We could consider different types of restrictions.

First, you can think about a survey where you asked the respondents whether they adopted a certain practice, and they responded to you by Yes or by No. If you try to explain the decision to adopt by some of the respondent characteristics, the dependent variable, here the adoption or the non adoption is the dependent variable, and is limited to two values (Yes or No). We will say we have a binary variable model.

Second, you may ask respondents a question where they have to choose among a limited number of exclusive options.

For example, you could ask whether they contracted some loans. In that case, there may be several solutions. First the respondent did not borrow money and we will code this as a zero. Second, she may have contracted a loan with a commercial bank and we will code it as 1. Third, she may have contracted a loan in the informal sector, and we will code it as 2. Finally, she may have contracted the loan with a bank and in the informal sector, and we will code it as 3. Again we can try to connect this decision to some of the respondent characteristics. The dependent variable will be limited to three values 0 ,1 , 2 ,or 3. We will say we have a multinomial variable model.

--- Slice 3

<speak>

<break time="1s"/>

Third, in a different type of survey, Shahzad Kouser and Matin Qaim, wanted to find out whether the adoption of Bt cotton had a significant influence on the frequency of acute pesticide poisoning related to sprays in the cotton crop.

<break time=".5s"/>

Since Bt is expected to reduce chemical insecticide use, they hypothesized a lower frequency of acute poisoning among Bt adopters. The dependent variable is the number of acute pesticide poisoning. It is a plain number, 0, 1, 2, 3, 4, etc.

<break time=".5s"/>

But we do not really have other restrictions on the number of poisinings. After all, some unlucky farmer may have experienced more than 6 poisonings. In that case, we will say that we have a count variable model.

</speak>

--- Slice 4

<speak>

<break time="1s"/><prosody rate="90%">

Finally, you may conduct a survey where you want to associate the quantity of fertilizers applied per hectare by the farmers to his farm characteristics. The dependent variable, here the quantity of fertilizer, is continuous, but it is limited on the left side, because you cannot observe a negative use of fertilizers.

<break time=".4s"/> Besides, you may observe a high number of farmers not using fertilizers at all, given an non-normal distribution fertilizer use.

<break time=".4s"/>

We can make the hypothesis that this is the result of an optimizing behavior leading to corner solutions for some nontrivial fraction of the population; that is, it is optimal to choose a zero quantity of fertilizers. We then need to model two decisions : do I want to apply fertilizer ? <break time=".2s"/> If yes, then what is the quantity that I want to apply ?

</prosody>

</speak>

--Slide Business as usual

<speak>

<break time=".2s"/><prosody rate="90%">

With a limited dependent variable, you could still use the standard Linear Regression.

After all, the dependent variable is a numeric value, and the regression will give you some results.

<break time=".5s"/>

However, you should be careful, since it is likely to give some biased results.

This is why other types of models have been developed to take care of the specific limitations in the data.

<break time=".4s"/>

It will always be a good idea to compare the specific model you have used with the standard linear regression. This will give you an idea of the extent of the bias avoided.

</prosody>

</speak>

--Slide Different types of LDV models

<speak>

<break time=".2s"/><prosody rate="90%">

There will be different types of models, depending on the type of limitation in the data.

When the dependent variable is a Binary variable, we will use a Logit or Probit model

<break time=".2s"/>

When the dependent variable is a Multinomial, we will use a Multinomial Logit or Probit model

<break time=".2s"/>

When the dependent variable is a Integer , we will use a Poisson model

<break time=".2s"/>

When the dependent variable is a Corner solutions , we will use a Tobit model

<break time=".3s"/>

Note that many other models exist. For example, when the sample did not permit to cover all the range of possible values for the dependent variables, we would use censored data models.

<break time=".3s"/>

In the following videos, we will mainly concentrate on the Logit and Probit models

</prosody>

</speak>

<speak>

<prosody rate="90%">

<break time=".1s"/>

In this course we will mainly talk about models where the explained variable is a binary variable.

Practically, this means that the explained variable is either a 0 or a 1

For example, we will use the data set used by Wooldrige, and study the participation in the labor force. We will try to relate this participation decision to the Age, Gender, and Education of the person.

A Person is either in the labor force or not. We will record this decision in a variable named Labor.

<break time=".1s"/> We could code Labor at either Yes or No, but we will use numeric values instead and enter one when the person works, and zero when he or she does not.

We then try to estimate the parameters of the following model :

</prosody>

</speak>

<speak>

<prosody rate="90%">

<break time=".1s"/>

Note that there are two binary variables in our model: Labor and Gender.

It is important to understand that the model is a binary dependent variable model because Labor is binary and not because Gender is binary.

For example, the second model is not a binary dependent variable model if the variable Income is a continuous variable

</prosody>

</speak>

<speak>

<prosody rate="90%">

<break time=".1s"/>

Let’s wrap up what we’ve learned

<break time=".1s"/>

First, limited dependent variable models are needed when the explained variable is restricted

<break time=".1s"/>

Second, Different types of models exist to avoid biased results that are specific to the type of restriction on the explained variable.

<break time=".1s"/>

In this course, we will concentrate on the binary dependent variable models

</prosody>

</speak>