chance that people may also change the rental location. In the case of company housing and government housing, location choice play a little role since the dwellings may be assigned to households or choice options may be limited. So, location choice is constrained.

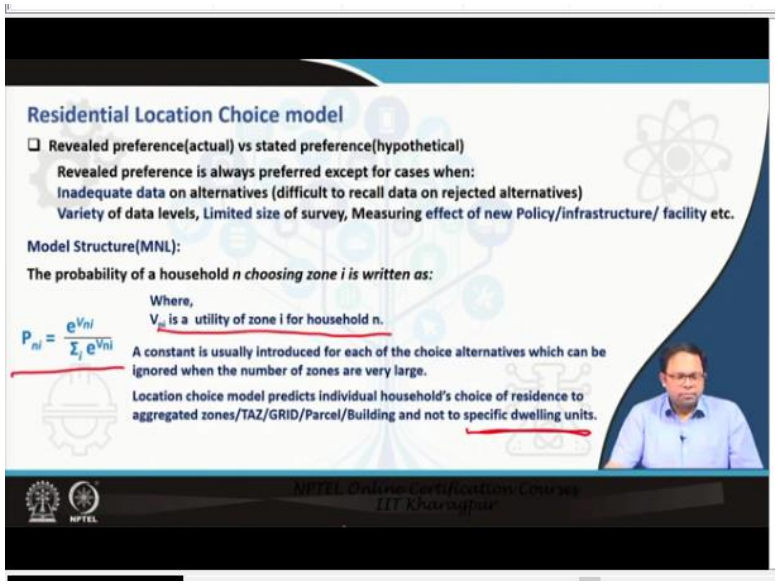
In the case of personal dwelling, the outcome of residential location choice behaviour differs based on household constraints, preferences, or stress threshold. Considering the intention to move, a family can either continue to stay, or extend one floor in the dwelling, or construct on the family's existing plot, or move. Generally, moving behaviour is less among dwelling owners due to higher transaction costs as compared to renters.

Likewise, the choice of dwelling can also vary. Generally, people buy plots and eventually construct a dwelling on it. In such cases, the location choice of plot is more important because the person's choice to build a house is constrained by the decision that he has already

bought a plot. So, he will build a house on that plot itself. The other choice can be buying available dwelling (old/used), because these dwellings are available at better locations and probably at lower prices. Thus many people do consider buying available housing. The final choice is a newly constructed dwelling, i.e., the choice can be only fresh housing stock provided by real estate developers during the planning period.

Thus, separate models should be developed for available housing, newly built housing, and other choices. In the case of newly built housing, the choice set constitutes only fresh housing stock whereas, in the case of available housing choice, the choice set includes existing empty housing stock that is available during the planning period. So, all the aspects discussed in this section should be considered when developing a residential location choice model in the Indian context.

(Refer Slide Time: 19:08)



Residential location choice model

Similar to the intent to move model, the residential location choice model can be developed using revealed preference data or stated preference data. The revealed preference data is always preferred because it represents the actual choice made by the household. In some cases, stated preference is more suitable. For example, when there is very inadequate data available on alternatives like people having difficulty in recalling what the rejected alternatives were. The other example is the case when the effect of the new policy, infrastructure or a facility is to be measured, since, these facilities or projects were not

available at that point of time when people made their choice for a location. So, both stated preference and revealed preference model can be developed.

The model structure for the residential location choice model can be written as;

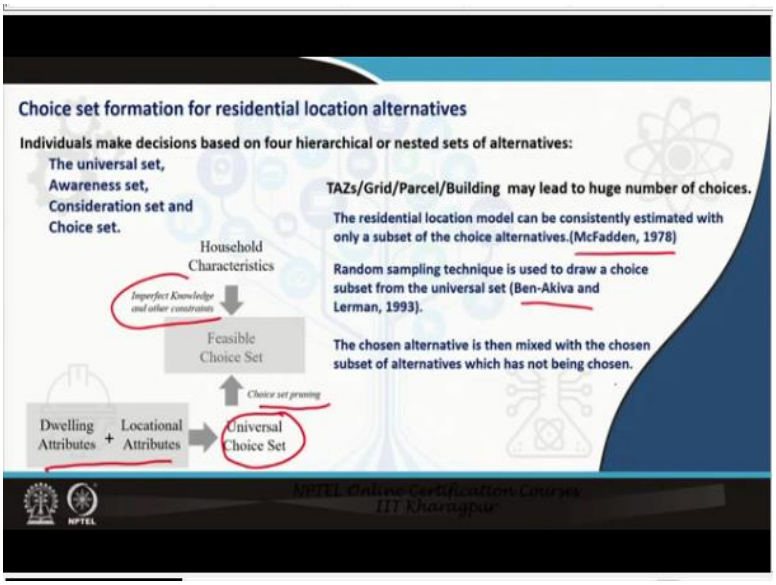
$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}$$

Where, Pni is the probability of a household 'n' choosing zone 'i'. Vni is the utility of zone 'i' for household 'n'. This equation is basically the utility of the selected alternative compared to the utility of all alternatives.

In the utility equation of an alternative, a constant β0 is usually introduced for each of the choice alternatives, this parameter can be ignored when the number of zones is very large. So, in the multinomial choice model, when the number of alternatives is huge, for example, a person may have chosen 100s of alternatives to choose from, in such cases, the constant parameters can be dropped.

The location choice model predicts a household's choice of residence to aggregated zones, TAZ, GRID, parcels, or building and not to specific dwelling units. Thus, all the location based or dwelling based characteristics introduced in the model, are for a particular zone, or parcel, or building, but it is generally not for a particular dwelling unit. If the level of analysis is dwelling level, then the number of alternatives will be huge, and it would be difficult to develop such a model.

(Refer Slide Time: 22:29)



Choice set formation for residential location alternatives

Individuals make decisions based on four hierarchical or nested sets of alternatives, which are the universal set, awareness set, consideration set, and choice set. Universal set constitutes all the available alternatives. For example, all the available dwelling units in an urban area. Awareness set includes the alternatives of which decision maker is aware of, because an individual does not possess perfect information about all the alternatives. He is only aware of a few alternatives. Consideration set includes those alternatives that clear decision maker's criteria, for example, budget constraints. Finally, the person selects 3, 4 alternatives, and chooses one alternative among these alternatives. This set is the final choice set.

When a household chooses a dwelling, he chooses the location as well. So, each alternative in a universal set is characterized by location attributes and dwelling attributes. Based on the household's characteristics, imperfect knowledge, and constraints, a feasible choice set is developed. This feasible choice set will be evaluated, and the decision is taken.

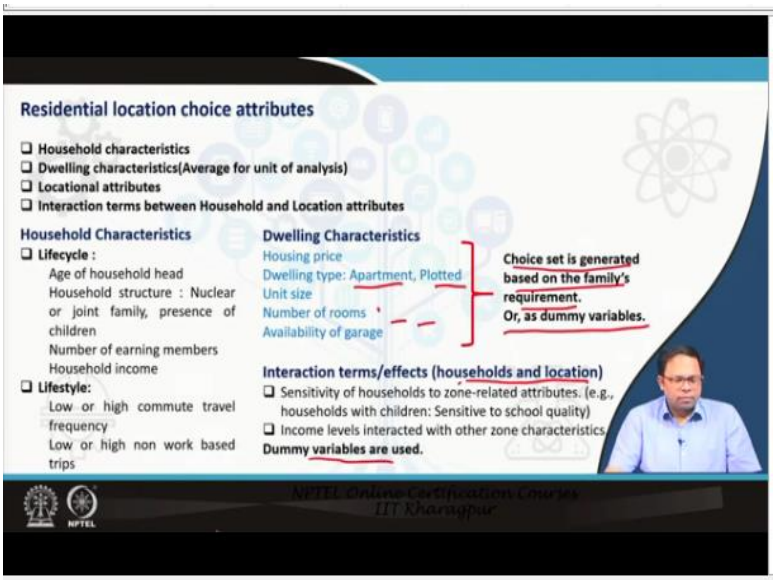
In residential location choice modelling, the level of aggregation can be at the TAZ level, parcel level, or building level. The level of aggregation has a direct bearing on the number of alternatives. Even, TAZ, GRID, parcel, or building level will lead to a huge number of choices. It becomes very difficult to estimate such a model. Also, the model specification would be very large, which means if 4 alternatives are considered then 4 models are to be built, if there are 100 alternatives, then 100 models are to be built.

So, the residential location model can be consistently estimated with only a subset of the choice alternatives. This can be easily translated to a multinomial logit model because of the IIA property. IIA property states that the probability of selection of one alternative compared to another is not affected by the third alternative. Therefore, the choice probability of two alternatives, that are being considered, will not get affected by the removal of some other alternative. This was given by McFadden in 1978.

In 1993, Ben-Akiva and Lerman proposed a random sampling technique to draw a choice subset from the universal set. The chosen alternative is then mixed with this subset of alternatives. Then, the model is estimated and the beta coefficients of each attribute for all the locations are determined. Also, the randomness in the selection of a subset of not chosen

alternatives does not change the estimation of the model due to the IIA property. This strategy is adopted in the seminal models that are being developed.

(Refer Slide Time: 26:57)



Residential location choice attributes

The different residential choice attributes considered in residential location choice models are household characteristics, dwelling characteristics, location attributes, and interaction terms between household and location attributes.

Household characteristics can be broadly categorized as life cycle characteristics and lifestyle characteristics. Life cycle characteristics are age of the household head, household structure (nuclear or joint family), the presence of children or not, number of earning members, household income. Lifestyle characteristics include different lifestyle choices of a household like low or high commute travel frequency or low or high non work-based trips.

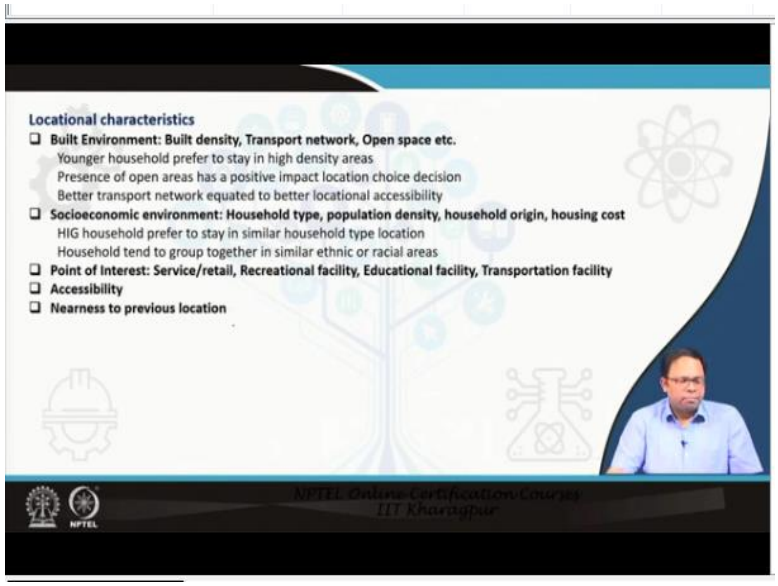
Dwelling characteristics include attributes related to the dwelling unit. For example, housing price, dwelling type (apartment or plotted), unit size, number of rooms, availability of garage, and others. It is important to note that each dwelling unit has different properties. So, a common property/attribute cannot be used for a particular zone or a housing complex. Therefore, dwelling characteristics are also included in the choice set. This choice set is generated based on the family's requirements. Also, the variable like housing type, number of rooms can be introduced as dummy variables. The number of dummy variables representing a single attribute is equal to the number of levels in the attribute minus 1. For example, the

dwelling type attribute has two levels i.e. apartment and plotted. So, one dummy variable will be introduced. In addition to this, the dummy variable help in understanding the influence of each level of attribute on the overall probability of choice. For example, how the probability will change if it is an apartment, or if it is a plotted housing.

The interaction terms between household and location include the variable which represents combined characteristics of household and location. For example, a household with children sensitive to school quality, here two variables are combined. The first variable is household with children which can take values of 1 (presence of child) and 0 (absence of child). The second variable is sensitivity to school quality. So, an interaction variable can be created by multiplying the first variable and the second variable. Also, automatically the value of the interaction term will become 0 for a household without children.

Similarly, other interaction effects like income levels interacted with other zone characteristics can be used. Because households belonging to a particular income group prefer to stay in certain zones.

(Refer Slide Time: 30:21)



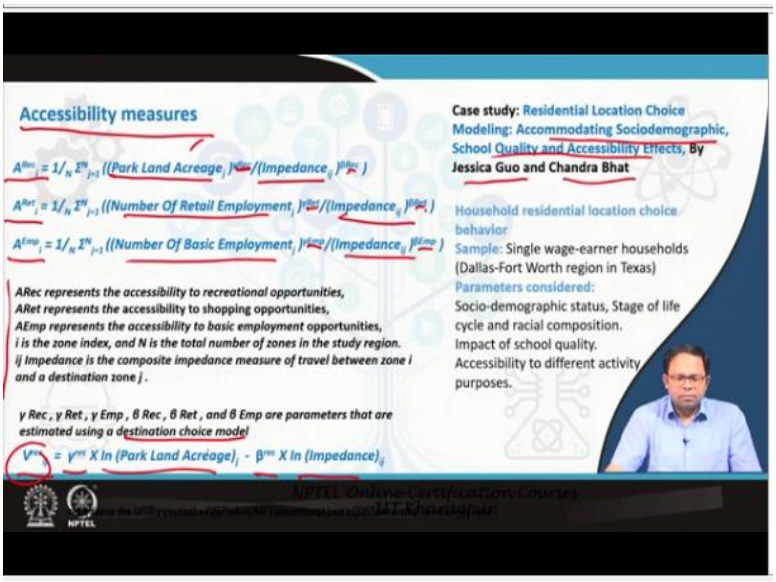
The other variables that are included in residential location choice are locational characteristics. These variables are the most important variables that influence the choice of a location. It includes built environment characteristics, socioeconomic environment, accessibility, point of interest, and nearness to the previous place.

Built environment characteristics such as built density, transport network, open spaces, can influence the location of a household positively or negatively. For example, younger households prefer to stay in high density areas, due to availability of many activities. The presence of open areas, better transport network has a positive impact on the location choices. Socioeconomic environment characteristics include household type, population density, household origin, and housing costs. For example, HIG household prefer to stay in similar household type locations. So, the predominant household type of a neighbourhood plays a role. Households tend to group in similar ethnic or residential areas. So, these kinds of characteristics are important variables in location choice decisions.

Point of interest like presence of service or retail, recreational facility, educational facility, or transportation facility, influences a location choice. For example, a household that travels by metro will choose a location where the metro station is available.

Nearness to the previous location also influences the location choice of a households, because people do not want to venture too much away from their existing location when they choose a new location.

(Refer Slide Time: 32:02)



Accessibility is defined as the ease to access opportunities available in an urban area using available transport facilities. It plays an important role in location choice decisions because people prefer to choose more accessible locations. Many researchers have developed different

measures of accessibility and incorporated these measures in the residential location choice model.

To understand the residential location choice behaviour further, the model developed by Guo and Bhat (2007) is discussed. They investigated the residential location choice behaviour of a single wage earner household in the Dallas Fort Worth region in Texas, America. The study included a) sociodemographic characteristics such as income of household, lifecycle stage, racial composition, b) school quality, and c) accessibility to different activity purposes.

In the present study, the author developed measures of accessibility to recreational facilities, shopping opportunities, and basic employment opportunities based on Hansen type measure. The equations are as follows:

$$A_i^{Rec} = \frac{1}{N} * \frac{1}{\sum_{j=1}^{N} \left(\frac{(Park\ land\ acreage)^{\gamma^{Rec}}}{(Impedance)^{\beta^{Rec}}} \right)}$$

$$A_i^{Ret} = \frac{1}{N} * \frac{1}{\sum_{j=1}^{N} \left(\frac{(Number\ of\ retal\ employment)^{\gamma^{Ret}}}{(Impedance)^{\beta^{Ret}}} \right)}$$

$$A_i^{Emp} = \frac{1}{N} * \frac{1}{\sum_{j=1}^{N} \left(\frac{(Number\ of\ basic\ employment)^{\gamma^{Emp}}}{(Impedance)^{\beta^{Emp}}} \right)}$$

Where A^{Rec} represents the accessibility to recreational opportunities, A^{Ret} represents the accessibility to shopping opportunities, A^{Emp} represents the accessibility to basic employment opportunities, 'i' is the zone index, N is the total number of zones in the study region, and $Impedance_{ij}$ is the composite impedance measure of travel between zone 'i' and a destination zone 'j'. In these measures, both the quantity and distance of the facilities are taken into consideration and summed up for all the surrounding zones to calculate the accessibility of the current zone.

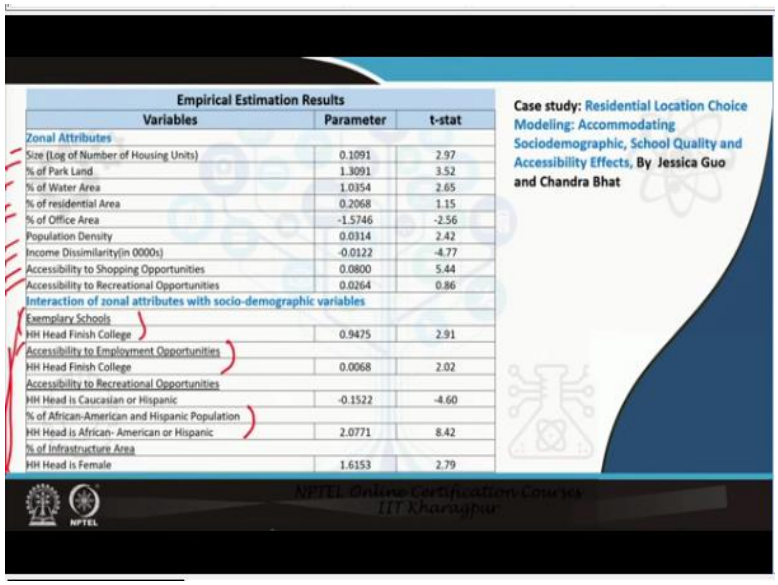
The variable γ^{Rec}, β^{Rec}, γ^{Ret}, β^{Ret}, γ^{Emp}, and β^{Emp} are calibrating variables. These parameters are estimated by developing another model, which is a destination choice model. The destination choice model basically determines the parameters that influence a person's choice of

destination. The variables γ^{Rec}, β^{Rec}, γ^{Ret}, β^{Ret}, γ^{Emp}, and β^{Emp} are estimated using the destination choice model of the form given below:

$$V_{ij}^{Rec} = \gamma^{Rec} * \ln \quad (Park\ land\ acreage)_j - \beta^{Rec} * \ln \quad (Impedance)_{ij}$$

Where V_{ij} is the utility of zone 'j' for recreational opportunity to an individual in zone 'i'. Similarly, the same formulation can be used for employment and shopping opportunities.

(Refer Slide Time: 36:06)



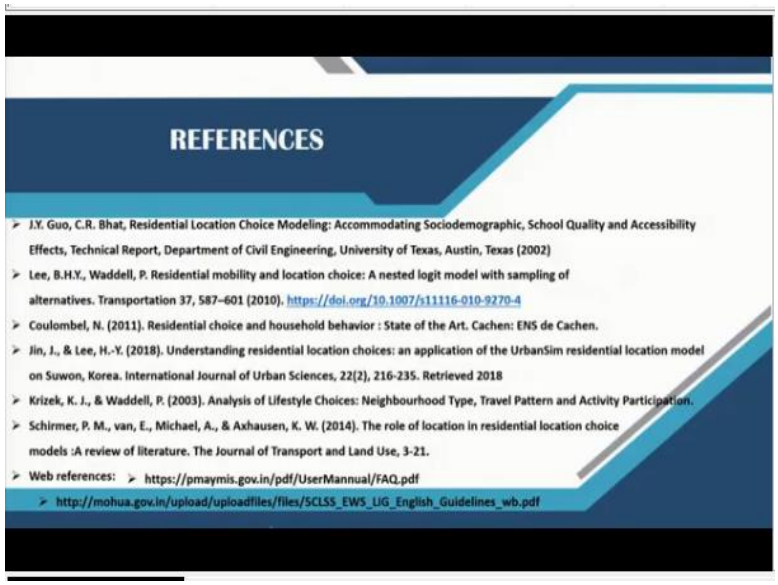
The variables included in the model are zonal attributes and interaction terms between zonal attributes and socio-demographic characteristics. Zonal attributes include percentage of park land, water area, residential area, office area, population density, accessibility to shopping and recreational opportunities, and so on. The interaction terms are household racial origin and accessibility to recreational opportunities because households with similar racial profile tend to stay together. Another interaction term is the percentage of infrastructural area and household head is female. The interaction terms are included to accommodate the sensitivity of household for zonal attributes.

The empirical estimation results of the multinomial logit model are given in the table. The given zonal attributes and interaction of zonal attributes with socio-demographic variables are found to be significant. These variables are either positively related or negatively related. For example, zonal attributes like percentage of office area, and income dissimilarity are negatively related to the choice of a particular zone. Whereas, zonal attributes like

accessibility to shopping, accessibility to recreation, population density, percentage of residential area, etc. are positively related to the choice of a particular location.

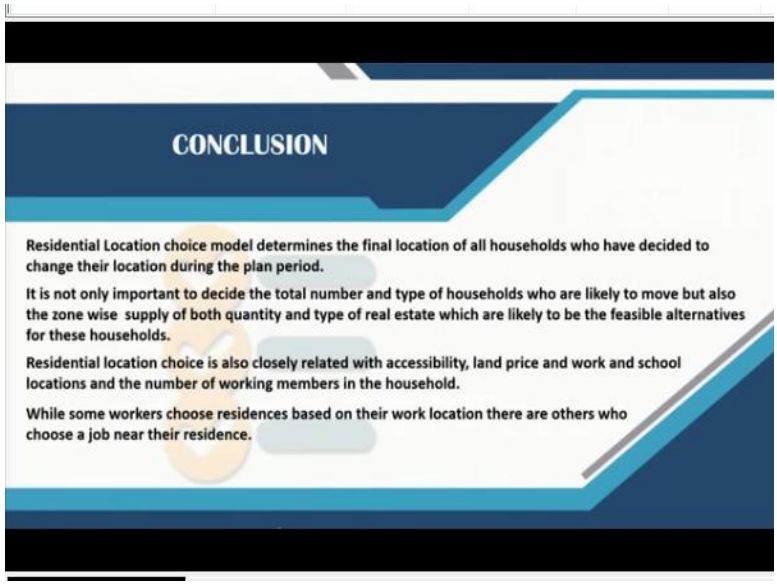
So, in this way, different locational characteristics, household characteristics, and interaction terms can be included in the model.

(Refer Slide Time: 37:27)



Some of the references are listed in the above slide.

(Refer Slide Time: 37:34)



Conclusion

The residential location choice model determines the final location of all households who have decided to change the location during the plan period.

It is not only important to decide the total number and type of households who are likely to move, but also the zone wise supply of both quantity and type of real estate, which are likely to be the feasible alternatives for this particular household.

Residential location choice is also closely related to accessibility, land price, work and school locations, and the number of working members in the household. So, these are the factors that influence residential location choice. While some workers choose residences based on their work location, there are others who choose a job near the residences particularly blue collar workers, plumbers, household maids.