

International Relations Lab #3: Analyzing Strategies for Development – Least Developed Countries and the Millennium Development Goals

PS 3210 – International Relations
Prof. Vanessa A. Lefler
Middle Tennessee State University
Spring 2017

Introduction

Improving a country's level of development, as inferred from our textbook, is a difficult task. There are a number of different strategies that both individual countries and the international community may pursue in order to improve the economic welfare of those least developed countries in order to raise the overall standard of living, promote production and consumption, and resolve some human security dilemmas.

Given the number of strategies that are suggested and that are being (often concurrently) implemented, an important question for social scientists and policy makers is whether these policies have their intended effect. How can we tell whether a country is developing? Does a country's level of development improve more when we adopt specific development strategies?

The purpose of this final lab is to introduce you to a basic, but powerful, tool to assess policy effectiveness: the Student's t-test.

Learning Objectives

The learning objectives for this lab are:

- To further practice and comprehend the empirical study of international relations.
- To gain the ability to interpret basic data analysis.
- To apply results of quantitative analysis to substantive, policy, and theory questions.

Development and IR

As our textbook explains, there are number of strategies that countries and international communities are using to target lower levels of development at the state-level. These include (Kinsella, et al. 2012):

- *Liberalization*: the adoption of capitalist, market-based reforms to encourage foreign investment and trade.
- *Basic Needs*: an approach to develop combining local economic initiatives with provisions of assistance essential to the population, like subsidized food, health care, and education.
- *Self-Reliance*: the removal or severe limitation of economic and financial connections to core countries, MNCs, and banks based there, and even Western aid-giving agencies.
- *Regime Change*: the installation of either democracy, in order to take advantage of the complimentary features of liberal capitalism and democracy, or authoritarianism, in order to create social and economic stability.
- *New International Economic Order*: the creation of economic arrangements more favorable to the global South, including commodity price stabilization, development assistance, and debt relief.

Each of these development strategies has demonstrated benefits, but they have drawbacks as well. For example, the appeal of a self-reliance strategy is that it eliminates the influence of larger, economic powers, such as industrialized countries and multinational corporations. It also allows a country to cultivate its domestic industries without unscrupulous trade practices, such as dumping.¹ Yet, closing off an economy may also limit a country's potential for growth, depending on the country's resource endowment.

More problematic for evaluating the effectiveness of any development strategy, however, is that development is not a clearly linear path. There will be fits and starts, booms and busts; some areas of society and the economy may not progress at the same rates. In other words, there is likely to be a great deal of variation in our monthly, quarterly, and yearly assessments.

This means that we need to have a set of rigorous tools to help make accurate comparisons so that we can note where we observe real progress.

¹According to the World Bank, dumping is the practice of selling goods at a price significantly lower than it would charge in its own markets. For example, it is "dumping" when the NFL gives the Super Bowl's losing team's t-shirts to developing countries (Kenny, 2011). Often, dumping has the consequence of killing domestic industries that would produce the same goods because they cannot compete with the prices of the dumped products. This makes these countries even more dependent on foreign assistance.

A Basic Needs Strategy: The Millennium Development Goals

Accurately measuring development progress is particularly important for the United Nations as it approaches its deadlines for the Millennium Development goals.

The international community's plan of action to improve development, especially among least developed countries, is housed in the United Nation's Sustainable Development Goals. The Sustainable Development Goals (SDGs) are a set of markers the international community – principally the United Nations Development Programme – hopes to achieve by 2025. The expectation is that if significant strides are made toward reducing poverty and improving basic human rights, the global community and least developed countries especially will have greater access to economic and political stability, security, and freedom.

In particular, the UN has established specific targets within each of these goals, such as halving the number of people “whose income is less than one dollar a day” or “without sustainable access to safe drinking water” (“Road map towards the implementation of the United Nations Millennium Declaration” 2001).

For this project, we are not going to be concerned about whether the UN is on track to reach its specific targets. Rather, the question is whether the particular strategy that the UN adopts – a combination of basic needs and liberalization – has been the key to the progress that they have made.

Analyzing Development Data

There are a number of methods that we can apply in order to study whether a policy has its intended effect. At the minimum, we might simply compare the values of something that we care about before and after a policy was implemented. But, as discussed above, this might be a faulty method of determining whether a policy has any potency as there could be fits and starts in the process. Furthermore, some things that we are interested in developing through policy, such as education and literacy, tend to be increasing regardless of our intervention through policy. Thus, determining whether the changes we observe are the result of the policy or what we simply find through natural processes can also be difficult. Thus, we need to adopt some more sophisticated tools.

The Student's T-Test

When we want to compare two sets of information, such as the period before and after a policy change or whether a treatment works for those who received it rather than a placebo, rather than relying on a visual or descriptive guess about the substance of the difference between them, we can employ a simple, statistical analysis to verify it. This statistical analysis is known as the Difference of Means test or, more commonly, Student's *t*-Test.

Student's t -Test begins with same comparison that we make in the most naive analysis: the difference of the averages between the two groups.

Let's say, for example, you were given grades for two different exam given in one class. You might, then, have a dataset that looks something like this:

Exam	Score (%)
1	40
1	70
1	70
1	80
1	90
2	35
2	55
2	80
2	95
2	70

Comparing Averages of Two Groups

There are two exams in this dataset, each with five scores. With simple calculation, it is apparent that the second exam was more challenging than the first, on average. The average score on the first exam was 70% while the average score on the second exam was 67%. An examiner who wants to give two relatively equivalent exams might be worried based on this information that the second exam is not a good comparison with the first and, therefore, she needs to go back to the drafting board. But, a closer looks shows that there is a great deal of variation in the scores on the two exams. The lowest score on the first exam was not as low as the lowest on the second exam, but the highest score on the first exam was not as high as that on the second.

Student's t -Test provides a way for us to take into account this variation that makes our certainty about the validity of simple comparisons difficult.²

Considering the Spread of Observations

Whenever we have multiple observations for an event or outcome we want to study – again, for example, a series of years before and after a policy was implemented, a number of patients who were given a treatment, a class of students taking an exam – those multiple observations constitute

²There are actually a variety of t -test methods that can be used. For simplicity, this assignment will only address the two-tailed, heteroskedastic t -test. For more information on Student's t -test, check out: Tacq, Jacques. 2004. "T-Test." *The SAGE Encyclopedia of Social Science Research Methods*, Michael S. Lewis-Beck, Alan Bryman, and Tim Futing Liao, editors. Thousand Oaks, CA: Sage Publications, Inc. pgs. 11146-1148.

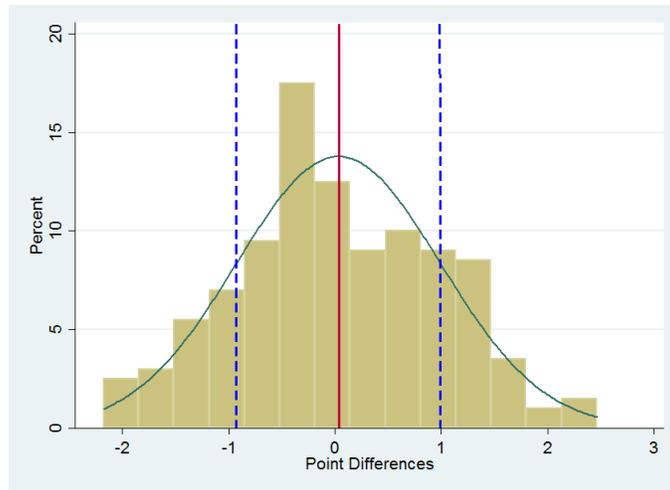


Figure 1: An Example Distribution of Score Differences Between Two Tests

a *set* of data. That set of data has a distribution, with some of the observations at the lowest end, some at the highest end, and several more in the middle. For example, consider Figure 1: Figure 1 is a hypothetical distribution of differences in points between one test and another. As you can see, some differences were negative, some were positive, and some had no difference at all. In this set of data is the average, indicated by the solid red line. But, with this average, there is also some variation; some had greater differences and others had lesser differences.

A useful method of managing these differences might be to simply calculate the range, the greatest value minus the smallest value. But, if someone wants to perform a test, we might not want to rely solely on the best performers and the worst performers to gauge our examination or treatment methods. Instead, our best approach is to also consider the way that the observations are distributed in the middle of the data and ask, in what ways to do our observations *typically* differ from the average?

To answer this question, we rely on the statistical measure of *standard deviation*. Standard deviation can be simply explained as a measure of how spread out the values in our observations are.³ It is calculated by taking the square root of the variance of the observations. When we look at a distribution of data and want to then know the range of values that we can most typically expect to observe if we continue to deliver this same treatment or exam, we look at the observations between the average minus the standard deviation and the average plus the standard deviation. These values are denoted in the dashed blue lines on Figure 3.

In more advanced statistics training, we can go into the mathematics behind both of these concepts, but, for now, it is easy to calculate the standard deviation in a set of data using the functions in most spreadsheet programs. The important thing to remember for Student's *t*-test, is that we calculate this value for both sets of observations. Thus, for our example exam data, the standard deviations are 18.7% for Exam 1 and 23.1% for Exam 2. Based on this, we know that there is quite a bit more

³<http://www.mathsisfun.com/data/standard-deviation.html>

spread in Exam 2 than in Exam 1, but, the ranges of expected values for both of these exams is quite large.

This last observation – that the ranges of expected values for both exams is quite large – is important for our last step. In our exam data, the normal range of values for Exam 1 are between 51.3% and 88.7%, while Exam 2 ranges between 43.9% and 90.1%. In other words, the ranges of expected values between the two exams have a substantial amount of overlap. In the graphical method of data analysis, we observed that overlaps are possible indications that the difference between our two sets of information are possibly the result of random factors, rather than the treatment or policy implementation. This is generally true for comparing the overlapping ranges of values according to their average and standard deviation. But, this type of comparison is quite sensitive to the number of observations that we have in each group. If a few more people took Exam 1 than Exam 2, we might not be confident that our comparisons using this approach are as accurate as they could be because the number of observations in the set affect how much variation is in the data.

Separating the Difference from the Noise

Student's t -test solves this last problem as it assesses whether the two groups that we are interested in studying are, with a certain degree of confidence, statistically different from one another.

Student's t -test simply divides the difference of the averages between the two groups by how much variation is in both groups.⁴ This allows us to separate some of the “noise” that is caused by variation in both groups so that we can determine whether the difference is statistically significant. Any difference that is statistically significant is one that we can say with a certain degree of confidence – usually at least 95% – that it is different from no difference at all, or a difference between the two averages equal to zero.

As with the standard deviation, there are ways to manually calculate Student's t -test. There are also easily accessible programs, particularly spreadsheet programs, that have built-in methods of calculating the t -test statistic. These methods make it fairly easy, given that you have:

- The observations for your two groups (e.g., before/after, treatment 1/treatment 2).
 - The observations for your two groups are necessary to calculate the *averages* of your groups.
 - This is, of course, necessary for us to tell, in general, what differences there are between them.
- Whether the range of possible differences in your data are simply either positive or negative, or whether they may be both.

⁴http://www.socialresearchmethods.net/kb/stat_t.php

- Consider the exam example again, it is probably reasonable to assume that some students will do better on the second exam, while others might do worse. Do your data follow the same pattern?
- When you can reasonably assume that the direction of change can turn positive or negative, then you want to conduct what is called a *two-tailed* test.
- Consider Figure 1 again. The tails are ends of the green, curved line that move from the center of the graph to its edges. In this case, there are two of them.
- In other cases, where the change might be assumed to only go in one direction, then you may be able to conduct a *one-tailed* test.
- A knowledge of whether the spread, or variation, of the values of your observations are equal or different.
 - This, of course, is based off of the standard deviation. The wider the standard deviation, the more potential “noise” or randomness there might be in our data.
 - For different sets of observations, this value is likely to be different. The only way to know is to calculate the standard deviation.
 - When the standard deviations of your two groups are different, you want to conduct a *heteroskedastic t-test*.
 - Heteroskedastic simply means that the variation in the two groups has different characteristics. This is important because if we assume that the variation characteristics are the same, when in fact they are different, we might mistakenly predict that a change in the first group will affect the second group in the same way.

Interpreting the t-Test Statistic

When you use a software package, like a spreadsheet program, to calculate Student’s *t*-test, what you are given is the *t*-test statistic. This can be interpreted as a measure of confidence that we have that the two values are statistically different from each other.

For example, using the spreadsheet’s *t*-test function, with a two-tailed test and heteroskedastic variances, produces a *t*-test statistic of 0.827.

What does this mean?

The most direct way of interpreting this number is to say that we are 82.7% confident that the difference between the two exams is statistically equal to zero. But, we typically do not interpret this value quite in this way. This is because we often care more about whether the difference is statistically *different* from zero, not whether it is statistically equal to it. So, rather than directly interpreting this value, we subtract it from one (1) in order to convey our confidence that they difference between the two groups is meaningful.

When we subtract the t -test statistic from one, we get the new value, 0.173. We can now interpret this to mean that we are only 17.3% confident that the difference between the two groups is statistically significant, or different from zero.

But, what does this mean?

Well, it means that we are not very confident that the second exam was, in fact, much more difficult than the first.

Typically, we want to be at least 95% confident that the difference between the two groups is different from zero. So, if we wanted to conclude that the second exam was demonstrably more difficult than the first, we would have to find a t -test statistic between 0.05 and 0.

In sum, Student's t -test allows us to compare two sets of observations, while taking into account the random factors that can also affect our data collection. It also tells us with what degree of confidence we can say whether the two groups are distinct, based on the values we are using to separate them, such as a policy implementation, treatment, or examination style.

Lab # 3 Assignment: Evaluate Development Strategies

Directions

General Objective: Using Student's *t*-test, evaluate four development hypotheses applied to Least Developed Countries (LDCs).

This exercise will be completed in a series of steps.

Select Data

1. Select two, general development strategies from those we've discussed in class: foreign investment and liberalization, self-reliance, basic needs, democratization, authoritarianism, or cartelization.
2. For each development strategy, identify, from the following list, a variable that measures a method of implementing the strategy. Briefly explain why the variable is an appropriate measure of the development strategy.
 - *ODA Rank* – a two-category ranking of the amount of Official Development Assistance the country received between the years 2000 and 2005.
 - Countries coded 1 received greater than average (among LDCs) amounts of ODA, 0 indicates that the country received less than average amounts of ODA.
 - *Democracy* – a two-category description of the country's political regime for the year 2005.
 - Countries coded 1 were considered democracies by the Polity project in the year 2005. A code of 0 indicates that the country was either an authoritarian or mixed regime.
 - *Authoritarian* – a two-category description of the country's political regime for the year 2005.
 - Countries coded 1 were considered authoritarian regimes by the Polity project in the year 2005. A code of 0 indicates that the country was either a democracy or mixed regime.
 - *Trade Openness* – a two-category description of the country's amount of trade openness between the years 2000 and 2004.
 - Countries coded 1 had greater than average (among LDCs) trade openness scores from Gleditsch (2004). Countries coded 0 were less open to trade than average.
 - *FDI Received* – a two-category description of the country's amount of Foreign Direct Investment.

- Countries coded 1 received greater than average (among LDCs) FDI, based on US dollar value, between 2000 and 2005 according to the International Monetary Fund. Countries coded 0 received less than average FDI.
3. Select two indicators of a country's level of development from the following list, based on the indicators given by the Millennium Development Goals from the year 2012:
- *Poverty Headcount Ratio* – The percentage of the population living on less than \$1.25 a day at 2005 international prices.
 - *Secondary Education Enrollment* – The percentage of the population enrolled in secondary education (e.g., high school), regardless of age.
 - *Female-to-Male Education Ratio* – Ratio of girls to boys enrolled in primary and secondary education in both public and private schools.
 - *Child Mortality* – The probability out of 1,000 that a newborn baby will die before reaching age five.
 - *Maternal Mortality* – The number of women who die during pregnancy and childbirth, per 100,000 live births.
 - *HIV/AIDS* – The percentage of the population aged 15-49 who are infected with HIV.
 - *CO2 Emissions* – Kilotons emitted from the consumption of solid, liquid, and gas fuels and gas flaring.
 - *Government Debt* – Total debt service, measured as the sum of principal payments and interest actually paid to the IMF as a percent of net government revenue.
 - *Internet Users* – The number of people, out of every 100, that have access to the world wide web.

State Hypotheses

4. Your development strategies and indicators make four, pairwise combinations: e.g., Strategy 1/Indicator 1, Strategy 2/Indicator 1.
- (a) Present appropriate theories that link your selected development strategies with your development indicators. Briefly explain the general logic and then pose the specific argument as a hypothesis.
- Make sure you pay attention to the direction of the effect that you expect the development strategy to have on the development indicator.

For example:

“First, democratic governments must be responsive to the basic needs of their populations in order to maintain political stability and economic

growth. These needs include maintaining a healthier population to sustain the growth of the labor pool. Therefore, I hypothesize that democracies will have *lower* maternal mortality rates than non-democracies. (50 words)”

“Second, democratic governments harmonize with liberal economic reforms that promote development by increasing access to industry and other means of production, rather than centralizing control of these resources in a few hands. This means that developing democracies may have larger numbers of people consuming fossil fuels – signaling growth, albeit environmentally unsustainable. Therefore, I hypothesize that democracies will have *higher* carbon emissions levels than non-democracies. (65 words)”

- (b) Repeat these steps for each of the four strategy/indicator combinations.

Hypothesis Testing

5. Using the data on `students.vanessalefler.com/developmentlab`, collect the results of the t-tests that have already been calculated for your combinations of development strategies and indicators.
 - (a) First, select a development strategy variable from the drop-down list, which corresponds with one of the variables listed above in 2.
 - (b) Next, select a development indicator or outcome from the second drop-down list, which corresponds with one of the variables listed in 3. Then click “Test.”

The page will report a difference of means analysis of the differences in between the two groups in your analysis by the development indicator you selected.

6. Report *all* the results provided in the table and interpret their meaning.
 - (a) Describe each part of the table.
 - (b) Pay attention to your object of analysis and the direction of the difference between your groups, looking not only at the calculated difference, but also the actual averages.

For example, the average maternal mortality rate among LDCs that were democracies was 366.647 deaths per 100,000 births. For non-democratic LDCs, this was 467.241. This means that the calculated difference of means is 100.594, indicating that more than 100 fewer women die during childbirth in democracies than in non-democracies.

In contrast, interpretation requires a bit more thought for democracy as a method to improve a country’s level of education. Secondary education enrollment in democratic LDCs is 49.975% and 37.87% in non-democratic LDCs. The results indicate that this is a difference of -12.106. The negative sign on this result, though, does not indicate

Difference of Means Analysis of Maternal Mortality by Democracy		
	Democracy	
	0	1
# of Cases	29	17
Mean	467.241	366.647
Std. Dev.	257.445	233.391
Difference of Means	100.594	
T-Test	0.1829	

Figure 2: Difference of Means Analysis of Maternal Mortality by Democracy. Example of a Positive Difference of Means.

Difference of Means Analysis of Secondary Education Enrollment by Democracy		
	Democracy	
	0	1
# of Cases	24	18
Mean	37.87	49.975
Std. Dev.	19.403	23.389
Difference of Means	-12.106	
T-Test	0.0838	

Figure 3: Difference of Means Analysis Secondary Education Enrollment by Democracy. Example of a Negative Difference of Means.

that democracies fare more poorly. Rather, the secondary education enrollment rate in democratic LDCs is more than 12% higher than in non-democratic LDCs.

7. Identify whether the difference between LDCs that received the development-promoting policy were *significantly* more developed than those that did not receive the policy.
 - (a) Interpret the difference of means. Is the difference in the direction you hypothesized?
 - (b) Interpret the *t*-test. With what degree of confidence can you say that the difference is due to the policy and not to other, random factors?
8. From these results, would you recommend that this policy be adopted to improve development among other struggling economies?
9. Repeat these steps for all four of the pairwise combinations of development strategies and indicators.

Formatting

The format of your paper should conform with the following rules:

- 700-1200 words in length.
- 1-inch margins, all around. Standard, serif (e.g., Times) or non-serif (e.g., Calibri, Arial) font, between 10 and 12 point size.
- The heading, which includes the “top matter” and title, should be single spaced, with double-spacing between the title and the body of the paper. In the header, each page should give your last name followed by the page number.
- If you use any sources outside the textbook or the article, cite them at the end of the paper.
- Save the document as a **.pdf** and deposit it in the Dropbox on D2L.

Papers not submitted in .pdf format will be penalized 10%..