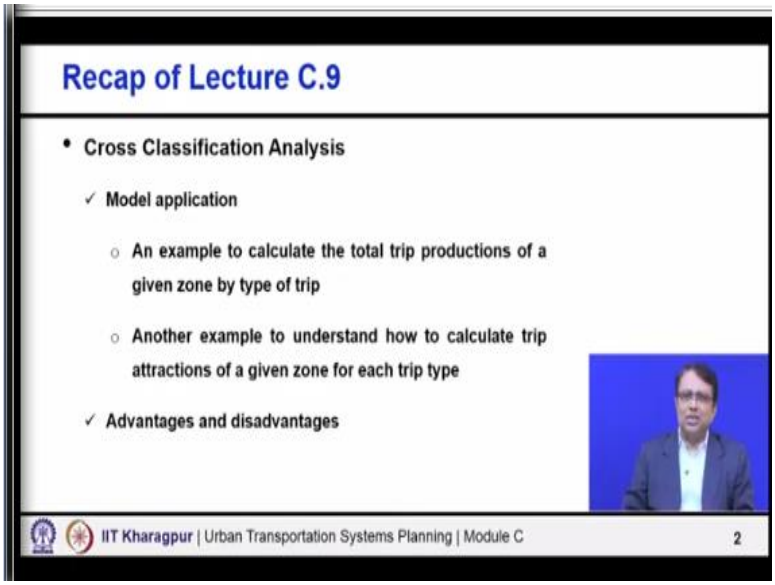
Urban Transportation Systems Planning
Prof. Bhargab Maitra
Department of Civil Engineering
Indian Institute of Technology-Kharagpur

Lecture-20
Matching Productions and Attractions;
Stability of Trip Generation Models

Welcome to lecture 10 of module c, this is the last lecture for this week 4 and also the last lecture for this module.

(Refer Slide Time: 00:25)

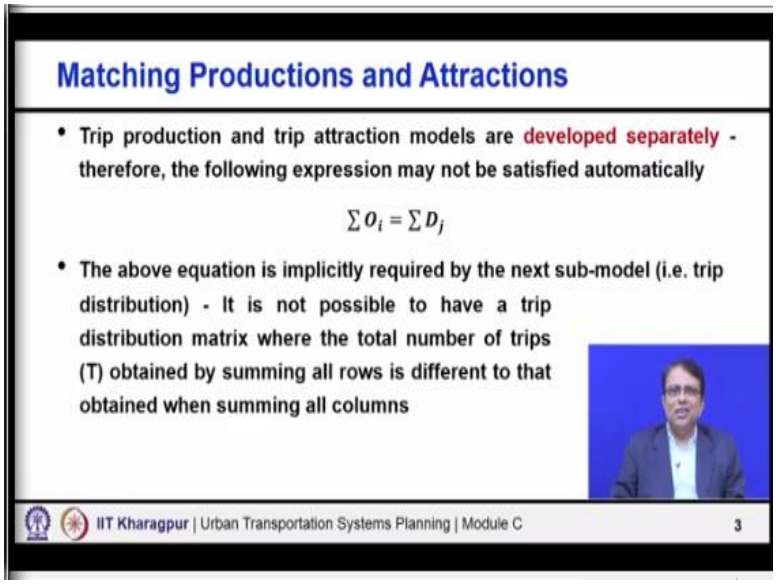


So, far we discussed about 2 approaches for carrying out trip generation analysis, the regression based model, the cross classification or category analysis, both approaches we discussed in details, various implications, various aspects which you need to remember, we try to discuss and highlight those. Then particularly in the last lecture, we are focused on the model application part related to the cross classification analysis.

Once you say the model structure, then you calibrate the model and the new actually apply those models. So, we discuss specifically about the model application parts related to cross classification analysis. And then finally we also discussed about the advantages and disadvantages associated with the cross classification analysis. So, that helps us to compare the cross classification analysis and the regression based model for trip generations.

So, there are both advantages and there are disadvantages and based on those one has to take a call that for a given condition or a given context, what approach should be followed? Should we apply a regression based model or should we apply cross classification or category analysis?

(Refer Slide Time: 02:16)



Now with all those today in the last lecture of this module, I want to discuss 2 specific aspects of trip generation modeling, applicable to both approaches, one is matching the productions and attractions, the second one is stability of regression models. Let us talk about the first thing matching productions and attractions. What we are trying to say here? That whether we are using regression based model or whether we are using category based or trip category or trip rate analysis across classification or trip rate analysis.

Both cases the productions and attractions model are different, if you are using regression based model separate model for productions and separate model for attractions, the related variables are also very different. Similarly when you are using category analysis, you have seen that examples that where one example we gave for trip production estimation and then another case trip attraction estimation.

They are different trip rates, different considerations, one case household characteristics another case the attraction into land use and further classification, characterization. So, ultimately both

approaches the models for productions and attractions are different and both are estimates, what you get out of application of those models are the estimates. So, you can consider this is something which is related to application or you know little bit beyond application also you can say.

So, last bit of application in essence. So, what we are saying, that since they are different models, how do you ensure that the total trips produced from all zones in a study area is equal to the total trips attracted in all zones in a study area or in a study. They have to match finally, why they have to match? The reason is very simple. If I say let us say I have n number of zones, so I am basically then talking about a n by n matrix, I have got in rows 1, 2, 3, 4 n.

So, trips are produced from each of these zones, and I have got n columns trips are attracted to each of these zones. So, the rows are indicating I have got 10 rows or n rows maybe in zones how the trips are getting produced and getting attracted to one of the j zones. J is also 1 to n, i's production is also i 1 to n. So, what we are getting, if I consider the T i j each cell of that matrix, then what is you take a row at all the columns, what do you are getting, you are getting productions from that zone.

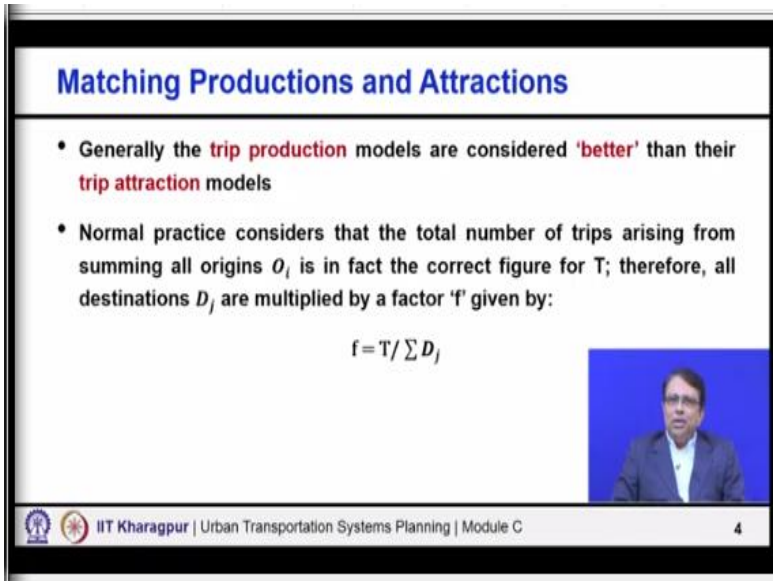
So, T i j sum over all j, that gives you the production from zone i. In the same matrix, if I take each column and sum all the rows in that column, what we have doing, T i j submit over all i, so what I get, I get the attractions. So, each zone the productions to take a row at all the columns you get productions. In the same matrix you take each column at all the rows in a column, you got the and you get the attractions.

Now sum of all productions in a zone in a study area from all the zones means what? Sum of all the elements in that matrix, if I sum it all the attractions take the total they also mean that sum of all the elements in that matrix, can that be different? Can they be different? The simple answer is no, I first I have a matrix with n by n matrix n rows and n columns. If I add the sales first column wise for each row, and then add all the rows.

And if I first add all the column wise all the rows and then add all the columns by total cannot be different, mathematically they cannot be different. You are true, we also know that we cannot proceed to a trip distribute developing a trip distribution matrix to match this row totals and column totals with a probable distribution of trips in that matrix. Unless the sum of the y, that means sum of the productions from all the zones is equal to sum of D j.

That means sum of attractions from all the zones within the study area, but I am saying they are possible because you are getting this from different models. So, estimates may be different, quite possible. So, if the estimates are different, I cannot go to my next step which is trip distribution with these sorts of estimate where my productions, total productions and total attractions they give different values, so I have to do something, that is the matter of discussion here. Now how to do it?

(Refer Slide Time: 09:01)



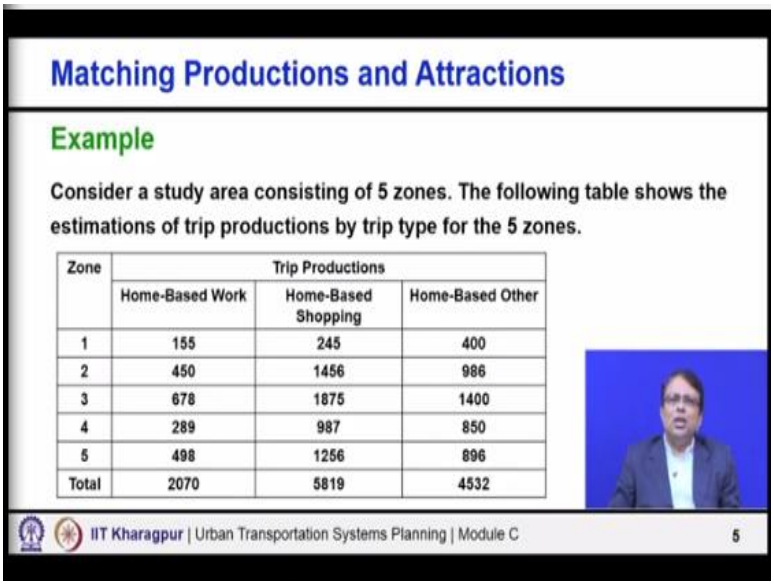
I would say that what is generally taken that the trip production models are generally considered superior, why? The simple reason is you have more explanatory variables, when you are developing a model. You remember you know the population, you can take car ownership, you can take income, you can take household size even when you are going for zonal based models you can take residential density, you can say value of land.

So many explanatory is variables are there. So, generally the production models are generally considered superior. Because of the kind of opportunity what we get in using variables and developing models. So, the normal practice is to assume that your total trip productions as you have got by taking sum of all productions by different zones, assume that as correct. Then I need to find out the factor, what is that factor?

Factor is that total trip produced assumed that is the correct figure that is 3 divided by sum over D j, sum over D j means sum over attraction from all the zones, you get a factor, this factor maybe more than 1, maybe less than 1. If your productions are more total productions from all the zones in the study area total production is more than the total attraction, then factor maybe something, if it is other way the factor may be different.

So, whatever you get this factor, then what we are saying multiply each attraction values you have modeled by this factor, multiply each attraction cell you have model, that for zone 1, this is the attraction zone 2, this is that traction zone 3, this is the attraction. Like that you have estimated attraction values each cell you multiply by this factor f, then what will happen? Then the modified attraction values the total will be same as the total for the productions from all zones, they will match. So, that is the way we handle this problem.

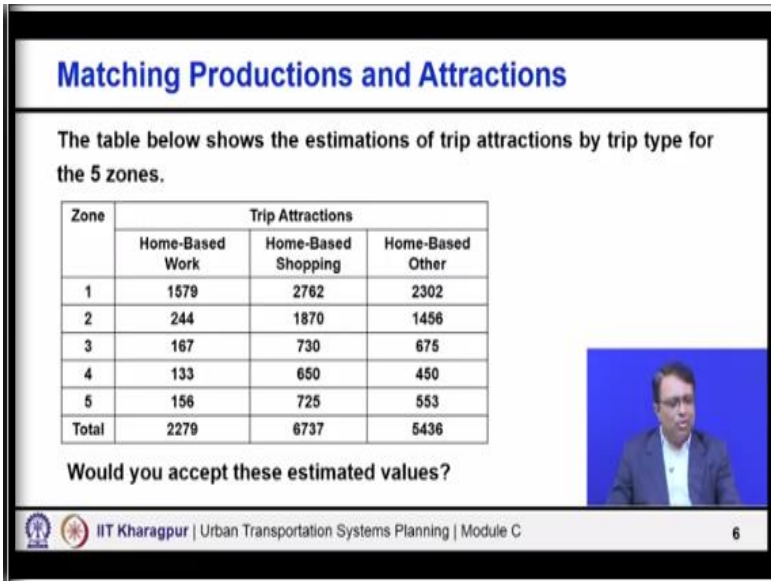
(Refer Slide Time: 12:04)



So, that is what I have told in this one. So, let us take an example to understand it a little bit better. We consider a study area of 5 zones and the following table shows that estimates of trip productions by trip type for all these 5 zones. So, I have 5 zones and you have modeled different model maybe you have got it from regression based model, maybe you have got it from the trip rates and using cross classification technique or the category analysis but you have got these estimates.

So, in each zone how much is home based work, how many trips are for home based shopping, how many trips are for home based others? These are arbitrary classification taken for example purpose only, not that you have to always consider these classifications and cannot consider any other, any other way you can express the trips, there is no bar. For example purpose I have considered 3 types of trips.

(Refer Slide Time: 13:20)



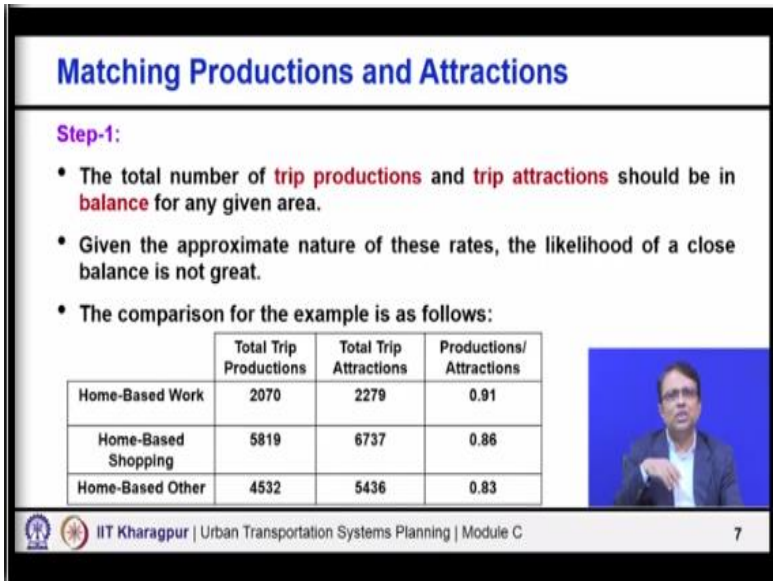
So, these values are known, exactly the same way, you also have trip attraction models, again maybe a regression based model, maybe a you use category analysis of cross classification techniques and use the rates and the quantity of the number to get the total trips attracted by different types of trip for different types of attractions. So, you got similar estimates home based work, home based shopping, home based other.

So, I have got one table which is share for the trip productions, for home based work, home based shopping and home based other for each of these 5 zones. I have a separate model which has given me the estimates of home based work, home based shopping and home based other in terms of attractions to each of these 5 zones. I am asking you would you accept this estimated values or can you just go ahead with these values to the next trip of trip distribution?

What is drip distribution? Now here we have row totals, row totals are there, so I know the productions, I have the column totals, so I know the attractions. And using this productions and using this attractions in trip distribution state, I want to find out each cell, T i j as I said. So, T i j sum over j is production, T i j sum over i is attraction. So, this row total and column totals are known and we want to go ahead to the next stage, that is getting the internal values, can you go ahead with this estimated productions and attractions to the next stage?

That is trip distribution or simply would you accept these values, so what we have to do? We have to check that each type of trip say home based work or home based shopping or home based other. Each type of trip whether the total productions and total attractions from the study area, whether they are matching or not, that is what we tried to do here.

(Refer Slide Time: 15:46)



So, the first step is the total number of trip productions and the trip attractions should be in balance for any given area. So, given that the approximate nature of the rates the likelihood of

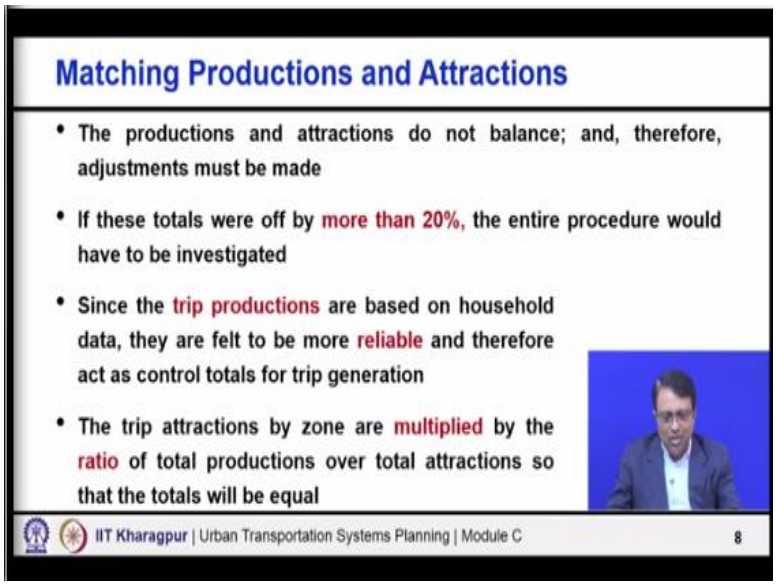
close balance is not great. Here, I am saying for example, maybe you have used category analysis. But similar type of thing is also possible maybe you have used to regression models as well, does not matter.

Because both cases you will get the trip ends, so this needs to be checked and something needs to be done irrespective of the modeling approach, whether you use cross classification analysis or use regression base model. So, what we try to compare here, when we try to compare the home based work, we find the trip production is 2070 from the complete study area, whether attraction is 2 to 79, they do not patch.

Home based shopping 5819, productions, attractions 6737 again do not match, and home based other productions are 4532 attraction is 5436, mismatch, so what we have to do? As I said earlier, we take the productions total as correct, taking that as the correct T, total trips, we calculate a factor, so what is this factor here? Factor here is first case home based work to 2070 divided by 2279, so 0.91.

Second case 5819 divided by 6737, 0.86, third case home based other 4532 divided by 5436, so you get 0.83. Now what I will do? I have to multiply these factors, home based work, home based shopping, home base other the corresponding factors to each cell corresponding cells in each zone.

(Refer Slide Time: 18:19)

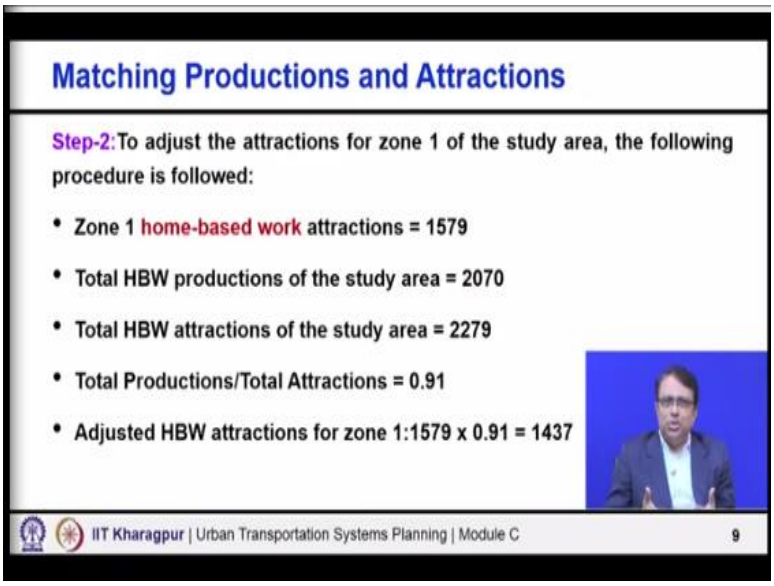


Let us say, first, but before we do that, I have stated what I have told you the productions and attractions do not balance and therefore we need adjustments. And here to remind you, this I have not mentioned when I was just discussing the things without referring to the slide, this particular slide. That mismatch is possible every case you will find a mismatch, you are it is unlikely that things will match.

Because there are different models and every model is approximate, you can never be 100% accurate in your models, model is always approximate. So, there will be difference but if you get the difference more than 20%, then what is advisable that you should thoroughly investigate. Somewhere something has gone wrong maybe in the modeling, maybe in calculation somewhere something.

So, if it is more than 20% then the entire process would have to be investigated. That means you may have to go back to trip production models and trip attraction models, if it is regression models. Otherwise you have to go back to the trip rates what you have developed, maybe somewhere some biasness is there, somewhere maybe the sample was not proper, so which has given some destruction in the trip rate. So if it is more than 20% you should go and reinvestigate, if not then simply do that judgments, what I said that is what is explained here.

(Refer Slide Time: 20:08)



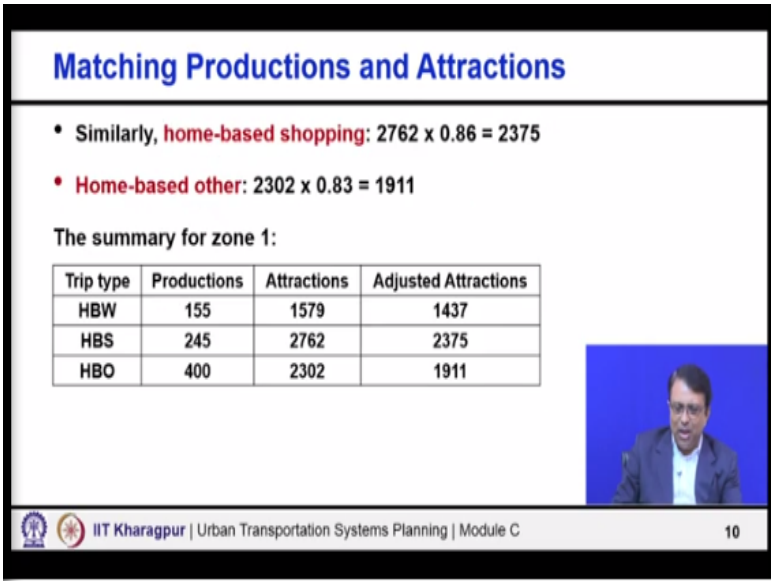
So, what we show, step 2, to adjust the attractions of zone 1 in the study area, why suddenly zone 1? Nothing special about zone 1, we have just picked up one zone as an example to explain you. And then we show the calculation for one zone and then similar calculations we do it for every zone. So, zone 1, home based attraction is 1579 that is to be adjusted, how we adjust it, we know that total home based work productions of the study area is 2070, this was already mentioned here.

You can see here in this slide, home based trip production 2076, I am mentioning it there again for reference. And then total trip attractions was 2279 and the factor was 0.91, that calculation is again shown here. So, our total home based productions home based work productions of the study area is 2070, and total home based work attractions of the study area 2279. So, what is my total productions by total attractions ratio is 0.91.

So, each cell in the study idea, each cell means home based work trip productions from each zone in the study area, we have to multiply it by 0.91, same thing we have done here. We know that home based attractions from this study area is 1579 equivalent has come around, that is for zone 1. So, home based work trip attractions from zone 1 equal to how much instead of equal you have explicitly share using colon, that is why the I got confusion.

So, it is equal to 1579 multiplied by 0.91, so my attraction gets adjusted to 1437, fine. Now exactly the same way for zone 1, I should also do the adjusted attraction for home based shopping and home based other. One case I have to apply 0.86 factor and another case I have to apply point eight 0.83, I have done exactly the same thing. So, exactly for zone 1, then shopping attraction was 2762, how this has come? You see for zone 1, 2762 is the home based shopping.

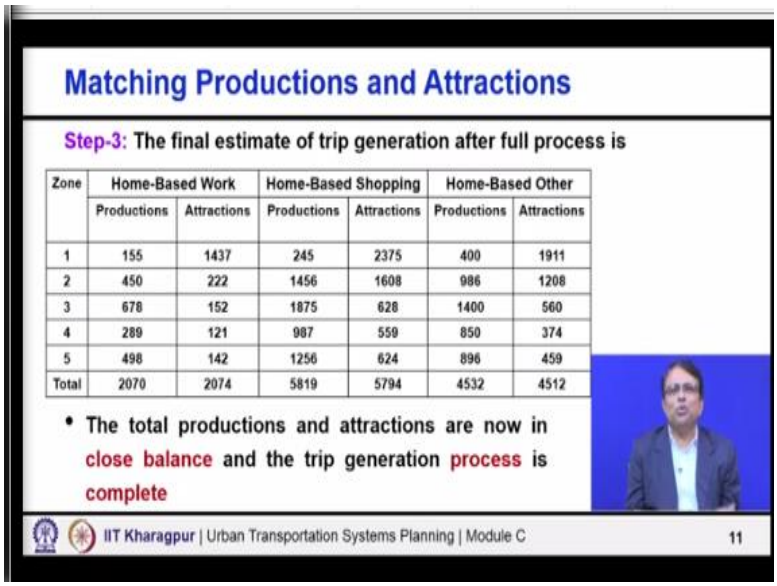
(Refer Slide Time: 23:09)



So that multiplied by what is the factor here, 0.86 it was calculation as shown earlier. So, this I get by modified attractions for home based shopping, and I also get the home based other 2302 and multiply it by 0.83, so get 1911. So, I have shown here the summary, my productions are 155, 245 and 400, this was originally given, all these remain intact without any modification. Because they are assumed to be correct, but the attractions were 1579.

Now modified attractions becomes 1437, home based shopping, original attractions were 2762, now the modified attractions are 2375, home based other the original value was 2302, now gets modified to 1911. So, what I will do, my final answer before we go to the next step, for zone 1 my productions are like whatever is shown here, and attractions are will be what, this adjusted attractions. So this explains you clearly how values are adjusted for a particular zone, exactly now following the same procedure for all the 5 zones we have to adjust the values.

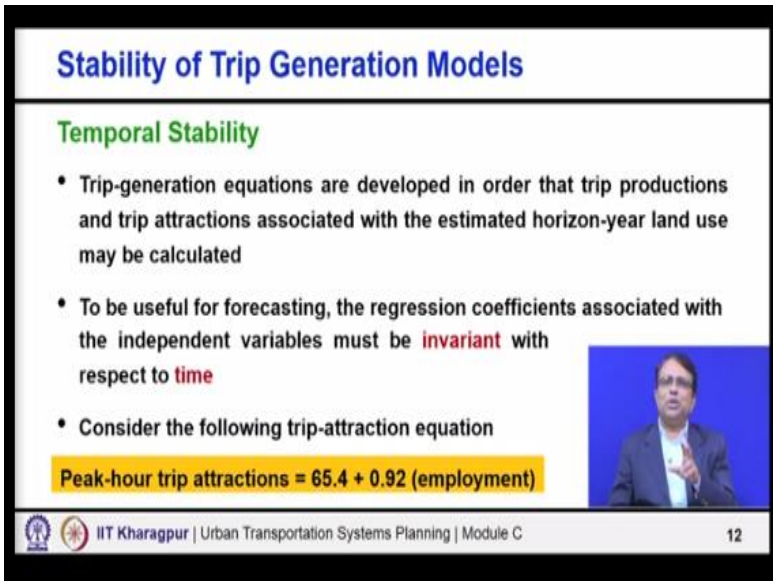
(Refer Slide Time: 24:48)



That's what I have done here and then I had plotted, so in each case, I have put it together productions and attraction, so that we can compare meaningfully. So, first say home based work the productions, attractions, you say the total is matching productions is 2070. Whereas attractions we are saying 2074 such kind of 2070 instead of that 2074 is coming because of the round off errors, so this kind of further little bit tuning you can do to match them exactly.

So, home base shopping also after adjustment 5819 and you get 5794, and home based other 4532 and you get 4512. So, these minor adjustments are because of the error in the round off. So, which may be considered further and you adjust the values marginally to make them exactly same. So, that way we can this is the procedure we can apply before we go to that next.

(Refer Slide Time: 26:06)



Now coming to the other topic, it is the stability of trip generation models, 2 types of stability. We say temporal stability and geographical stability, what is temporal stability? Temporal means over time, so the temporal stability means stability or applicability of the model over time. What sir is saying, it is confusing, because any model we develop we want to apply it for the future only, I have told you so many times.

Then why I am talking about temporal stability or the applicability of modeling the future. Model is always done for application in the future only, other is if I cannot apply for the future why you do a develop model, because base year anything you want you can measure, why is also measured, can be measured, dependent variable, independent variables also can be measured, then why we need to build relation?

Relationship because we want to apply the model after calibration, validation everything for the future, so then why we are talking about temporal stability? Actually most cases it will be temporarily stable, that means the coefficient estimates what you are using being a trip rate, being if the coefficient estimates in your regression model. They are likely to be temporarily stable, temporal stability will be there.

But some occasions may come which may say that the models are coefficient estimates are temporarily not stable, that is the kind of thing. So, it is just a point of caution maybe out of 100

case 99 case, at least if not 99 more than 90 cases, temporal stability will be there. But you must be aware of certain situation which may question the stability of the coefficient estimates, how? Let us give a few examples.

Say, let us say peak hour trip attraction model was been developed 65.4 + 0.92 into employment, what does it mean? 92% of the employee in a simple sense will make a trip in the peak hour and other 8% they come not in the peak hour but other times, very good, very simple model, very simple explanation. Now let us say, now you have 6 day week, everybody is following 6 day week or 5 and half day week let us say.

Now in fact some years back, I have seen that when the offices used to work, all offices use to operate with maybe 5 and half day week, 5 full days, Saturday half day. But then many of the government offices started operating with a 5 day week, so weekly work hours is defined. So, whether you work for 5 day means you have to totally do the work for the week. So, per day the work hours will be little bit more.

Suppose you say everybody shifts no problem, suppose you said now also it is their central government or government of India they operate say 5 and 5 day week, some of the state governments or maybe so many state governments, they operate sometime 5 days week, sometimes 5 and half day or 6 day. Now what will happen? Earlier total employment 92% of them were coming in the peak hour.

Now will all of them will come, because the office time will be very different, office time for a person for a 5 day week and for a 5 and half day week or a 6 day week, starting and closing time may not be same, so all employee may not come same time in the office. So, that one has to keep in mind or let us say during this COVID break India got very much accustom to work from home.

IT industries particularly they all work from home, others also work. But then it may happen also that if you are applying such kind of model in a situation where the COVID influence is there, many of the people maybe working from home. So, you cannot say if you are estimating peak

hour trips, you cannot use 0.92 this is also an example. During COVID maybe it is 0.92, during different phases of unlocking, these coefficients are expected to vary and will be much lower.

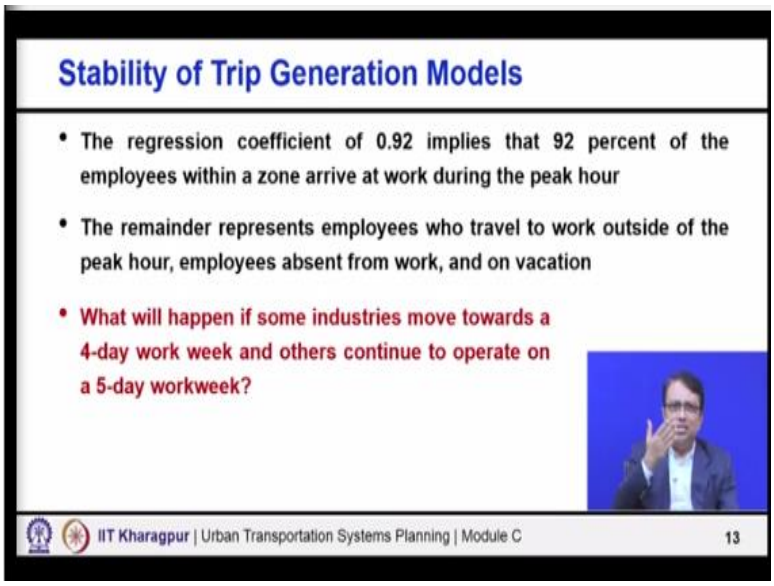
During unlocking also not all offices allowed all employees to come back, I mean alternate day or some working hour, some restrictions because they wanted to maintain social distancing. So, as per the government rules and regulations things were governed things regulated. So, not suppose you have developed a model earlier which were giving you a 0.92 coefficient, now will you also use the same coefficient, you should not use.

Or suppose maybe the once this COVID phase is completely over then also still many of the IT companies may operate. That they may not really take back all their 100% employee in the office every day as it used to be in the pre COVID situation or why only IT industry maybe even other offices also, many other offices, not all. So, they may still say that, ok, let by I will use the smaller office space and let some of my employees work from home every day, why should I take this burden of this physical infrastructure, higher end, higher maintenance charges and everything.

So, the infrastructure I can still reduce, I can use a smaller office and every day every employee need not even come to the office. Many cases people do it in a broad work from home was not a new thing. But in our case it was new but and we very quickly, I should say that we adopted it. So, I can see in the future maybe not all employees in the IT sector and many other sectors, not all employers will come back to the office, the way they were doing it earlier.

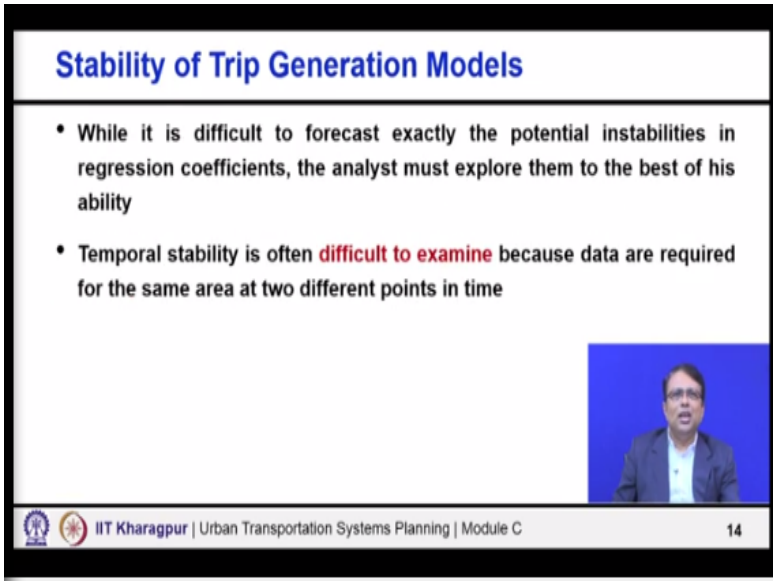
So, maybe this 0.92 will not be valid, yes, I understand that COVID does not happen every day, COVID outbreak or this kind of situation does not you do not face every day. But for you to remember that yes, it will be applicable but some cases, some context it may not be applicable as well. So, before you apply just think little bit about stability, temporal stability.

(Refer Slide Time: 34:19)

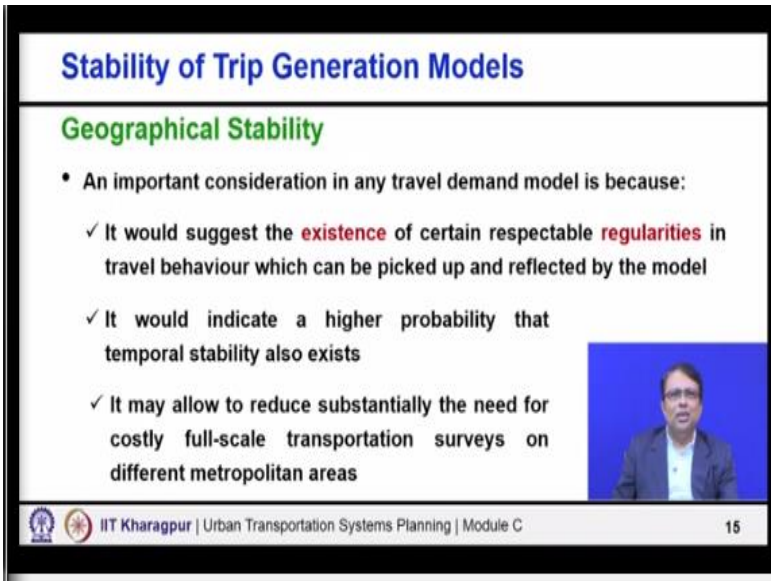


The same way what I said, I wrote here the same thing.

(Refer Slide Time: 34:25)



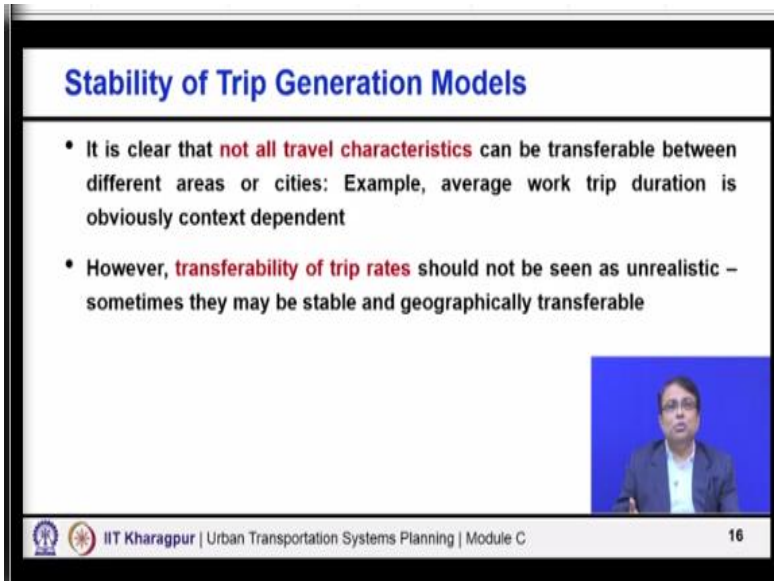
(Refer Slide Time: 34:27)



The same way is the geographic stability, what is geographic stability? We develop the model in for one area or one city, can I transfer the model geographically? That means can I transfer the coefficient estimates whatever I got here, are these coefficient estimates geographically transferable from to another city? Maybe I have done the work in the context of Delhi, Can I use the same coefficients for Mumbai, for Mangalore, for Chennai, for Calcutta, can we do that?

That is the question we are trying to answer here. Now there are why they are important, I have explained some points here which I need not read every point. There is certain deeper implications of course.

(Refer Slide Time: 35:33)



But can we really transfer it, I would say that it is clear that not all travel characteristics can be transferable between different areas or cities. Say for example, average trip duration work trip duration is obviously context dependent. So, different cities have got different thing you go to one city where it is very common, maybe it is a common culture. Evening people go out to restaurants, shopping and recreational trips maybe more. In another cases go to another city, it may not be really so vibrant, right.

One city may work really if you go to Bombay maybe it is very different than what you find in Calcutta. So, it is absolutely true that not everything is geographically transferable. City characteristics, the city specific travel behavior, context, ambient, policy many things will actually influence this rates or the coefficient estimate. So, it is not really possible that you develop one and the cities are so much heterogeneous in terms of income, in terms of approach, in terms of attitude of people, behavior of people, it is never a simple task.

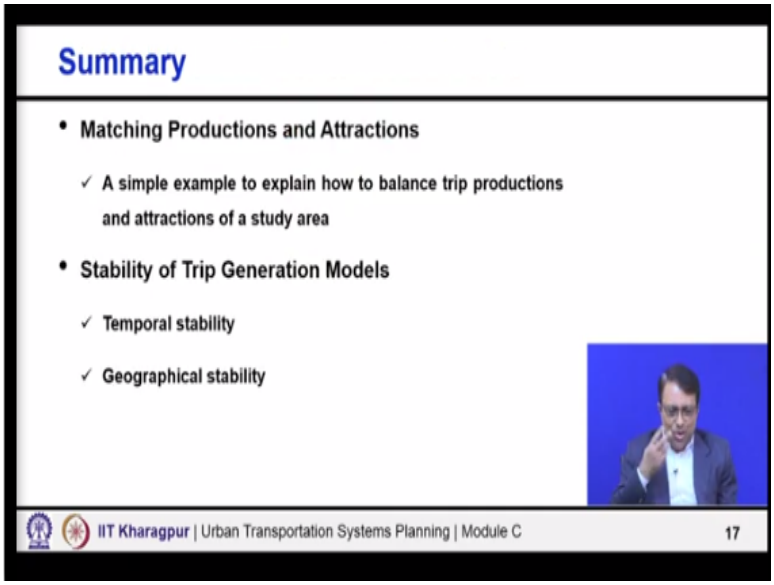
So, it is really too simple to assume that we do it in one case, and then everything we will get will be geographically transferable, normally I would say no. However also things considering or acknowledging that, I should also say that transferability of trip rates should not be seen as completely unrealistic, maybe sometimes they may be stable and geographically transferable but not in all cases.

So, if you are trying to use the rates, transferring the rates or transferring the models and the coefficients for trip generation models in the context of trip generation model. Please be careful, duly consider that I am transferring these rates, what you are transfer that matters a lot. As I said not everything is transferable unlikely to be transferable but some cases some specific aspect maybe transferable as well.

So, before you transfer any value, any coefficient, any rate, please carefully think. Apply your logical mind to see that are these 2 things comparable does it sound alright to assume that this is same thing is valid here. Even you across towns even in one region, you consider one mega city and all the other cities, small towns, other levels of cities different types of cities, things may not be completely transferable unlikely.

But as I said, the other part is it is not completely unrealistic, that nothing is transferable, geographically transferable, I do not want to be so stringent as well in terms of delivering my lecture. I would say in a humble way, that many things are not clearly not transferable. So, many things are not transferable very clearly, some of the things may be transferable. So, one has to think what we are trying to transfer? What specific context and the context of the cities we are trying to compare and then apply it. So, with this I close.

(Refer Slide Time: 39:44)



So, what we discussed here is 2 important aspects. In the closing part one is matching the productions and attraction we said, why the problem may comes and we said what are the way outs with an example and then said also the stability of the trip generation models. Again applicable for regression based model applicable also for the cross classification or category analysis or trip rate base model.

Both cases they are applicable and applicable means both cases they are (()) 40:20) in a way temporal stability and geographical stability. As I said, most cases temporal stability will be there, but you must carefully examine. Because some cases, the models, coefficient estimate, rates may not be applicable. Similarly geographical stability is also important, not everything can be geographically transferable to another place.

But of course, some values may be as well transferable. So, you have to apply your judgment and think carefully before you make the models temporally transferable and geographically transferable. Thank you so much, thank you for your kind attention.