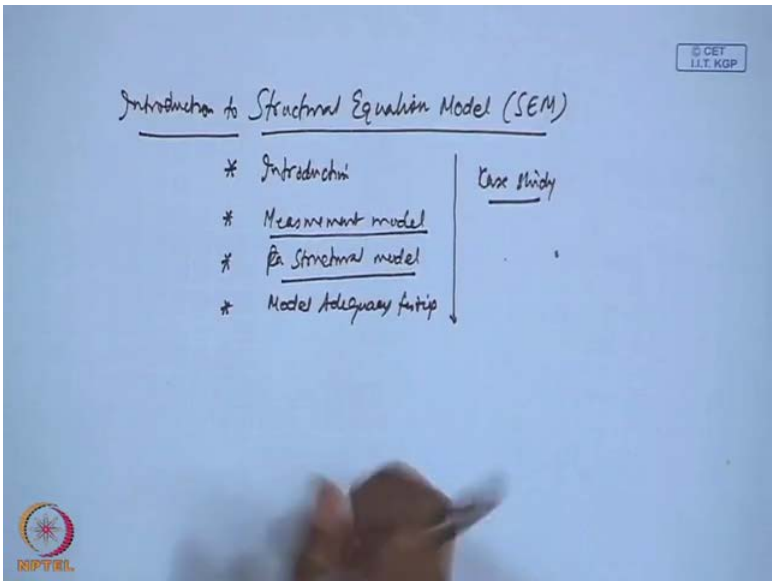
Applied Multivariate Statistical Modeling
Prof. J. Maiti
Department of Industrial Engineering and Management
Indian Institute of Technology, Kharagpur

Lecture - 38
Introduction to Structural Equation Modeling

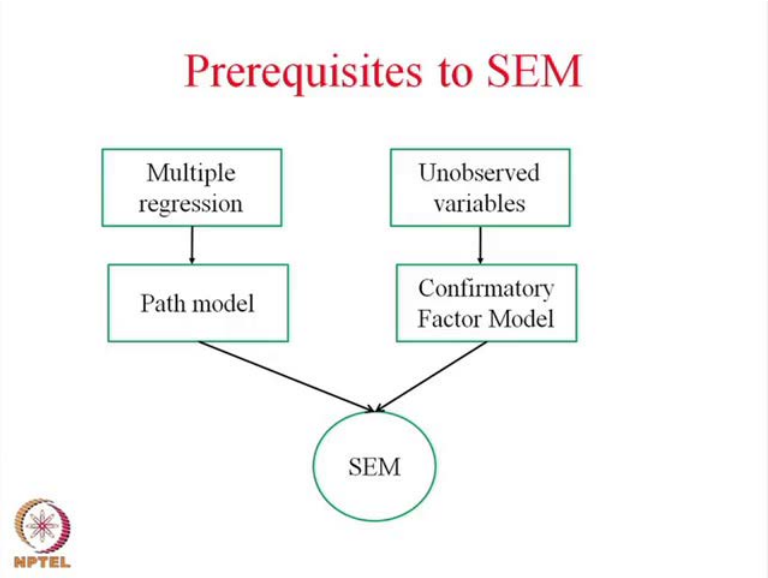
(Refer Slide Time: 00:27)



Good evening. Today, I will discuss structural equation modelling also known as SEM. It is a little bit complex, advanced. It requires knowledge of some other tools and techniques what already covered in this course. I have partitioned this total lecture into three primary headings. One is today's introduction. So, today I will show you that what is structural equation modelling and how it is applicable to real life problem solving.

Then also we will discuss that what are the components of structural equation modelling. For example, it consists of measurement model, it also consists of path model or structural model. So, we will elaborately discuss this measurement model in next class. Then we go for path model discussions, and then there will be the model adequacy testing, adequacy testing and obviously I will discuss case study, one case study all through. Today's lecture is introduction to structural equation modelling.

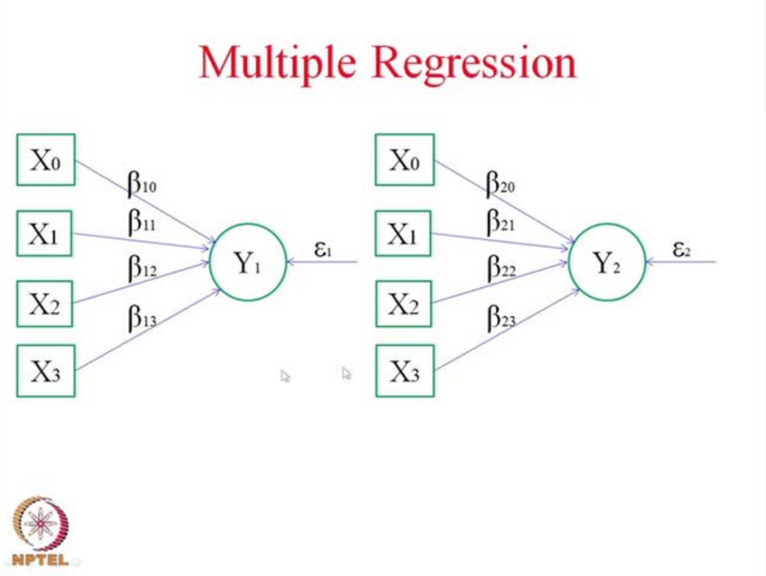
(Refer Slide Time: 02:39)



Let us see that what are the prerequisites to structural equation modelling you must know multiple regression. You must know path model. Although I have not described path model till date, but within the system, we will be describing path model. You must know confirmatory factor model whose prerequisites are exploratory factor model so that unobserved variables can be measured through exploratory factor model.

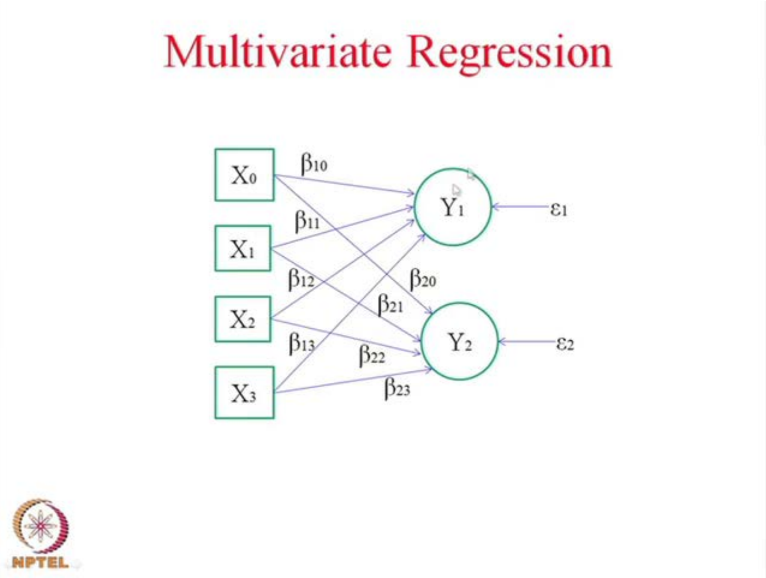
Now, what is the difference between confirmatory and exploratory factor model that we have already discussed? So, that means if you know confirmatory and exploratory factor model, you know path model, which is basically the structural model. Then you will be able to apprehend structural equation modelling.

(Refer Slide Time: 03:36)



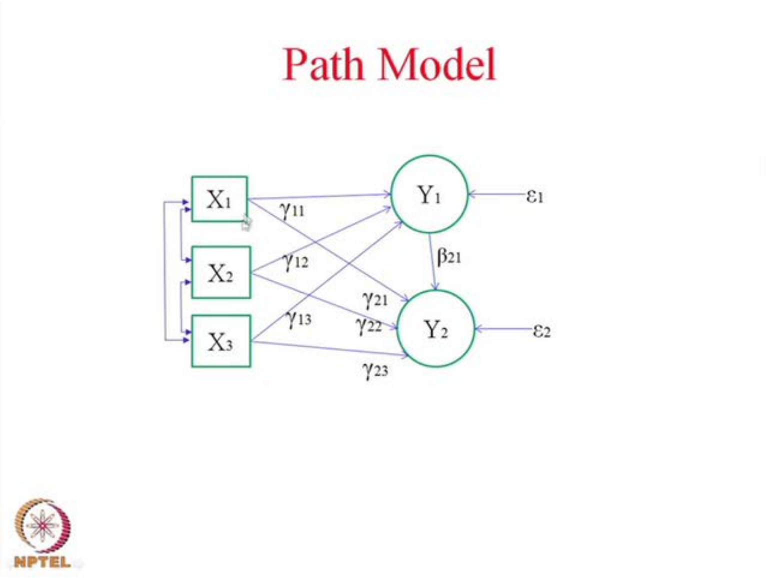
Let me recapitulate little bit of the prerequisites. I know that all of you have seen this type of pictorial representation for multiple regression, where Y1, here Y1 is affected by several variables as well as Y2 is also affected by the same sets of variables X1, X2, X3.

(Refer Slide Time: 04:09)



When we have more than one dependent variable, we have seen that that multivariate regression is preferable. Now, the question comes if these variables are also correlated and the independent variables are also correlated or if there is even the causal relationship between that Y1 and Y2, so what will happen?

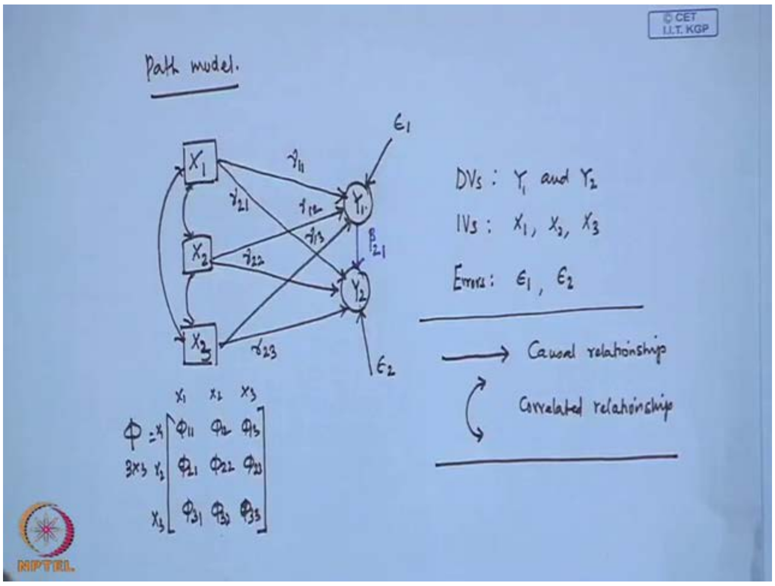
(Refer Slide Time: 04:37)



So, ultimately in order to incorporate the covariance relationship between the X variables, so you create something like this where X1, X2 and X3, these three are the independent variable. Y1, Y2 are dependent variables, but the difference of this structure with multivariate regression structure is that in multivariate regression, you are assuming that X1, X2 is equal to X3 is truly independent.

They are not correlated and there is no causal relationship between Y1, Y2 also. But, in path model, what you are making? You are considering that X1, X2 as well as X2, X3 or X1 X3, they can vary. In addition, what we are seeing that Y2 can be affected by Y1 also in addition to the effect of X1 to X2 to X3. Now, let us see equation for this path model. So, we are deriving a migration from path model.

(Refer Slide Time: 05:51)



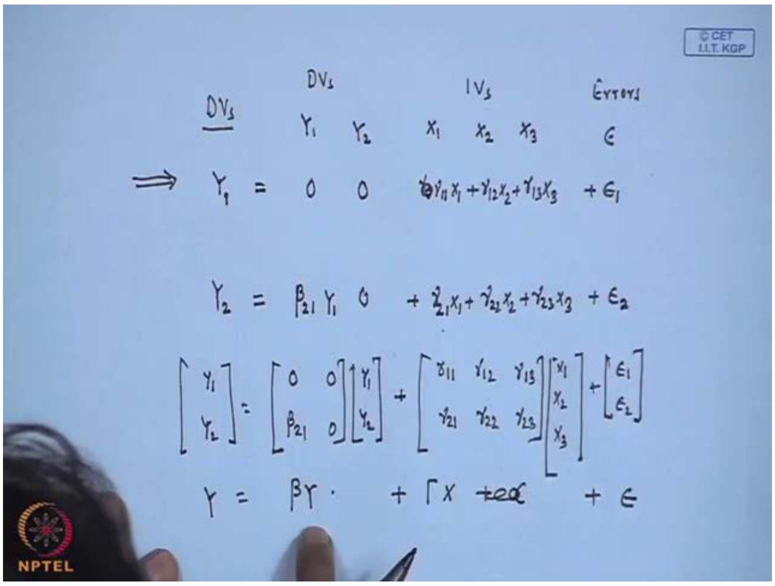
Given like this X1 is an independent variable in this model, but correlated weight other independent variables. So, that is basically this correlation can be captured through a matrix, which is phi. If we can say this phi is a 3 cross 3 matrix, so we can write like this, so X1, X2, X3, X1, X2, X3, it is phi 11, then we can write phi 21, we can write phi 22, and we can write phi 31, phi 32, and phi 33. Definitely, this is a symmetric matrix and this will be phi 12, which is basically phi 21 equal to phi 21. Then phi 13 is equal to phi 21 and phi 23 is equal phi 32.

In case of multivariate and multiple regression, we have assumed that this is structure of diagonal elements are 0. Now, we have considered here two independent variables Y1 and Y2 and we are assuming that X1, X2 and X3, they are affecting Y1 with coefficient like this. Gamma 1 is affected by 1; gamma 1 is affected by 2, gamma one is affected by 3. Similarly, from independent side point of view, this gamma 2 is affected by 1. That means Y2 affected by Y1, Y2 affected by X2 and Y2 affected by X3.

Let me repeat here, Y2 affected by X1 that is coefficient is gamma 21, Y2 affected by X2, coefficient is gamma 22, Y2 affected by X, coefficient gamma 23. As these are regression lines like equations, so this is having an error term. Now, this also will have an error term. In addition, here we are assuming that this Y2 is also affected by Y1. If this is the case, you have to give a symbol like beta Y22 and Y11, this beta 21.

So, if you want to write an equation for this, then, there are how many variables are there? How many variables are there? There are different types. One set is DVs that is Y1 and Y2. Another set is IVs that is X1, X2, and X3. Another set is errors that are epsilon 1 and epsilon 2. Then what types of relationships are available here? One relationship is like this. If you see that X1 toY1, this single arrow, this is known as causal relationship. Another type of relationship with this giving here like this card bidirectional, decision on edge correlated relationship. So, these two types of relationships are available in this. These two types of relationships are estimated in path model.

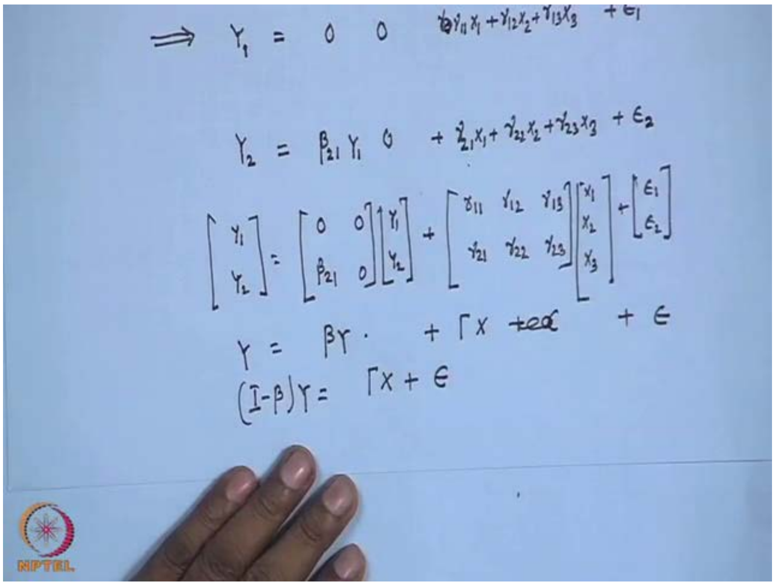
(Refer Slide Time: 10:44)



Now, if you want to write down the equations for these, straightway we can write one is DVs. I can write Y1 and definitely Y2 is there. This side also you can write DVs, you can write IVs, you can write IVs. How many DVs are there? Two DVs, Y1 and Y2, X1, X2, X3, then errors are definitely this. So, the equations will be like this, Y1 equal to that, Y1 is not dependent on that, Y1 again it is not like this. Similarly, if I see that Y1 is dependent on X1, X2 and X3 and this arrow term, not Y2, so nothing will be written here. It will be straightway gamma 11, then X1, then plus gamma 12 X2 plus gamma 13 X3 plus epsilon. This is the regression equation for Y1. Similarly, for Y2, if we see that Y2 is also affected by Y1, so we can write here that beta 2 is affected by 1, Y1 plus gamma 21 X1 gamma 22 X2 gamma 23 X3 plus epsilon 2.

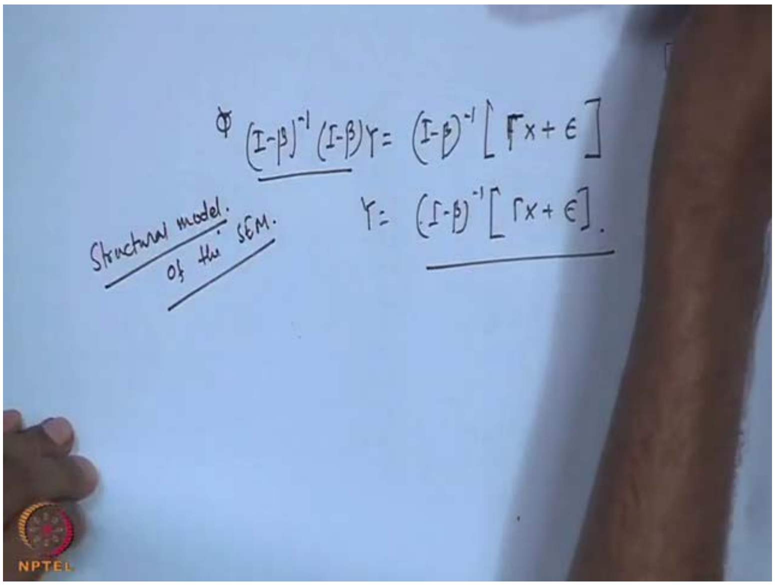
So, in matrix form, you can write Y1 is equal to Y2. Then this side you see here it is nothing means 0, 0, 0 so 2, 1, 0, definitely again you are writing Y1, Y2, which is the difference from regression model. Then here you have gamma 11, gamma 12, gamma 13, gamma 21, gamma 22, gamma 23 into X1, X2, X3 plus you are getting epsilon 1 epsilon 2. This is your path model or Y equal to beta Y plus gamma X plus epsilon, beta Y plus gamma X plus epsilon.

(Refer Slide Time: 14:00)



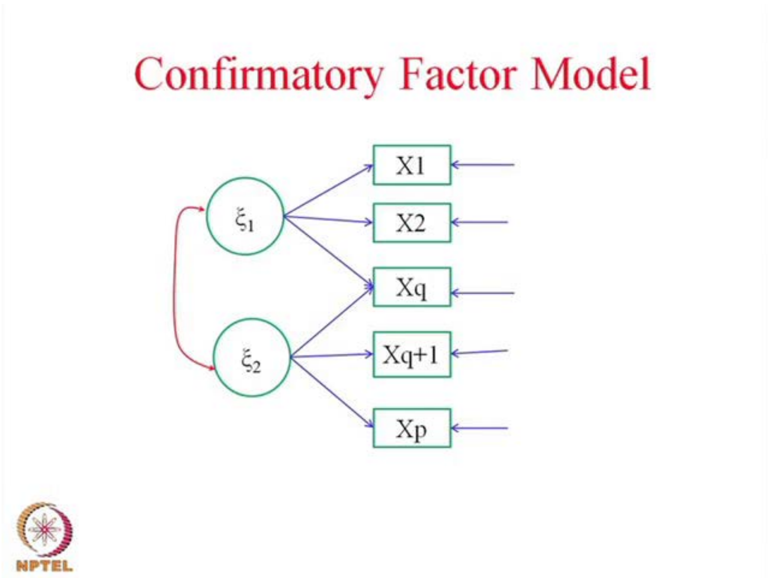
If you do little more manipulation, then this will be I minus beta Y equal to gamma X plus epsilon.

(Refer Slide Time: 14:15)



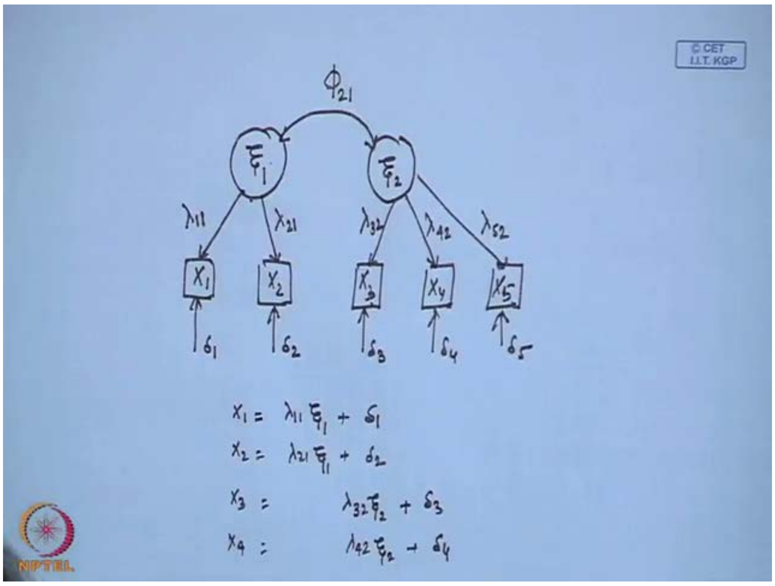
Then, further you may be interested to write this that Y. If I multiply Y minus beta inverse here with I minus beta Y, then I minus beta inverse, you are getting that gamma X plus epsilon. This will be I. So, Y equal to I minus beta inverse F X plus epsilon. The details of this how to estimate this parameter beta, psi and all those things, epsilon, phi, all those things will be discussed under structural equation model, structural model under structural model of the structural equation modelling. So, let us assume that we know this for the time being that this correlation is possible to understand.

(Refer Slide Time: 15:18)



Then, you come to confirmatory factor model. I think you can collect yesterday's lecture will particularly I think last two lectures were on cluster analysis before that.

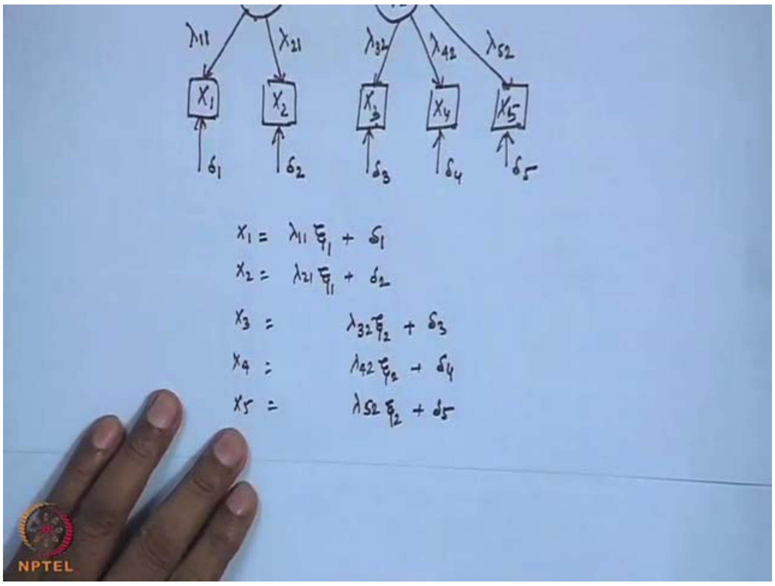
(Refer Slide Time: 15:35)



Then, if I write like this what you are doing, then here in this case, let us concentrate on only two factors and let they are related with X1. First one is related with X1, X2 maybe, second one is related with X3, X4 and X5. There is definitely covariance between the two, that is phi 21 and each of the observed variable X1 are caused by psi 1here. These are caused by psi 2, psi 1 and psi 2. Now, if I write this one is lambda, then for X1, I am writing 1 for psi 1. I am writing 11 X1 is affected by psi 1, then lambda 2 X2 affected by psi 2, then lambda 32, then lambda 42, then lambda 52.

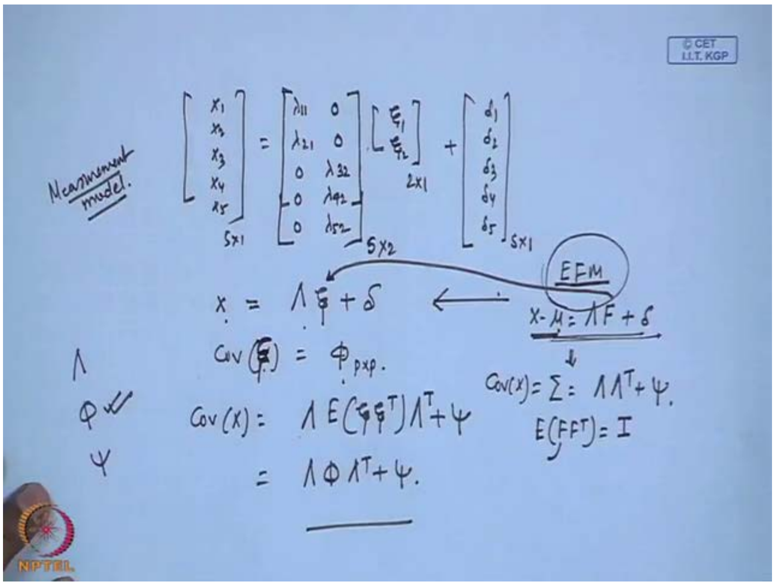
So, this one will be delta 1, this will be delta 2, this will be delta 3, this will be delta 4, this will be delta 5. This is what you have seen in confirmatory factor analysis. We have written this one in this manner that X1 equal to lambda 11 psi 1 plus delta 1, X2 equal to lambda 21 psi 1 plus delta 2, X3 equal to lambda 32 psi 2 plus delta 3, X4 equal to lambda 42 psi 2 plus delta 4.

(Refer Slide Time: 17:53)



X5 is equal to lambda 52 psi 2 plus delta 5. So, the resultant matrix you have written like this.

(Refer Slide Time: 18:07)



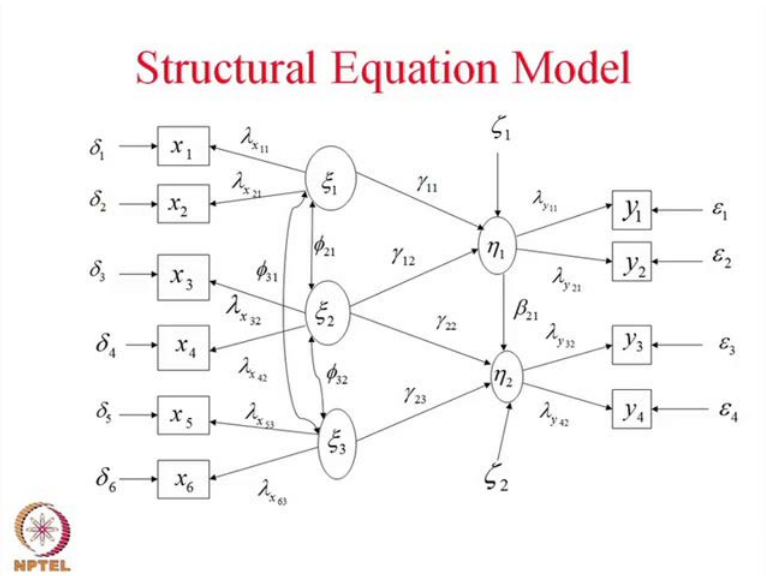
So, X1 X2 X3 X4 X5, this is equal to your psi 1 psi 2 is there, so we write down lambda 11, 0, lambda 21, 0, then 0, lambda 32, 0, lambda 42, 0, lambda 52. So, this is 5 cross 1. This is your 5 cross 2 and then psi 1 and psi 2, this is 2 cross 1 and plus delta 1 delta 2 delta 3 delta 4 delta 5, this is five cross 1.The resultant you can write is X equal to, this is you lambda, this one is psi plus we can write this as delta. If you recollect gamma exploratory factor model, what we have written there is X minus mu is equal to lambda F plus delta. It is coming similarly; only that minus mu is not there, but here also we assume that this subtraction is made.

So, actually what happened is that you are getting the same factor model. There is one difference in this case. The difference is in this exploratory factor model, what you found out, when you found out that covariance of X, which is sigma that one is lambda lambda transpose plus psi because we assume that expected value of F F transpose, this is I that covariance between the factors. In this case, we are assuming that covariance of X is phi depending on, so there are p variables, it will be p plus p.

So, as a result, the covariance here we are assuming covariance of psi, I am extremely sorry covariance of psi, this one correspondence to this; this corresponds to this covariance of psi is this. Then covariance of this will be lambda expected value of psi, psi transpose lambda transpose plus phi lambda transpose plus phi. So, for the time being, we stick to this only.

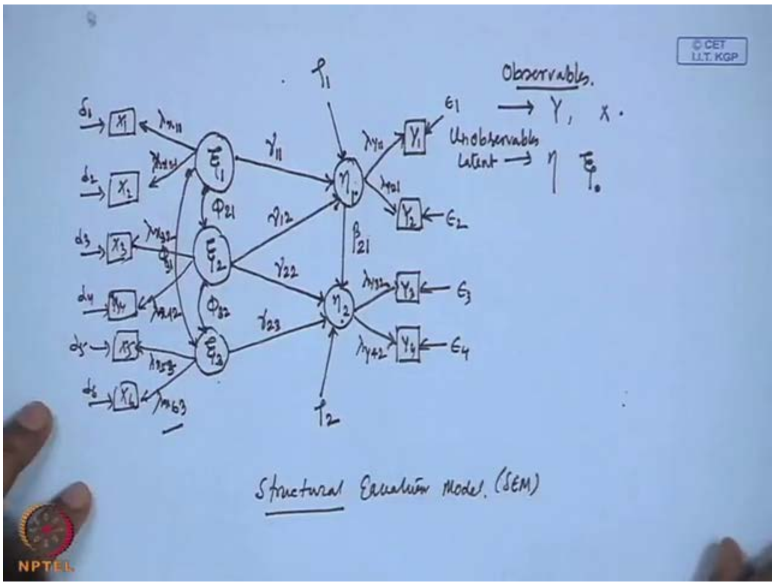
Now, question is how to estimate this lambda phi and psi and all those things related to exploratory factor analysis. We have seen that how this can be estimated, but in exploratory initial direction, this phi matrix was not there. Here, phi matrix is there. Also, we will be discussing this under measurement model.

(Refer Slide Time: 21:54)



So, these two are very essential for structural equation modelling. Now, see one big picture where we are doing like this. First, let us think of that there is one variable eta 1, another variable eta 2.

(Refer Slide Time: 22:14)



These two variables are affected by some variables called psi 1, psi 2, and psi 3. Let the link is like this eta 1 is affected by psi1 as well as psi2 and eta2 is affected by psi2 as well as psi3 and eta 2 is also affected by eta 1. So, under this condition, if you see the structure path model equation, so in path model equation let me find out this path model.

Yes, this is our path model equation. So, if I keep side-by-side this to side-by-side, so what is happening here is you see, we are writing eta 1 in place of Y1, eta 2 in place of Y2, psi1 in place of X1, psi2 in place of X2, psi 3 in place of X3. Here, everything is linked with all the X variables are linked by Y. Here, we are restricting that there were that eta 1 is affected by psi1 and psi2 and eta 2 is affected by these two. So, there has been this structure is similar to the structure except the some of the links are omitted.

So, I can write this that gamma 1 eta 1 is affected by psi 1 means gamma 11. So, this one is gamma 12. Then this is gamma 22. This is gamma 23 and this one is beta 2 affected by 1. The coefficient as usual we are writing here. Now, the difference between this path model, what I have shown, I have shown in path model in terms of Y and X. When we talk about in terms of Y and X, we say they are observable.

Now, when we are saying in terms eta and psi, they are latent and unobservable or latent that is the difference. So, if this is the case, then what you require to do? You require measuring these unobservable or latent things. We will see this, but before that, some more things that these are cohering, let me write. This is nothing but phi 21, this is my phi 32, and this is my phi 31. These structures are there. So, as I told you that this eta and your psi, they are immeasurable, unobservable, latent. We require measuring this. How do you measure? We have seen earlier that our confirmatory factor analysis or exploratory factor analysis will help us to measure. What it is? These are latent, this is latent variable, this is also latent, and these are manifest variables.

So, these manifest variables are used to measure these latent variables. Now, we will do the same thing here. We may create here like this. Later, we denote it is as Y1, this one is Y2 here, also Y3, and Y4 and definitely there will be error. That error will be suppose epsilon 1, this error is epsilon 2, this error is epsilon 2, and this error is epsilon 4, error head. Please give it properly. Then what is the difference between this structure and this structure?

If I place these to linking to three indicators, linking to two indicators, variables here eta 1 is linked with two indicator Y variables. Eta 2 is linked with two Y variables. What we are saying these two are latent, these are the manifest variables. This eta 1 and eta 2 can be measured using this. This is nothing but a confirmatory factor analysis. This is what confirmed confirmatory factor model is.

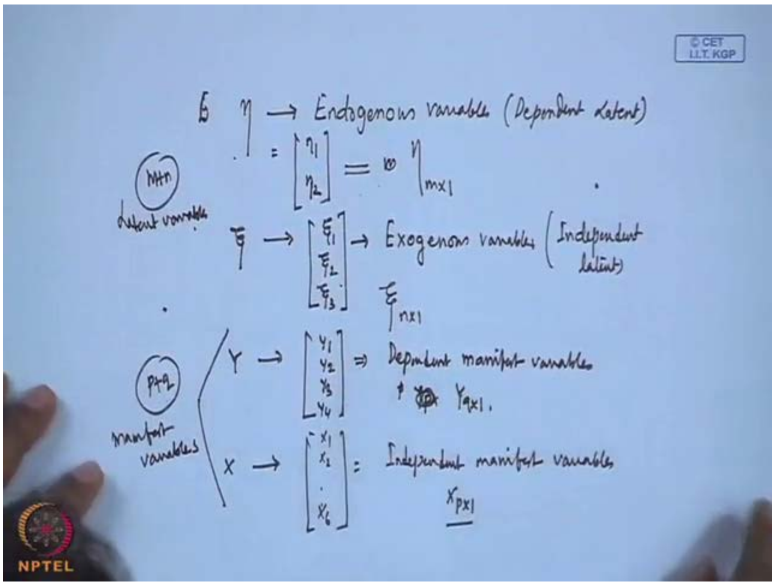
So, you can give a notation here. So, we have given lambda I think. Let us use one another subscript Y and this is 11. Then this one will lambda Y2 is affected by 1. Then this will be lambda Y3 is affected by 2. This is lambda Y4, is affected by 2. So, we also require measuring this. So, let us assume that this can be measured by indicated manifest variable X1. Then your X2, then this is your X3, this one is X4, then your X5, X6 and this side delta 1, this is your delta 2, this is delta 3, delta 4, delta 5 and delta 6.

Now, if you look into this side, what is there? Is there any difference between this, between this and this here psi1, psi2? You have taken three latent variables and these are the indicator five and indicators two are linked with psi1 and three linked with psi2. Here, there are sixteen indicators; two each are linked with psi1 and psi2. So, that means this is also a factor analysis part, confirmatory part.

So, then this one, we can write like this lambda, we are taking as X11, this is your lambda X21, this is your lambda X32, lambda X42, lambda X53, lambda X63. What is missing here? These are from confirmatory factor point abilities, these are the dependent, and these are the causes here. These depend on these causes. So, arrow head, arrow terms are given, arrow terms are given, but here from the path model point of view, these are the dependent side and this is independent.

So, you have to give eta. This is what is known as in path diagram, this is known as structural equation model. Why? It is structural equation model because the latent variables, they also have certain structure, some structural relations, which they are. For example, psi1 is affecting eta 1, but eta 1 is affecting eta 2. So, there is a structural path from psi1to psi7, though here psi1 is not directly affecting eta 2, but psi1 can have effect through eta 1 to eta 2. So, that means that the relationship structure is correctly preserved. So, that is why, this name is structural equation modelling is given here. Now, how to have the estimation of this? This is a big issue. We will be describing later on that all those things, but here, we will want to give you some names in the structure.

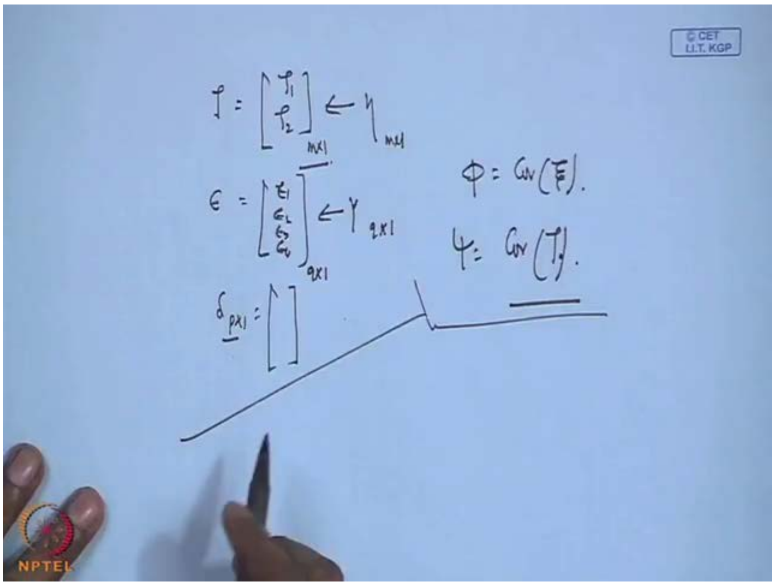
(Refer Slide Time: 31:33)



I want to give you some names of the variables. In structural equation modelling, we will use the term like these eta and all eta will be termed as endogenous variable. Endogenous, this is basically dependent latent in the example given, this one each this eta, this is basically eta 1 and eta 2. Now, there is another variable, which is psi, which is in this example. This is psi1, psi2, psi3, which is known as exogenous variable, exogenous variables. This is independent, latent.

So, in general, your eta can be m cross 1, psi can be n cross 1, those many things will be there or any other means, you can use something else. Then we have Y, which is in our case, in this example Y1, Y2, Y3 and Y4. So, this is known as dependent manifest variable. It can be p, q cross 1. Similarly, you have X here, we have X1, X2 to X6. These are independent manifest variables. It can be X that is p cross 1. That means in totality, you will be having p plus q manifest variables and m plus n latent variables. In addition, you have other variables like psi.

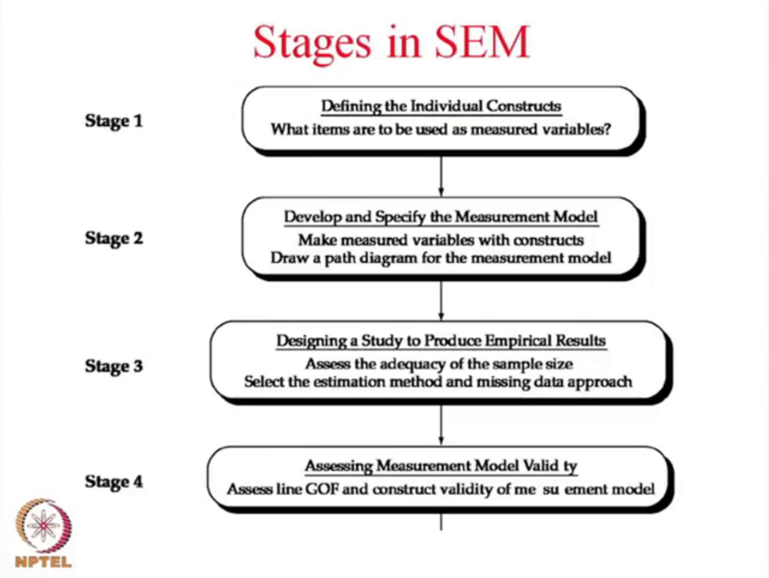
(Refer Slide Time: 34:37)



Other variable in this case, for example is zeta. Zeta is nothing but the error terms. In this case, it is zeta 1, zeta 2. It is related to eta. So, depending on the eta, if eta is m cross n, then this also be m cross, m cross 1, m cross 1. Then another one is your epsilon that will be related to epsilon 1 to epsilon 2 to epsilon 4. In this case, this is related to your Y. It all depends on what is the value of Y. So, this will be same.

Similarly, related to X, there is delta. The delta is your p cross 1 variable vector, these three. You have another parameter called phi which is basically covariance of this. You will also have, you will be getting psi as your covariance of your zeta. So, there are a huge, every large number of, what I can say, parameters to be estimated. Let us assume that yes, everything is possible and we are going for its stages.

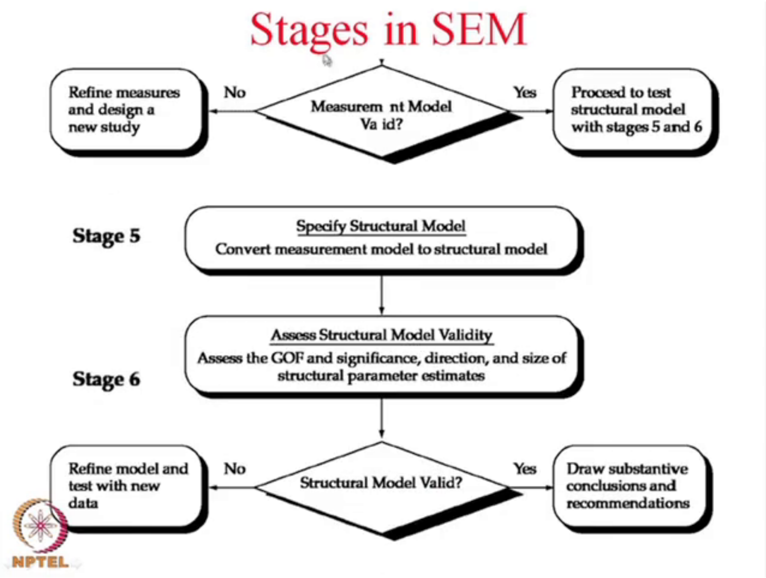
(Refer Slide Time: 36:10)



What is stage one is? It is defining the individual constructs, what items are to be used as measured variables. Then stage two is develop and specify the measurement model, what is the confirmatory factor model, make measured variables construct, draw a path diagram for the measured variables. Then stage three is designing a study to produce empirical results. All multivariate modelling, as such, any model building case, the design, study design is very important. Here, what we are saying, assess the adequacy of sample size, select the estimation method and missing data approach.

This is true for multiple regression, multi variable regression and path model, factor analysis, compound principle, compound analysis, structural model, cluster model, everywhere, it is applicable. Then assessing measurement model validity means these few first four stages; we are talking about the measurement model. Then what is this measurement model? Measurement model is this part, this factor analysis part; this is your measurement effect. Then what is your structural side? Structural side is this one. So, first what we have, it is said that in the stages and find out the manifest variables, then develop the constructs and also measure this, the adequacy of this model.

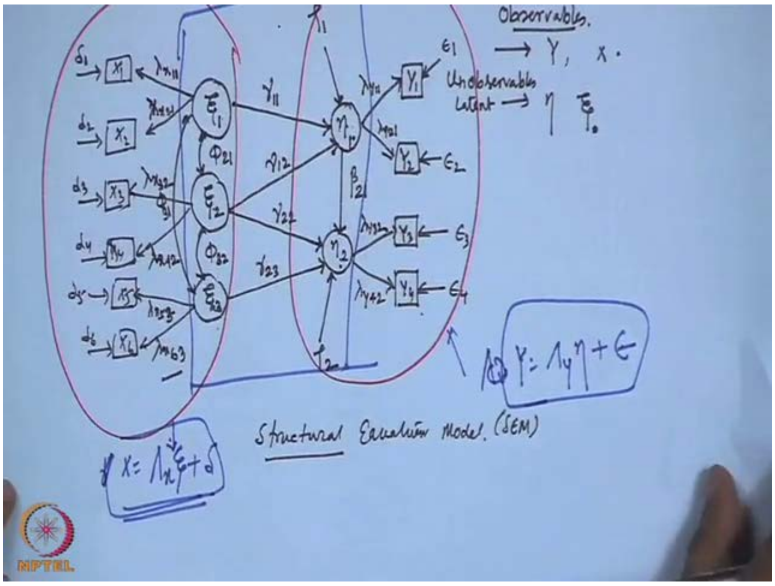
(Refer Slide Time: 38:06)



Then, you have to see if the measurement model is valid, then you have to go through five to six. If the measurement model is invalid, then you have to redefine the measures and other things. Now, specify, then your structural model is coming. Specify the structural model. Convert measurement model to structural model. I think you have seen here ultimately. I will repeat this again. Then stage six is assessing structural model of validity. Then if structural model is valid, then draw substantive conclusion. If it is not valid, refine the model and test with a new data.

In this case, what is it actually given? It is given in a two stage model building. One is first you develop the measurement model. First you develop the measurement model irrespective of your, in this case, forget about the structural part. What you have? You have these many X variables, these many Y variables. Irrespective of eta and psi, you first go for it on measurement model where you will basically factorise, confirmatory you factorise this eta 1, eta 2, psi1, psi2, psi3 from Y1 to Y4. This is the first.

(Refer Slide Time: 39:41)

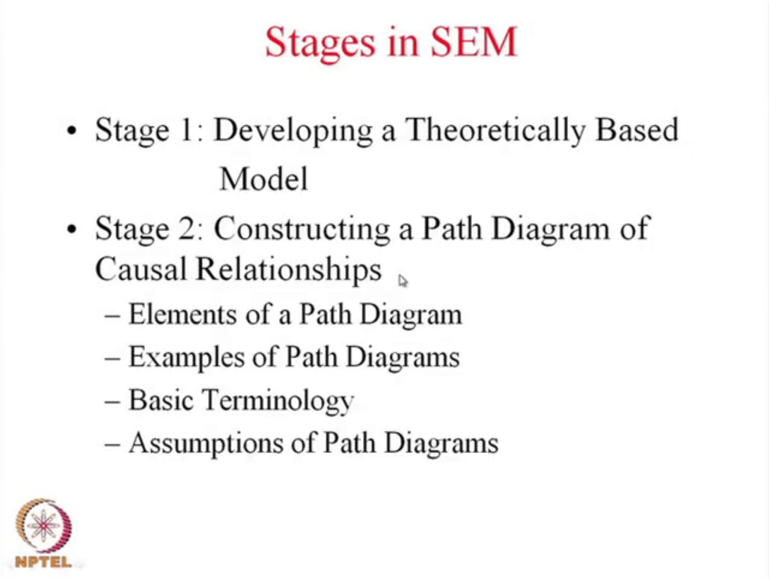


Once you are happy with this, so that means that if I just follow this one this side, it is X is lambda x psi plus delta. This is my measurement model for this site. This side, it will be, Y equal to lambda y into eta plus epsilon. This is my contrary path. But, the process of model building stages is that that you go for measurement model. First consider all these, and then do a factor, confirmatory factor analysis. Find out the correlation or covariance structure of all the latent variables. See here. Then once you have all your latent variables, everything is crystal clear, and then you go for structural model.

The structural model is the building between this path this one you will do and you will estimate and all those things. So, ultimately you will first prepare measurement model. Then from the measurement model, you get the covariance or correlation matrix for the latent variables. Using this correlation or covariance matrix, you will estimate the regression like coefficients, regression coefficients for this structural model. This is one.

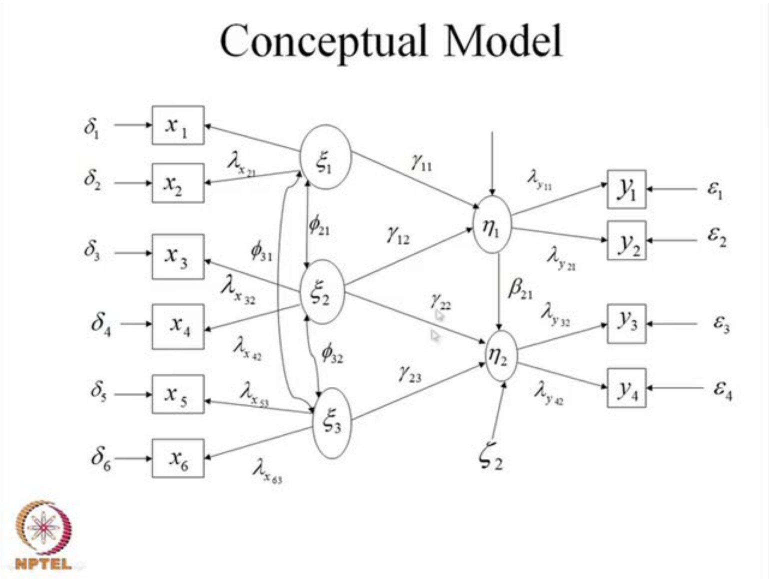
Second one is you may be interested to do everything at a time. That means you will feed the model on both structural and measurement part keeping intact at one go. But, because of so many parameters to be estimated and so many computations, permutations, combinations, what will happen ultimately? It is highly likely that you may get several of offending estimates. How to delete offending estimates it is an issue also in structural equation modelling.

(Refer Slide Time: 41:55)



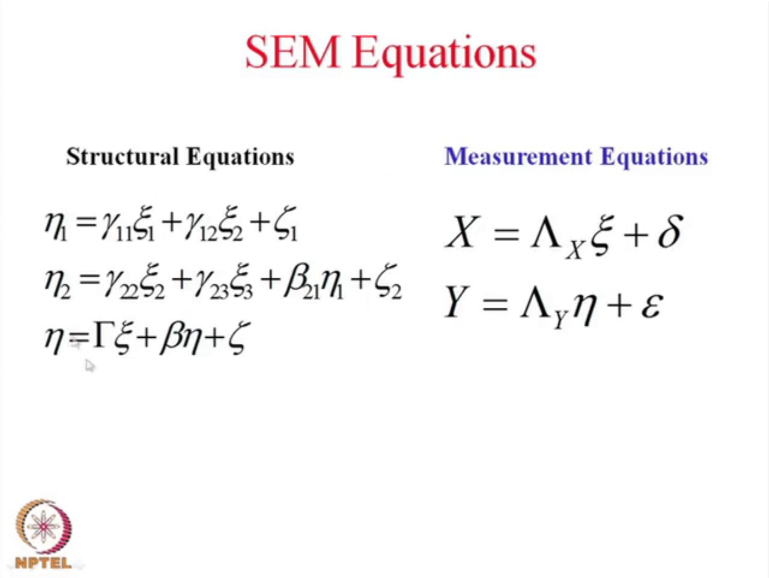
So, now what I will do?

(Refer Slide Time: 41:59)



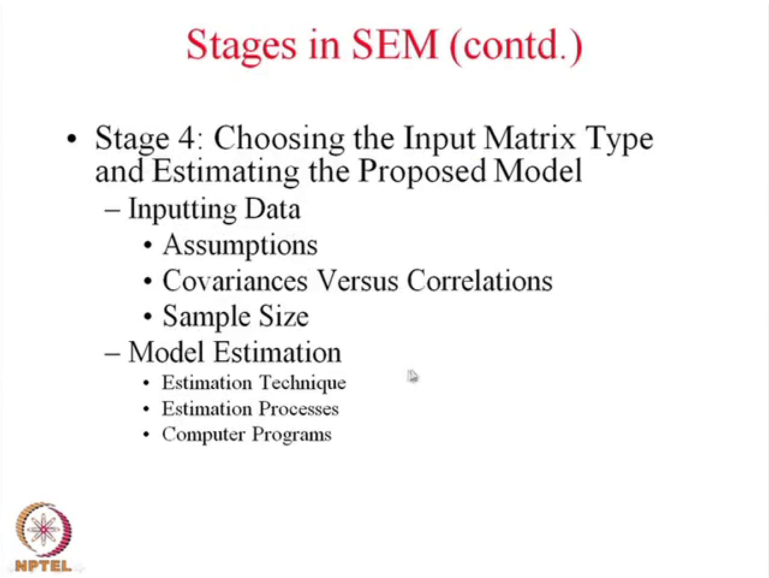
This is the structure I have shown you. Already, I have shown you. These are the stages that what we have discussed.

(Refer Slide Time: 42:05)



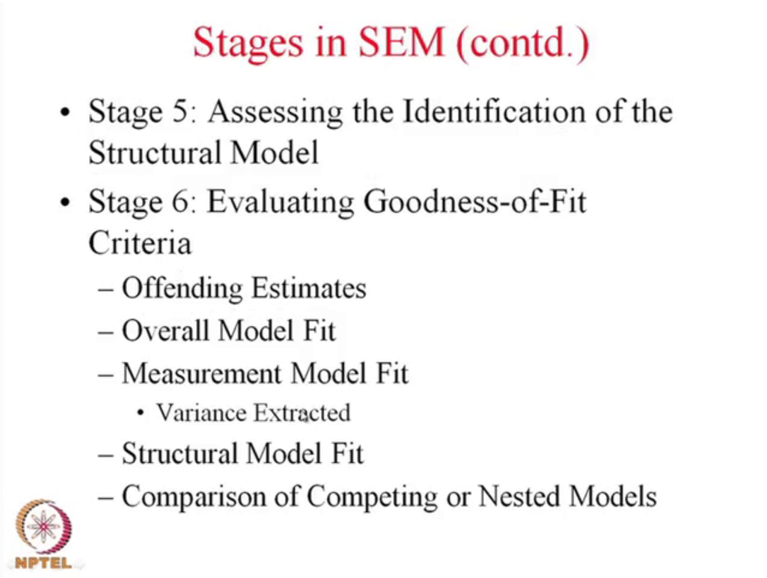
These equations also you have seen X and Y, these equations you have seen.

(Refer Slide Time: 42:10)



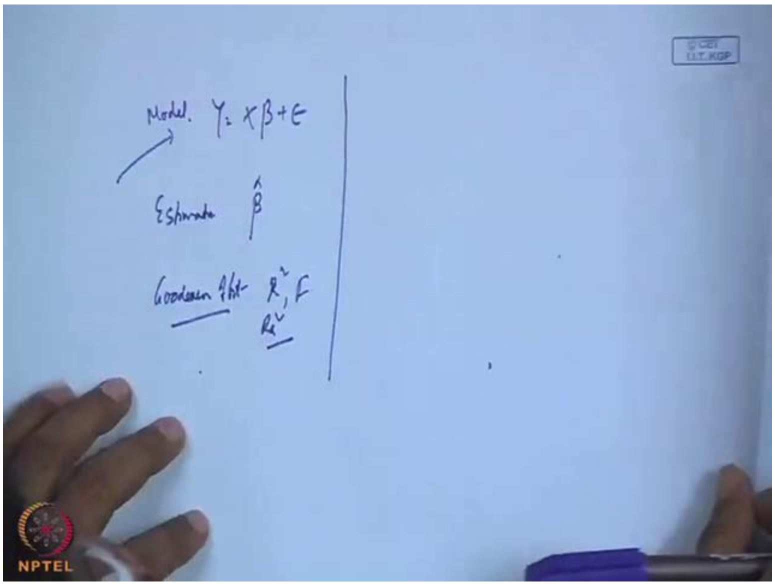
Then, the stage four is that what matrix you are going to use, you have to choose one input type of matrix that is correlation or covariance matrix. There are different assumptions. Then there is your estimation technique to be used, process, computing programs. These are all the stages one by one you have to do.

(Refer Slide Time: 42:30)



Then, suppose this is like your regression model also, you have seen that regression equation is y equal to x beta plus epsilon this is my model.

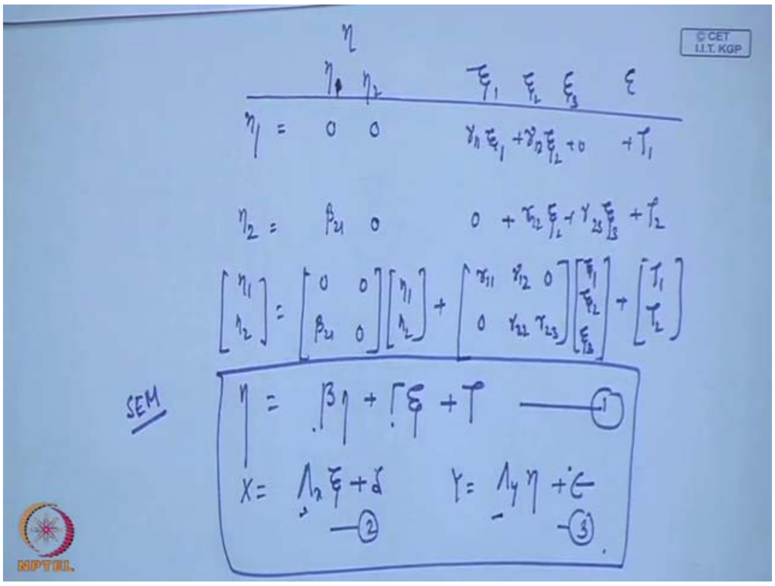
(Refer Slide Time: 42:41)



Then, suppose this is like your regression model also, you have seen that regression equation is Y equal to X beta plus epsilon. This is my model. Then what we have gone for estimation that is your beta cap. Then what you have done? You have gone for goodness of fit, model adequacy, goodness of fit or model adequacy that is R square F

Re square, all those things that we have used. So, here also, if I see that this is my structural equation modelling, then are you not getting a structural equation like this?

(Refer Slide Time: 43:17)

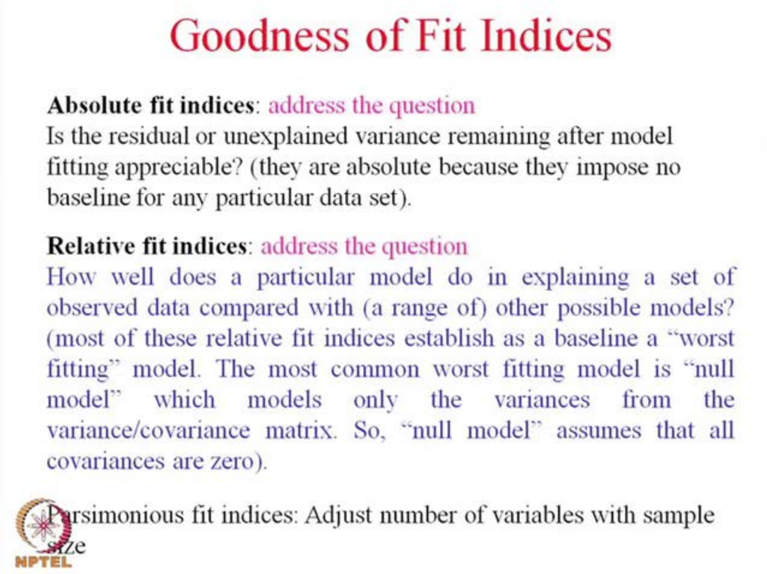


The equation will be your first one that eta 1 eta 2 in this case and definitely here will be eta 1 and psi, sorry eta and psi and error that path structure. Under this eta 1 and eta 2, if I write eta here, then this is psi 1, psi 2, psi 3 and all those things. So, you have to write the equations. If I write for this particular model the equations, eta 1 is not affected by any of the data 0, 0. So, gamma11 psi 1 plus gamma12 psi 0 plus, plus zeta 1 and this one is beta 21 and this is zero. Then this one is also 0 plus gamma22 psi2 plus gamma 23 psi3 zeta2.

Then, what is your matrix? It is eta 1 eta 2 equal to 0, 0 eta 21 0 eta 1 eta 2 plus, you are getting gamma 11 gamma 12 0, 0 , gamma 22 gamma 23 into psi1, psi2, psi3 plus zeta 1 zeta 2 that is my structural part. So, eta equal to beta eta plus gamma psi plus zeta, this is one equation. Another equation is X equal to lambda x psi plus delta and Y equal to lambda y eta plus epsilon. This is equation number two. This is equation number three. So, your model, this is my structural SEM model and you require to estimate beta gamma, then gamma x, lambda x, lambda y, covariance of this, covariance of this, covariance between these; so many parameters you have to estimate.

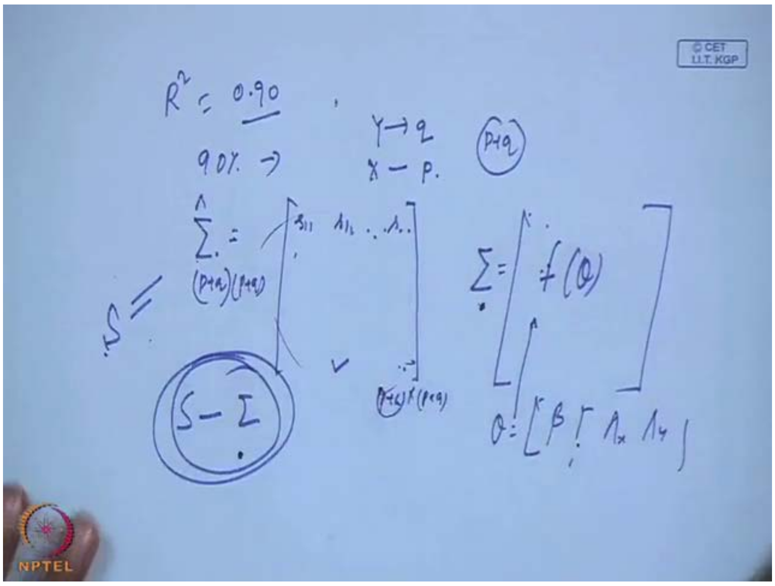
Second is parameter estimation. So, we are not discussing here the details of parameter estimation. We will discuss later on separately. So, once you have the model, you require to know that what will be the goodness of fit measures. So, my model is this and I know that somewhere, I will estimate all those things.

(Refer Slide Time: 46:39)



Then, I have to go for fit measures. There are three types of fit measures used in structural equation modelling. One is absolute fit indices, relative fit indices and parsimonious fit indices. In absolute fit indices, the question is the residual or unexplained variance remaining after model fitting appreciable. They are absolute because they impose no baseline for any particular data similar to R square.

(Refer Slide Time: 47:11)



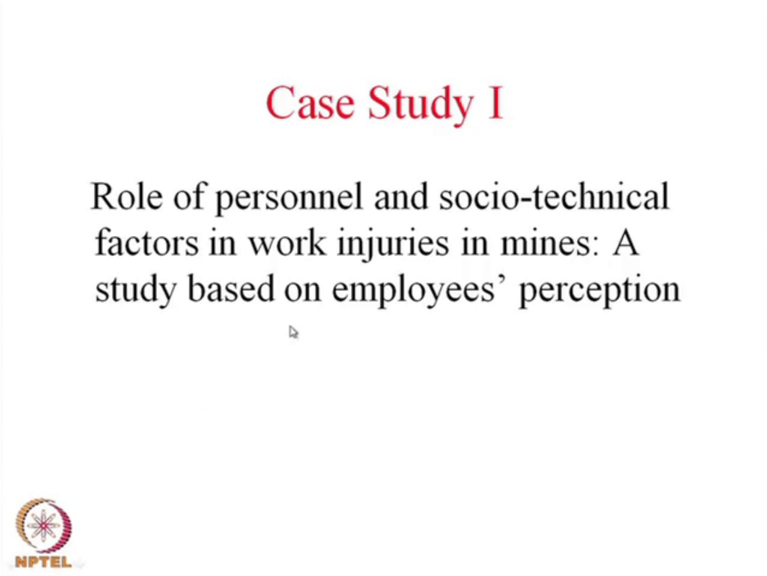
R square, if it is 0.90, it says 90 percent variance of Y is explained by this. Here also, we are basically talking about what will happen to this. Actually, although we will be describing in detail later on, the estimation what is observed with you is this. This is basically if I, how many Y variables, Y is q observed variables and X is p observed variables. You have total p plus q observed variables. So, you have a covariance matrix of p plus q or p plus q. So, this is the observed value and this is nothing but this is what is observed.

Now, using these equations, this eta X and Y, this equation, you will also be able to estimate or other way I can say fit this one, you can calculate this. Here, if I use S, then it is s11, s12 like s that p q up to p plus q. So, similar values we will be getting. Here, instead of, now I am writing instead of writing s here, we are writing the population sigma. Then that will be in terms of this, beta gamma lambda X lambda Y. So, there will be function of the, sorry, so what I mean to say is we can find out the theta in the nutshell.

I am talking about that beta, gamma lambda X, lambda Y, like this, it is possible to frame this, and then this s and these are compared. There are several ways to compare this. Then here p plus q into p plus q, this is the end. How many unique elements are there those will be compared. When we initialise some value here, some value here, those will compared.

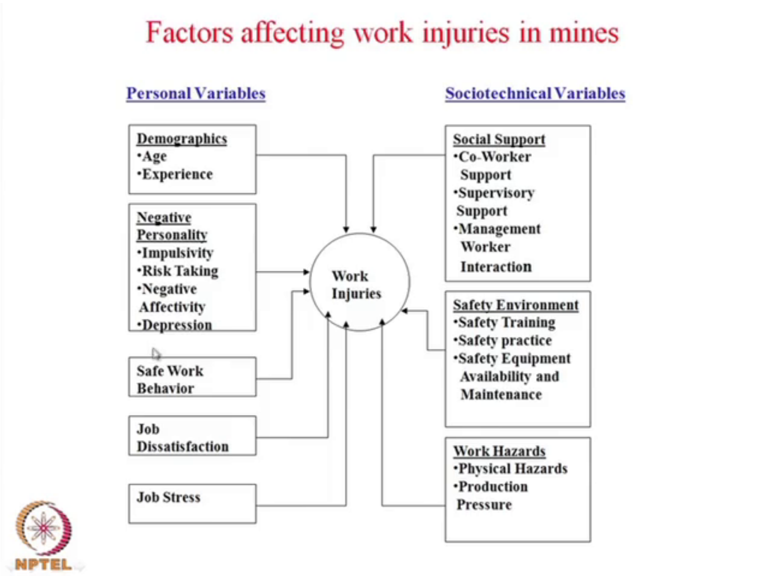
Then, we finally get the value. Then based on this difference, when we need something like this, this is total. Relative feed index that is the question that how well does a particular model do in explaining a setup of observed data compared with other possible models. You can go for several models and then compare which one is your best. The reference one is null model where we assume that there is no covariance value, relationships in the observed variables. Parsimonious fit index is similar to your adjusted R square where it is the number of parameters estimated, are adjusted. There are different fit indices.

(Refer Slide Time: 50:25)



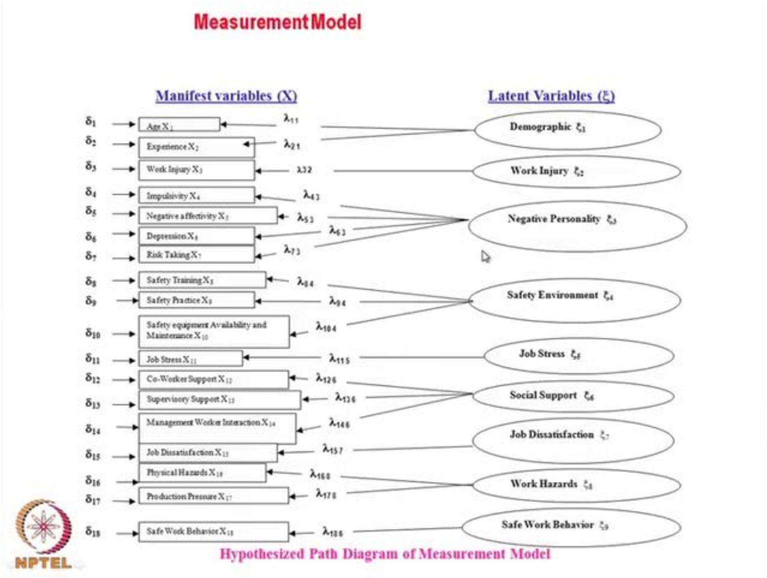
I will show you one case here by five minutes of time. This is role of personnel and social technical factors in work injuries in mines, a study based on employees' perception. This is published in organic general in 2008.

(Refer Slide Time: 50:41)



So, we have captured several variables under personal and social technical variables.

(Refer Slide Time: 50:52)



Then, what we have done? We have first gone for measurement model. All the manifest variables are used simultaneously and then the latent variables, first latent variables are identified and link to them was done. Then then confirmatory factor was done.

(Refer Slide Time: 51:18)

Correlation Matrix

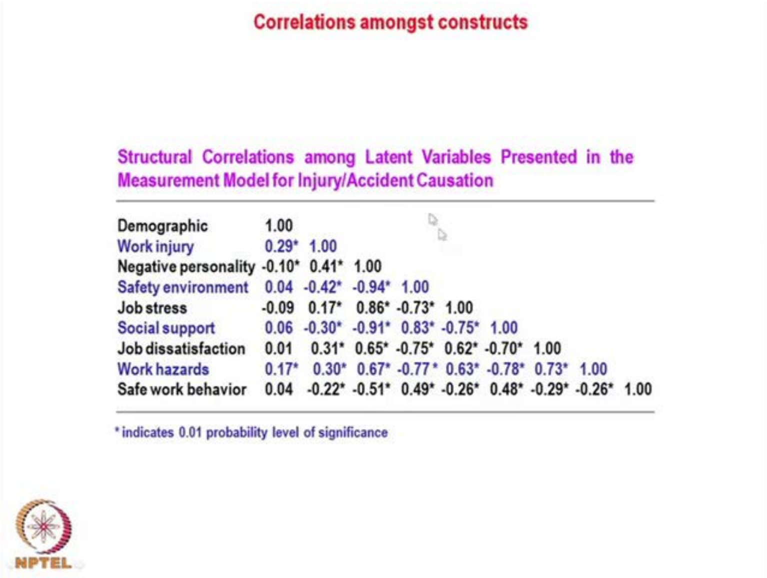
Correlations amongst the Major Variables for Accident Path Model

Age	1.000																	
Exp	0.816	1.000																
Work injury	0.271	0.249	1.000															
Impulsivity	-0.143	-0.085	0.214	1.000														
Negative affectivity	-0.114	-0.078	0.373	0.600	1.000													
Depression	-0.046	-0.059	0.084	0.183	0.409	1.000												
Risk taking	-0.001	0.033	0.349	0.467	0.586	0.224	1.000											
Safety training	0.074	0.060	-0.251	-0.577	-0.444	-0.112	-0.464	1.000										
Safety practice	0.052	0.040	-0.358	-0.607	-0.521	-0.251	-0.668	0.523	1.000									
SEAM*	-0.024	-0.056	-0.296	-0.350	-0.404	-0.277	-0.557	0.344	0.653	1.000								
Job stress	-0.103	-0.054	0.169	0.630	0.672	0.375	0.568	-0.543	-0.600	-0.460	1.000							
Co-worker support	0.118	0.087	-0.078	-0.456	-0.315	-0.084	-0.349	0.300	0.501	0.368	-0.375	1.000						
Supervisory support	0.042	-0.001	-0.269	-0.600	-0.547	-0.285	-0.660	0.631	0.716	0.560	-0.660	0.357	1.000					
M_W_INT**	0.053	0.012	-0.292	-0.588	-0.515	-0.295	-0.627	0.543	0.799	0.673	-0.659	0.451	0.817	1.000				
Job dissatisfaction	0.001	0.026	0.306	0.386	0.475	0.347	0.524	-0.436	-0.573	-0.625	0.617	-0.215	-0.653	-0.665	1.000			
Physical hazards	0.094	0.107	0.238	0.301	0.329	0.211	0.424	-0.200	-0.497	-0.514	0.455	-0.251	-0.465	-0.511	0.482	1.000		
Production pressure	0.113	0.124	0.169	0.119	0.240	0.413	0.424	-0.293	-0.362	-0.385	0.398	-0.099	-0.542	-0.466	0.507	0.459	1.000	
Safe work behaviour	0.055	0.010	-0.217	-0.368	-0.303	-0.362	-0.370	0.166	0.539	0.307	-0.258	0.361	0.396	0.387	-0.288	-0.172	-0.181	1.000
Mean	37.34	14.58	0.50	16.02	20.88	8.707	18.58	13.22	39.67	15.73	16.82	12.86	14.91	20.36	23.54	24.10	8.78	20.27
Standard deviation	9.01	9.25	0.50	4.12	5.95	2.55	5.41	3.27	7.53	3.93	3.99	2.18	3.68	5.44	6.16	4.36	2.82	2.99

* Safety Equipment availability and maintenance ** Management workers interaction
Correlation coefficient ≥ 0.113 indicates 0.05 probability level of significance

The output of the confirmatory model, this is the correlation matrix.

(Refer Slide Time: 51:22)



Correlations amongst constructs

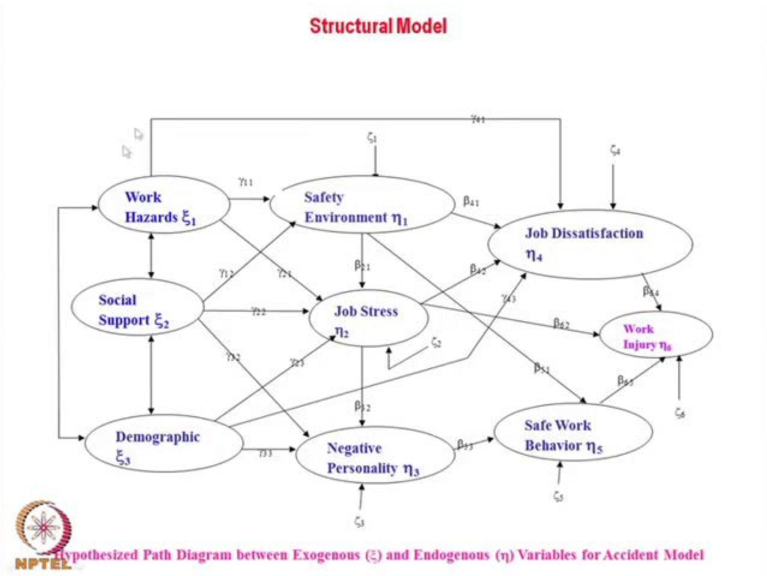
Structural Correlations among Latent Variables Presented in the Measurement Model for Injury/Accident Causation

Demographic	1.00								
Work injury	0.29*	1.00							
Negative personality	-0.10*	0.41*	1.00						
Safety environment	0.04	-0.42*	-0.94*	1.00					
Job stress	-0.09	0.17*	0.86*	-0.73*	1.00				
Social support	0.06	-0.30*	-0.91*	0.83*	-0.75*	1.00			
Job dissatisfaction	0.01	0.31*	0.65*	-0.75*	0.62*	-0.70*	1.00		
Work hazards	0.17*	0.30*	0.67*	-0.77*	0.63*	-0.78*	0.73*	1.00	
Safe work behavior	0.04	-0.22*	-0.51*	0.49*	-0.26*	0.48*	-0.29*	-0.26*	1.00

* indicates 0.01 probability level of significance

This is what the output is.

(Refer Slide Time: 51:29)

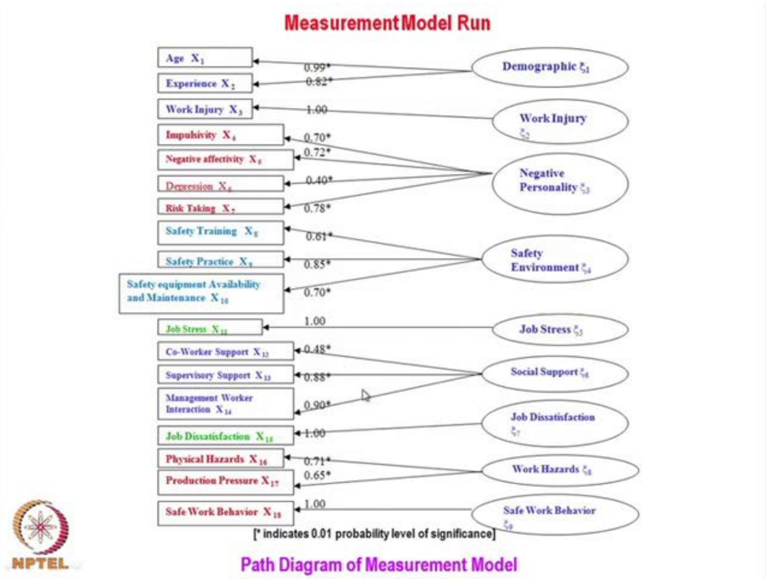


Hypothesized Path Diagram between Exogenous (ξ) and Endogenous (η) Variables for Accident Model

Once we get this output, this is input to our structural model. This is the first phase where you have to identify the manifest variable and the latent constructs. Then you do confirmatory factor analysis. Find out the correlation or covariance structure of the latent variables because in SEM, your input is either correlation covariance structure. So, for the measurement model, you give the covariance structure. They will give you the latent correlation or covariance structure.

Now, for structural model again, use the latent correlation or covariance structure, whatever you are getting from the latent variables and then you feed like this. So, you will be getting picture like this. Then definitely these are the some issues what we may be discussing later on if time permits in other classes. Ultimately, this is what is the observed correlation.

(Refer Slide Time: 52:29)



This is the factor and measurement model.

(Refer Slide Time: 52:31)

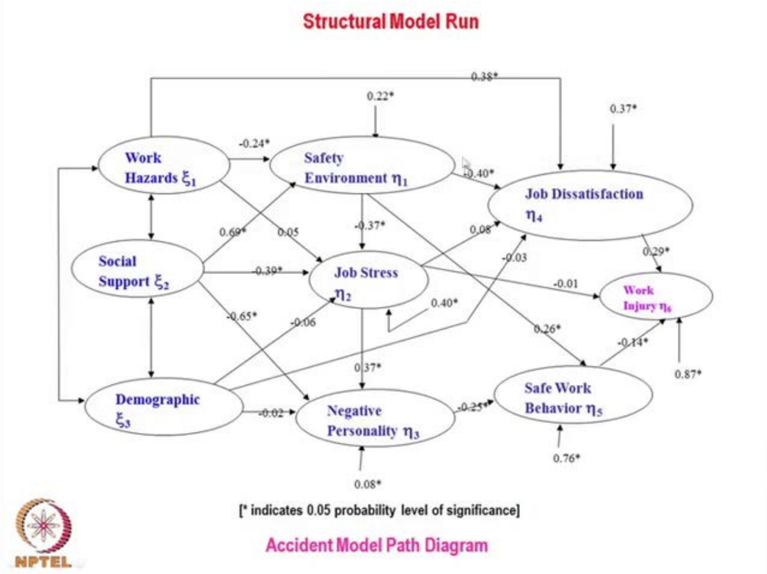
Measurement Model Fit Indices

Goodness-of-Fit Indices for the Measurement Model of Causal Accident Model

Parameter	Values
Chi-square with 99 degree of freedom	257.24
Root Mean Square Residual (RMR)	0.06
Goodness of Fit Index (GFI)	0.98
Normed Fit Index (NFI)	0.97
Comparative Fit Index (CFI)	0.99
Incremental Fit Index (CFI)	0.99

This is the fit of the measurement model that I told you that I will be discussing fully in the next class. Then this is the correlation among the constructs.

(Refer Slide Time: 52:41)



Then, this is your structural part and these are the structural regression coefficients.

(Refer Slide Time: 52:47)

Structural Model Fit Indices

Goodness-of-Fit Indices for the Structural Model of Accident Path Model

Parameter	Values
Chi-square with 15 degree of freedom	212.23
Root Mean Square Residual (RMR)	0.06
Goodness of Fit Index (GFI)	0.87
Normed Fit Index (NFI)	0.88
Comparative Fit Index (CFI)	0.88
Incremental Fit Index (CFI)	0.88

NPTEL

Then, you see whether the structural model is fit or not. If your structural model is fit, then go for interpreting the things. Here ultimately, there are several structural that fit measures like chi-square with 15 degrees of freedom. This should be as small as possible. Root mean square is 0.06. This should be less than 0.05, less than or equal to

0.05. The goodness of fit index is 0.87, 0.9 or more is better means adequate non fit index, but this is almost 0.9. We can say a structural model is fitting the data. Similarly, we have also seen that the measurement model also fits the data.

(Refer Slide Time: 53:33)

Contributions of constructs

Total Effect of the Significant Variables on Work Injury

Variables	Direct	Indirect	Total	Rank Order
Work hazards	----	0.15*	0.15*	3
Social support	----	-0.14*	-0.14*	4
Safety environment	----	-0.16*	-0.16*	2
Job dissatisfaction	0.29*	----	0.29*	1
Safe work behavior	-0.14*	----	-0.14*	4

* indicates 0.05 probability level of significance

Total Effect of the Significant Variables on Safe Work Behavior

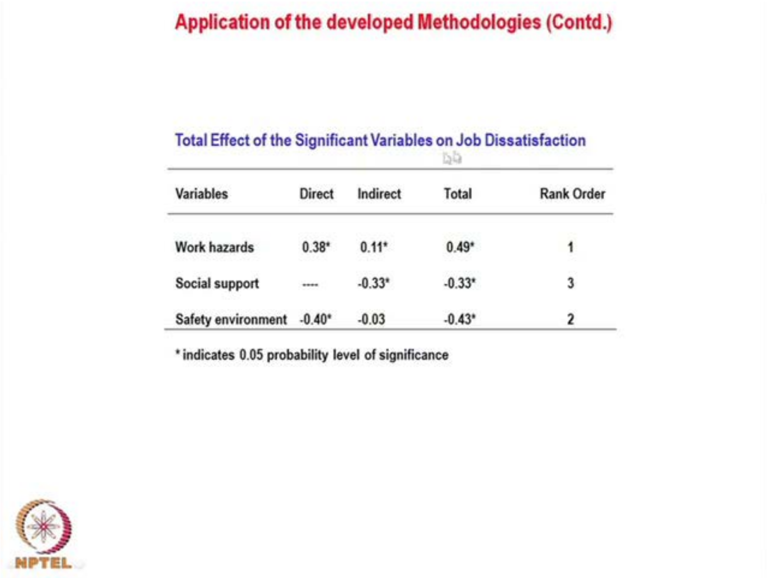
Variables	Direct	Indirect	Total	Rank Order
Work hazards	----	-0.07*	-0.07*	5
Social support	----	0.40*	0.40*	1
Safety environment	0.26*	0.03	0.29*	2
Job stress	----	-0.09*	-0.09*	4
Negative personality	-0.25*	----	-0.25*	3

* indicates 0.05 probability level of significance

So, that means our model or my model fits very much to the data. Then using this structural equation modelling, you will be also able to find out the total effect that is direct effect plus indirect effect. Also, you will be able to, based on the effect; you will be able to rank the variables. Now, what do we see that we found that work hazard vis a vis work injury, work hazard has no direct effects in the model, we have failed, but it is indirect effect. Total effect is this, social support this, indirect affect this, safety environment indirect effect this and total effect this, job dissatisfaction direct effect and safety behaviour, this is the case.

So, ultimately based on this total of effect, we have given the rank one, two, three, four, and five, like this. Similarly, as there are many variables, many dependent variables or dependent constructs, you can find out here. See these are the exogenous, but all are endogenous. So, for all any of the endogenous level variable, you can find out that what are the direct and indirect effects. This will definitely help you in making better policy.

(Refer Slide Time: 55:01)



Application of the developed Methodologies (Contd.)

Total Effect of the Significant Variables on Job Dissatisfaction

Variables	Direct	Indirect	Total	Rank Order
Work hazards	0.38*	0.11*	0.49*	1
Social support	----	-0.33*	-0.33*	3
Safety environment	-0.40*	-0.03	-0.43*	2

* indicates 0.05 probability level of significance

Total effect, so we calculated for this case, work behaviour, for work injury, for job dissatisfaction, what are the variables and what is the rank order. So, thank you very much. Next class, we will discuss confirmatory factor model with respect to structural equation modelling followed by my structural model, which is basically a path model.

Thank you very much.