Introduction to Economic Growth- I

Dr. Sohini Sahu

Department of Economic Sciences

Indian Institute of Technology Kanpur

Lecture-8

So, after discussing the various facets of GDP, now that we have taken a look at some data, we have also talked about how to calculate the basic growth rate, etc., comes a very pertinent question, especially when in research we are working with data. Sometimes it does happen that we cannot access the particular data that we are looking for. So, say for instance this looks a little unlikely, but it might be possible that GDP data is not available. So, what do we generally do when, let us say, GDP data is not available? Now, this might look a little, you know, unfamiliar as to why GDP data might not be available, but there might be a variety of reasons, okay? In some countries, let us say that it is undergoing war or there has been a huge catastrophe, let us say. So, during such times, you know data collection is perhaps not possible.

In some countries, maybe the data collection procedure itself might not be very sophisticated. So, during those times, there is a lot of misgivings about the data that comes out, right? And there are very specific instances that we will cite now, due to which sometimes economic activity data might not be available. So, what do we do under those circumstances? So, under those circumstances, that is in the absence of a variable of interest. Now, I would also like to clarify here that what we are going to discuss here is not just specific to GDP data or economic activity data.

This is true for any other macroeconomic variable. Well here we are talking about GDP because this is our variable of interest. So, in case GDP data is not available for our work or for some conclusion that we would like to draw, some examples we would be citing shortly. In that case, we use something known as a substitute variable, okay? And more popularly it is known as a proxy variable. Now, proxy, as the term suggests, I am sure all of us we are familiar with that say, for instance, like, in college, you know typically if someone is absent and then you know, that person requests the friend that you know maybe you know just give the attendance on my behalf right that is what we call proxy attendance, do not we? So, proxy means substitute.

So, here also proxy the term proxy has the same connotation. So, if particular data on a particular variable is not available, then we substitute it by a somewhat similar variable, which we call a proxy variable, and then one might ask that why is it that we might require

a proxy variable as we said that non-availability of data is the primary reason and data might not be available for a variety of reasons as we had just said it might be unreliable or maybe there is just lack of collection of a certain kind of data. Sometimes, there are also what we call missing points, okay? For instance, the monthly GDP data for India. We have quarterly GDP data for India that is published. So, I am showing a screenshot from FRED here.

So, if we are trying to find India's GDP data, monthly GDP, Let us say we are working on some project. Where we require monthly GDP data. What are we going to do? We are going to Google, right? We are trying to find the monthly GDP data. And as we had discussed earlier that we generally stick to official sources of data. So, FRED is a very popular source that we generally use.

So, that is why I am showing a screenshot from FRED here. And what do we see here? The GDP for India, you see there are two options that are here. Either it is annual, which we had earlier seen from RBI as well. Or there is also quarterly data. So, what do we do? Because we require monthly data for our work, then during that time, this is very popular in the existing literature; most researchers then use it instead of monthly GDP because that data is non-existent.

At least, you know, in the public domain, IIP instead is used. What is IIP? IIP is the Index of Industrial Production. So, we will call IIP as a proxy for monthly GDP data for India. So, this is one instance where we might require a proxy variable, okay? I hope we are good till now. So, this was one instance where we showed that maybe data is there, but it is not available very easily or it is not in the public domain.

Now, sometimes for some of our work, we also require data at a very granular level. What do you mean by granular level? It means at a very small level, small here in the sense of a particular example that I am citing here, small in the sense of administrative unit, okay? So, for instance, in India, we already have all India-level GDP, which is the aggregate GDP. We also have state GDP numbers and sometimes district GDP numbers are also available. So, you see how, in terms of an administrative unit, we started from the all-India level, which is at the topmost one can say. Then the next level is at the states, states are smaller than the entire economy per say and districts come within states.

So, this is what we mean by granular that is we are gradually moving towards smaller units. Now let us suppose for some work, we now require data on village-level economic activities or block level economic activities, or constituency-level economic activities, right? Now, that kind of data or that kind of GDP data for these smaller units, what we call the granular units, is generally not available in most cases. But as we said, you know, as researchers, when we are looking at interesting questions and economic activities, it is one of the most important variables, right, because it is the most important macro indicator. So,

what do we do because GDP data or economic activity data at such granular levels might not be available? One of the proxies for economic activities that we are going to present here. And let us again remind ourselves it is just one of the proxies.

We will also talk about this towards the end of this presentation, how there can be more proxies. But I chose this particular proxy because not only is it interesting, but there is a very wide application of this particular proxy variable for GDP or for GDP growth these days. As you can see on the screen, this is the night lights data. This is lights at night. This is for the entire world and this is a satellite image that we can see here.

And why are we using this as a proxy for economic activities? The reason being as we are aware of this that higher economic activities. They require more electricity. So, although one might argue that this is what we get to see only at night, it is still fairly true that well-lit regions do indicate something compared to regions that are not that well-lit. So, for instance, here in this particular figure that we see, there will be more figures later on as specific examples. As we can see that particular parts of the world are quite well-lit and also, in some cases we see that the intensity of the light is also vary.

So, India, as we can see, is fairly brightly lit in this particular figure. Now, where did this idea originate from? Now this idea has been around in the subject area of geography for quite some time, but in economics this is I would not say it started from here, there have been a couple of, you know, papers prior to this as well, but this is one paper that was published in 2012 and the authors are Henderson, Storeygard and Wiel. I would also like to quickly remind you here, if you recall, in the first week, when we mentioned the textbook that I generally follow, it is one of my favorite textbooks on economic growth; this is the same author. The textbook that we mentioned is by David Weil and we see David Weil's name here as well. So, these are the three authors and the paper is titled Measuring Economic Growth from Outer Space.

This is a paper that was published by AER, American Economic Review. In the year 2012, and I am presenting the abstract here, you are not required to read or go through the entire paper. Whatever is important from the paper for our particular purpose, I will be mentioning that here. But may be very quickly first couple of lines we can read out from the abstract so that we understand how they are presenting night lights as a proxy. So, it reads like this: we develop a statistical framework to use satellite data on night lights to augment official income growth measures.

For countries with poor national income accounts, the optimal estimate of growth is a composite with roughly equal weights on conventionally measured growth and growth predicted from lights, okay? You can read up on the rest, but the point is, as we mentioned, that there are several instances where official data might not be very reliable. There are various one can say situations where that kind of situation might arise. So, in that case,

instead of completely foregoing the GDP data, this is perhaps one possibility, okay? Now, this is what the night lights appear to be like. I will try and show as many graphics as possible so that we can drive home the point that we are talking about here. So, as we can see, this particular figure is taken from the paper.

In fact all the figures that I show here. There will be a couple of tables as well following the figures. They are all taken from this particular paper that I mentioned, okay? So, if you are interested, you can check out this paper, and you will find all the figures and tables in this particular paper. So this, for instance, is Eastern Europe in night lights. Now, in one panel, these are the lights of the same geographical region in the year 1992.

The other panel is of the night lights of the same region, Eastern Europe, at a gap of 10 years. So, this is 1992, and the other is 2002. So, I will give you a moment to just take a look at this figure and what we see here is that the lights are much brighter in the year 2002 and there are much more lights in the year 2002 compared to 1992. So, the researchers are trying to make this point that when we see something like this, it is clearly telling us a story, okay? So, the difference in the lights, not only in terms of their dispersion, but also in terms of their brightness. That is telling us a story here and this is what they have used to measure economic activities and also economic growth at the same time.

So, I hope up to now, we are all on the same page about how night lights are being used as a proxy for economic activities, okay? This is another example of Korea. We have both North Korea here and South Korea here, right? So this I will again give you a moment to look at the two figures and to try and interpret that what are these figures indicating to us. On one panel, we see both North and South Korea during 1992. And on the other panel, we see again the same two countries but this is for the year 2008. So there is a considerable time gap between the two figures that we see here.

Well what do we see here? There are a number of things; maybe if I miss out on something, please feel free to add. As far as North Korea is concerned, as we can see from the two panels here, there is not much difference between the lights. In fact, a lot of zones of North Korea are unlit, and that has not pretty much changed even in the year 2008; that is the first observation. When it comes to South Korea, we see that South Korea was fairly lit during the year 1992 when the first image was taken. And if you are comparing that with the year 2008, what do we see? We see that there are more lights now overall.

Do we all see that? What else do we see? We also see that the brightness or what is known as luminosity of these lights, they are much brighter in 2008 for South Korea. If we are more observant, what we will find is a lot more lights along the coastal zone of South Korea here, okay? These are at least the points that immediately came to my mind, as I said if we spend some more time. We might be able to draw more inferences. Now, look at the beauty

of the night light data. This is one of the reasons why I chose night lights data to be presented here as a proxy variable.

So far, see, we have not talked about any data per se. We are just looking at pictures. We are drawing inferences based on these pictures that we are seeing here. So, this is another instance where we see the earlier one was about Eastern Europe, but I think this one is even starker, right? So, this figure is also taken from the paper. Now, as we said that it is not only of interest because sometimes there is missing data or, you know, sometimes for some countries maybe official data is not coming on time, but there are also instances where maybe catastrophes where things like genocide in Rwanda you might be aware Rwanda had a very horrendous genocide episode.

Now, how would that show up right, or what would have been possibly the impact of that? That is also represented by lights here. As we can see, in 1993, at least, there was one bright spot, but in the year 1994, that spot almost disappeared. And then, within a gap of again 2 years by 1996, that light somewhat came back. So, these are the varied applications of night lights. We are not just talking about unreliable data or missing data.

but also if there have been catastrophes and natural calamities due to which data could not be collected as well, but you know, these are satellite images, so that is not affected by these things. At least we have some source of data. Now, naturally, one question that comes to mind is just like proxy attendance. Your friend is giving attendance on your behalf.

Now your friend is not you. So similarly, here, one might ask that we are not using the hardcore GDP numbers like the hardcore GDP numbers we had seen from the RBI website, remember? So we had very specific numbers there as GDP. Here, up to now, we have not seen any numbers. So, how do we know that how good is this particular proxy variable? Now, that is a very valid question. So, we just cannot take up any variable as a proxy variable because the point being it has to be able to or it should be able to capture the essence of the original variable that we are talking about. So, how do we make this comparison? So, this is what we see in the graph in this particular slide what we can see.

The authors have compared the actual log of GDP. So, to do these exercises we have to begin from somewhere where we already have the data, the actual GDP numbers and compare it with the night lights. So, this is what we have done. So, what we see here is in the solid line that is the log GDP that is predicted by the lights, and the dotted line or the dashed line that we see here that is the actual log GDP. So, obviously, there is some gap between the two; we will talk about some of the limitations because, remember, at the end of the day, it is a proxy variable.

So, it cannot capture the entire essence of GDP; if it could, then that itself would become the GDP. So, some gap is permissible; in fact, some gap will always be there, but as long as it does a fairly good job, we would kind of call it a good proxy variable. So this is the

first table in that particular paper, and the reason why we are looking at this table is when we are using night light data for any purpose because we are drawing some very strong inferences; we also should be careful about what we call some of the limitations. For instance, in night lights, now if we go back to the figure here or the figure here. Now, it is difficult to distinguish just by looking at these images whether light is not present because there is, you know, maybe a lack of development or light is not present simply because maybe it is a water body, right? So, that is why We are looking at this image here, you know, North Korea, South Korea surrounded by water bodies.

So, how do we distinguish between the two that is a very important question. Otherwise what would end up happening is we would possibly be looking at the image of a water body and there we see no night lights and we would infer that there is very little development here which is a totally fallacious you know conclusion that we are drawing. So, what is important here is we can see from this table that population density gives us a good clue about that, okay? So, if we look at one of the rows here, we see for these different countries. So, you see, this is a mix of different kinds of countries. We have Bangladesh on one hand, we also have US and Canada, we have Netherlands, we also have Madagascar, Mozambique.

So, this is like a mixed sample, okay? And what is this DN? DN is the light unit and it varies from 0 to 63. So, this is not like our regular GDP data, this is another thing. When we were talking about GDP data, it is generally in terms of US dollars, or in the Indian context, it is in terms of INR, Indian rupee. Now, Night Lights is not in terms of any of these currencies per se. It has a very different unit altogether because it is measuring the luminosity or the intensity of light, and that unit is d n, and it takes a value between 0 and 63, okay? So, we see the different ranges.

So, say, for instance, 0, which means the absence of light. And what does this indicate? So, for Bangladesh almost 67 percent of the area did not have lights that is the DN is equal to 0. But the same we can see is also true for Canada in fact much more than 67 percent almost 94 percent area of Canada where DN is 0. Now, this is what we would like to distinguish here that then how do we know that you know a good part of Canada is unlit? Is it because it is not developed, that is why it is unlit, or is it unlit because you know the topography? Because of the geographical location, you can say you know the distribution of land, etc., is such that it is sparsely populated. So, to distinguish between the two, that is why it is not just enough to look at the night light's number because otherwise, it will be very misleading. It is to be coupled with, or it has to be adjusted with the density of the population, right? And how is that visible here? Again, if we go back to the same data for Bangladesh, we see that the population density, which is per square kilometer terms, is 1080, okay? Whereas, the population density in Canada is only 3. So, this itself tells us a story that a lot of unlit portions in Canada that we have seen is not due to low development. So, these are some of the points that we have to be careful about.

Also, for our convenience the authors have presented. The actual per capita GDP numbers are both in PPP; remember we had talked about PPP earlier? There are also per capita GDP numbers in terms of constant US dollars so that we can compare the night lights with the actual GDP numbers, okay? This is another interesting table that the authors present, which is specifically meant for countries where data is available, or data is published, or data does come out from these countries, but maybe data is unreliable as they write here, and if we follow the explanation of the table that comes in the text if we look at some of these countries here. Now, the growth rate that has been predicted by WDI is for Myanmar. Let us say it is about 10 percent, okay? So, WDI is World Development Indicator published by the World Bank. But as we said that many a times there might be some doubts about the official data. And typically, this happens in low-income countries; you see LCU, which is why it is written here.

So, what if someone would like to verify that the growth rates that are presented are they the actual growth rates yes or no. So, to do this exercise what the authors did was for these bunch of countries, you can go through the list here. And Haiti is one country remember, we did mention Haiti also when we were talking about the stylized facts. So, for these countries they also predicted the growth rates based on night lights data. And as we said that obviously, there will be some gap; in some cases, the gap is larger than the other.

And the contention of the authors is that in the case for some country if the growth rate, the official growth rate, or if the growth rate that is announced based on GDP data, if there is any doubt about that, then ideally, what should be done is it should be combined with the growth rate that is predicted by the night lights data, okay? So, they say that this might be a somewhat better representation; one might get a better picture based on growth rates, which is based on night lights as well. So, this is the kind of application that night lights have in the current literature. Now this is also another representation of you know how different are growth rates based on night lights vis-a-vis the actual growth rate numbers that is based on the GDP numbers. So here on the x-axis, what is being measured is the annualized percentage change in GDP, which is based on data, the world indicator, and development indicator data. On the y-axis, the same thing is measured, but it is in terms of the night lights.

Now, for the same country, if the two numbers are the same, we would expect that particular coordinate to be on the 45-degree line, right? This bold black line that we see is the 45 degree line. And as we can see, of course, there are a bunch of countries that are either on the 45 degree are very close to the 45-degree line, which means that the numbers or the growth rate numbers coming from GDP, where GDP is calculated in a regular manner, is very similar to the GDP growth rate that one obtains from night lights. But there are also countries which are quite far from that 45 degree line. So, then that clearly indicates that the two numbers do not match, and in those cases, the authors suggest that in these

particular cases, the ideal growth rate should be a combination of the GDP numbers and the night lights numbers as well.