Foreign Direct Investment and Development in Africa

Members: Nikolai Sie, Ankita Singh, Tanay Choudhary

December 15, 2017

1 Introduction

We have measured the effect of Foreign Direct Investment (FDI) on development in Africa, using regression analysis. With a landmass greater than the United States, India, China, and Europe combined and a population of almost 1.3 billion, development in Africa matters greatly to the world's welfare. Its young population and abundance of land and natural resources mean Africa is well positioned to harness its endowments for its people's benefit. Its generally weak governance and low level of technological development gives it the opportunity for great catch-up growth. But growth needs to be financed, and technology needs to come from somewhere. Africa is relatively poor in both capital and knowledge, so an alternative source of financing and technology is FDI. Our hypothesis is that FDI causes largely positive development effects in Africa.

Previous researchers has used aggregate FDI data and aggregate measures for development. By breaking up our FDI data by industry, we could find industry-specific relationships that were hidden from previous researchers. By evaluating multiple development indicators, we find a more nuanced perspective on FDI's effect on development. They also use FDI data based on Balance of Payments accounting, which means complex accounting rules distort the values in the data. Our FDI data is econometric estimates and broken down by individual FDI projects, so it is not distorted by accounting rules. It also includes jobs created, as well as dollars invested, for each project.

We use regressions, the bread and butter of economists, for our analysis. With data covering every African country in the period 2003-2016, we include control variables to adjust for outside effects, and our regressions can tell us how FDI in any given year affects some development indicators in the same and following years.

2 Data

FDI Intelligence has collected and estimated data on the value of and jobs created by every FDI project in the world since 2003. The projects are categorized by industry (e.g.: agriculture) and activity (e.g.: manufacturing). We have collected all the data on FDI into Africa.

The World Bank's World Development Indicators dataset has many development indicators for all countries, and we have selected 137 of them for all African countries. Most will be regressed against their relevant FDI statistics, some, especially those measuring governance quality or absorptive capacity, will be used to measure interaction effects, and some will be control variables.

We have processed this data into panel format. It is organized in a matrix such that each row contains all of the FDI data and all of the development indicators for one country in one year. The FDI data is reduced from three to two dimensions in order to make the regression analysis easier. Thus, each row has data on FDI in each industry and all combinations of industries and activities.

3 Summary statistics

One needs to explore the data. This phase of our project was the data exploration and exploratory data analysis parts. Firstly the data was loaded into a jupyter notebook. We started with a combined dataset called "Combined.csv" which was comprised of variables of type: float64, int64 and object. The size of this data file was 3.4 MB and was compromised of 756 rows X 666 columns. The combined dataset compromised of FDI investment data in every industry category and the responses in created in a country. For summary statistics we needed just the investment part.
The dataset was loaded into a dataframe and was sliced to the column where investments in industries end. The column "allall" was replaced by the name "Total". "allall" held the total investments made in a million dollars in the given countries for every year in the range 2004-2016. Some summary stats we derived after we explored our data were:

- Africa has 54 countries recognized by the UN as of 2017.
- The Total FDI in Africa over the 10 year range is a staggering 1,657,388 Million USD or 1.66 Trillion USD.
- The highest FDI over a 10 year period was in Morocco 222.97 Billion USD while the least non-zero value was in Comoros 0.324 Billion.

The Gravity Model: This economic model states that organizations and countries usually invest

in countries that are close to them as its a shorter distance, similar cultural and similar language spoken. This can be seen in the case of FDI as the 3 highest investment statistics are over in Morocco: 222.97, Egypt: 215.64, Tunisia: 89.08 and Algeria: 96.28 (All figures are in Billions, USD). These countries are on the coast of the Mediterranean sea and are close to Europe allowing quick movement of goods and labor. The languages spoken in the countries apart from the local languages are French, Arabic and English. This allows investors to easily communicate with their counterparts in these countries.

With comparing the total FDI over a 10-year period another group of countries emerge that have high FDIs. Theses are countries on the coast. South Africa: 206.24, Ethiopia: 94.37, Mozambique: 54.63,Kenya: 51.98, Angola: 38.89. These countries seem to have the next highest Investment figures. We can infer that investing organizations prefer these due to the sea connectivity to transport goods.

Another interesting inference is that countries closer to South Africa: 206.24 have just a fraction of the figures invested in South Africa. Example: Botswana: 21.58 , Zimbabwe: 25.95, Namibia: 23.46.

The sub-Saharan region of Africa has very small investment figures. One if the major reasons investors avoid this region is because of weak governance and lack of infrastructure. Adding to that is the threat of extremist organizations such as the Boko Haram. A few of these countries are:Senegal: 19.34, Guinea: 14.16 ,Chad: 5.82, ,Niger: 3.53, Gabon 5.42. Although Countries like Senegal, Guinea and Gabon are on the coast they aren't immune to extremist organizations. An outlier in this group is Nigeria: 109.05 where Boko Haram is present but this hasn't stopped investors from seeking opportunities.

Another Region affected is that of east Central Africa where issues such as lack of connectivity and unstable governments have discouraged investors. An example is the Civil War in Sudan: 15.37 and South Sudan: 4.42. While nations such as Uganda: 21.99, Rwanda: 10.96 Burundi: 2.62 and Central African Republic: .67 have very low figures. Highest in this region being the Congo (DRC) 36.34, a small figure compared to the northern and coastal regions of the continent.

The summary statistics code also offers it users to Enter the name of a country and get information in 2 different ways:

1) A plot that holds the total FDI over a 10 year period and has the options to add up to 3 industries to it.
2) A pie chart that shows total investment in every single year.

(All figures are in USD and Billions)

4 Economic background and literature review

FDI is when a Multi-National Company (MNC) invests in a foreign country, the host country, to begin operations there. This can be done by purchasing more than 10% of another company in the host country, or by buying or constructing property or equipment to begin operations there. It is generally thought that FDI causes economic development, and many governments spend precious resources on attracting FDI. A more thorough understanding of the relationship between FDI and development will help to inform governments of whether they should increase or reduce their efforts to attract FDI. In order to contribute to this information, we review the current literature on the topic.

There are several theoretical reasons why FDI could increase development. Most are based on the hypothesis that FDI increases productivity, especially through technology spillovers. In Global integration and technology transfer, Hoekman and Javorcik (2005) explore and provide evidence for technology spillovers caused by FDI. However, their results are mixed. They evaluate many studies and find that some show no statistically significant effect of FDI on productivity, but overall, they conclude that FDI does cause technology spillovers. They describe four specific mechanics through which FDI can cause technology spillovers.

Demonstration effects: When a foreign firm with access to new technology enters a new economy, if the firm is successful, its success demonstrates the economic viability of its technology. Then, local firms can copy the technology and adopt it.

Labor turnover: The foreign firm needs to either hire and train workers to operate its technology or bring workers from its origin economy. Some of these workers are likely to quit eventually and some will start their own firms or take jobs with local firms. They will then use the knowledge gained from the foreign firm to make use of its technology. Labor turnover could also cause spillovers of management techniques from the foreign firm.

Backward linkages: The foreign firm might need supplies that make use of technology that has not yet existed in the host economy. Then, it can either import these supplies from a different country, or it could share its knowledge with local suppliers to help them produce the technology. Sharing its knowledge to buy from local firms saves on transport costs relative to importing, but could dilute the its competitive advantage if local competitors adopt the technology.

Forward linkages: If the foreign firm's products make use of its advanced technology, its customers would need to understand it, and the firm would likely provide them instructions. They would also learn about the technology by using the product.

Keller and Yeaple (2002) use firm-level panel data in the US manufacturing sector from 1987-1996 to measure how FDI and imports' prevalence in the firms' industry have affected their total factor productivity (TFP). Their results are inconclusive for imports, but find that FDI has contributed TFP growth in US manufacturing. They estimate that FDI caused 11% of the TFP growth over the period. Their results are strong for firms in high-tech industries, but insignificant in low-tech industries.

The results on FDI's effect on productivity are mixed, but mostly weakly causally positive. The ambiguity in the results can be explained by some potentially negative effects of FDI on the host country's economy. Aitken and Harrison (1999) use plant-level panel data from Venezuela to gauge FDI's effect on productivity. They find that FDI correlates positively with productivity for small plants, but that it negatively affects the productivity of domestically owned plants through negative spillovers. They propose a market stealing effect as an explanation, meaning foreign competition forces domestic firms to reduce their production, which raises their cost per unit produced.

Lehnert, Benmamoun, and Zhao (2013) use data from 1997-2007 for 150 countries to estimate FDI's effect on the HDI and knowledge infrastructure, as measured by patent applications. They find a positive effect on HDI, and a negative effect on patent applications. As an explanation for the negative effect on patent applications, they propose that investing companies are reluctant to share their knowledge with competing domestic firms, while taking market share from them, thus crowding out domestic innovation.

Soumaré (2015) uses panel data from North African countries to estimate the effect of FDI on the HDI and GDP per capita. He finds a positive effect on both. Using a Granger causality test, he finds that the results are stronger for HDI, as he finds unidirectional causality in Egypt, Mauritania, Morocco, and Tunisia. He proposes that the industry concentration of FDI is the main reason for the differences between the results by country, and that investment focused in primary and manufacturing sectors are likely to have the greatest impact on development. FDI is mainly directed to the tourism industry in Morocco, utilities in Tunisia, and petroleum in Egypt and Mauritania. FDI in North Africa generally is largely directed to the petroleum industry.

We can use many of these papers to inform our hypotheses. Since most results indicate that FDI affects productivity and human development positively, we hypothesize that the causal relationship between FDI and most of the development indicators we have collected for Africa is positive.

5 Panel data regressions

Our main analysis is a standard panel data regression. The basic regression equation is:

$$DEV_t = \beta_1 FDI_t + \beta_n X_t + \epsilon_t$$

Here, DEV is a development indicator, FDI is the variable for FDI, X is a vector of n control variables, and ϵ is the error term. Our experiment uses 96 different development indicators. It uses 6 different FDI variables. The first four are total dollars invested, total jobs created, dollars invested for manufacturing, and jobs created for manufacturing. We have also gone through all the development indicators and found the industry that is most relevant to each of them. Our last two FDI variables are dollars invested and jobs created in the industry most relevant to each development indicator.

We use multiple variable to control for different factors. FDI could be an endogenous variable. Endogeneity means a positive relationship between FDI and development could mean that more developed countries attract more FDI, rather than that FDI causes development. To control for this, we include variables for *GDP per capita* and *trade openness* in the year before the year of our FDI variable. We also want to control for other things that could cause development to make sure our FDI estimator does not look more positive than it really is. For this we have chosen variables for *net aid received* and *gross domestic savings*, both of which could be used to finance development. We also want to control for things that could harm development, so our FDI estimator does not look worse than it really is. For this, we have chosen a variable for *external debt* and a *war dummy* that is equal to 1 if there is war in a given country in a given year, and 0 otherwise.

Thus, our program runs 600 complex and carefully designed regressions at a time. In order to handle the control flow, we used Python, and our regressions use the Statsmodels package, and its OLS module. Our program is also dynamic; we can change the control variable and many other setting to run regression equations that are different from our base model. The various setting we can include are a choice of 13 control variables, 4 lag structures, fixed effects controls by year and country, 5 interaction effects, and 2 instrumental variables. In total, this allows the program to

use over 420 billion different regression equations.

5.1 Control variables

Other control variables our program includes are: the *shares of FDI in manufacturing, high value-added industries, and extraction industries* to see if these affect development results, *internet penetration* since much of development is built in connectivity, *weapons imports* and *share of the labor force in the military* since this is indicative of a government's priorities, and a *financial openness index* as an alternative to control for endogeneity.

5.2 Lag structures

The program can evaluate development indicators up to 3 years later than the FDI variable, in order to find effects over time.

5.3 Fixed effects

Unobserved factors in a year or a country could affect our regression results. To control for this, we include dummy variables to indicate whether an observation is in a given country and year.

5.4 Interaction effects

An interaction effect is used to see if a combination of FDI and some other factor affects development. This is done by multiplying the FDI variable with another variable in the regression. The interaction variables we're using are: *political participation* and a *governance index* since better governance is likely to make FDI have a greater effect, *education enrollment* since a better educated workforce could take better advantage of the benefits of FDI, *access to electricity*, and *mobile phone penetration*, since a better connected population has a stronger ability to develop.

5.5 Instrumental variable

This is an advanced regression technique. It is an alternative to using regular variable to control for endogeneity. We find a variable that should cause more investment in a country, but should not be related to development. This is the instrumental variable. Then we run a regression with this variable against our FDI variable. Then we take the fitted results from this regression and put them in a new variable. Then we use the new variable as the FDI variable in a second regression

that is otherwise performed normally. Since all the variation in the new variable comes from the instrumental variable, all the endogenous parts of the FDI variable should be sucked into the error term in the first regression, and our new variable should not have endogeneity problems. Our instrumental variable options are *imports*, since companies often invest in a country as a substitute to exporting to it, and the *financial openness index*, since more financial openness should attract FDI, but might not affect development.

6 Difference in Differences

The difference-in-differences method evaluates the effect of individual projects, rather than the total FDI flow in a given year. The idea is to take a country which received a large FDI project and see how some characteristic of that country changes after the project relative to the change in other African countries. Thus, if a country receives a large project, we can see how its GDP changes from before to after the project, we can see how its neighbor's GDP changes from before to after. If the country that receives that project has a bigger change, this indicates FDI caused the difference between their changes.

The difference-in-differences approach is practically a t-test to see whether the difference between the changes in the two countries is equal to 0. We do this with a regression:

$$DEV_{ist} = \beta_1 Time_t + \beta_2 Treated_s + \beta_3 (Time \times Treated)_{st} + \epsilon_{ist}$$

Here, DEV is the development variable. The regressors are all dummy variables. $Time$ is 0 if we're looking at development before the FDI project and 1 if after. $Treated$ is 1 if we are looking at development in the country that received the FDI project and 0 if not. The diff-in-diff term is $Time$ multiplied by $Treated$. β_3, the coefficient on this variable, tells us whether the difference in differences is positive or negative.

6.1 General Discussion about this method:

$$Y_{ist} = \alpha + \beta T + \gamma Time_s t + \delta (Time \times Treated) + \epsilon_{ist}$$

Here, α and β are constants and account for average differences between the control and treated groups. γ accounts for average difference over time and δ is the coefficient of interest reflecting the true effect of the treatment on the treated sample.

6.2 Dummy Variables

Dummy variables are the variables that take only 2 values which are typically 0 or 1. For example in if the dummy variable is time t over which the sample has been treated then t is 0 before the treatment and t = 1 after the treatment. This helps in splitting the sample into sample sets, one before the treatment and one after the treatment.

6.3 Implementing Difference in Differences

We made use of 3 dummy variables, one for time one for treatment and the other one for interaction. Value of 0 was assigned to the variables before the treatment (Large FDI projects) and 1 after the treatment. Here the value of Y was represented by development indicators.

6.4 Issues with Diff in Diffs

- The key assumption undermining all difference in difference research is the common trends assumption, the trends in the treatment and control are the same.
- If the assumption does not hold then all the diff in diffs will be countered by any differences in trend between the treatment and control.
- Usually makes use of one pre and one post observation it is hard to make use of this method.

7 Results

Economic experiments are complicated. Every psychological feature of every individual is a variable, and billions of those variables are aggregated into that individual's behavior. Then millions of people's behavior is aggregated into a country's macroeconomic activity. Then hundreds of statisticians try to measure this activity in a way that is standardized, sound, and interpretable. Because of this complexity, it is hard to prove that any economic relationship is real. Thus, instead of focusing on our coefficients, we focus on the t-statistics of our regression results. If a variable has a t-stat with an absolute value above 1.96, its relationship with the dependent variable (the development indicator) is statistically significant within 95% confidence, and we say that the relationship is real. We highlight the three most interesting of our significant results here.

FDI seems to have a strong effect on *Gross Domestic Product (GDP)*. The table below shows the t-statistic for FDI on GDP. It includes all four lag structures and all 6 FDI variables used to

determine the robustness of our results. The industry-specific FDI variables for GDP are in the business services industry:

T-stats	jobs, industry	dollars, industry	jobs, total	dollars, total	jobs, manuf.	capex, manuf.
No lag	10.7	4.18	2.35	2.2	2.09	0.05
1 year	-4.79	-0.43	-0.34	-0.06	-0.6	-2.2
2 years	2.05	-0.82	0.24	-0.39	0.56	1.56
3 years	1.87	-0.58	-0.07	-0.73	0.64	-3.72

With large, positive t-stats with the lagless regressions, FDI seems to have a strong positive effect on GDP in the same year as the investment project. However, its effects seem to wane over time, as the t-stats are mostly insignificant after the year of the investment. This indicates that governments that want fast increases in GDP should attempt to attract FDI.

We also find uplifting results to the development indicator: *Individuals using the internet (% of population).* The industry-specific FDI variable here is investment in the software and IT industry:

T-stats	jobs, industry	dollars, industry	jobs, total	dollars, total	jobs, manuf.	capex, manuf.
No lag	2	3.87	0.79	0.8	1.16	0.07
1 year	3.16	2.96	2.84	0.83	2.74	0.29
2 years	3.22	2.79	3.22	0.95	3.78	0.3
3 years	3.01	2.67	3.56	1.08	4.06	1.06

Jobs created by Foreign Direct Investment seems to cause more people to adopt internet usage. Notice especially that all investment in the IT industry has a strong causal effect in the initial year. This is to be expected, since productive IT companies that invest in Africa should be able to provide IT products at a lower price than local competitors, making the internet more accessible to the local population.

We find some curious results for the indicator: *International tourism, number of arrivals.* The industry-specific FDI variable here is investment in the hotels and tourism industry:

T-stats	jobs, industry	dollars, industry	jobs, total	dollars, total	jobs, manuf.	capex, manuf.
No lag	-3.88	-3.08	4.08	4.25	2.14	0.67
1 year	-2.46	-2.14	0.36	1.61	-0.31	-0.6
2 years	-2.56	-2.15	0.23	1.7	0.04	-0.34
3 years	-2.77	-2.47	-0.28	1.98	0.38	0.66

The amount of tourist arrivals is consistently negatively impacted by investment in the tourism industry. Our t-stats for this results are large and stable, so it is a robust result. On the other hand, total FDI in the initial year seems to have a strong positive effect on tourist arrivals. This is an odd result. It may be that investment in the tourism industry causes a lot of construction of hotels, which makes a country a less attractive tourist destination. At the same time, countries tend to invest in countries that a culturally similar to themselves. It is reasonable to expect that people tend to vacation in places that are fairly cultural similar to their homes. This could cause the positive effect of total FDI on tourist arrivals.

We have tested each of our interaction terms in our standard regression equation, letting GDP remain the dependent variable in each of the interaction terms. The interaction variables are: *Political participation*, *Ibrahim Index of African Governance*, *University enrollment*, *Access to electricity*, and *Mobile phone subscriptions (% of population)*. A positive result for the interaction term means the interaction variable combines with FDI to cause development:

	Participation	Governance	University	Electricity	Phones
T-stats	2.64	2.47	-0.03	2.22	14.9

We find strong positive results for most of our interaction terms, especially Mobile phone penetration, but remarkably insignificant results for our education variable. This tells us political participation, good political governance, access to electricity, and connectivity all amplify the positive effects of Foreign Direct Investment. However, university enrollment does not. This is surprising, since it is generally thought that a country's ability to take advantage of FDI depends on its ability to make use of foreign technology, which again depends on its level of education.

8 Conclusion

Our results are broadly informative. They tell us Foreign Direct Investment has significant positive effects on African host countries. It can help spread the use of the internet and boost GDP. We also find reason for caution, given that it can negatively affect tourist arrivals. Equally important lesson from our research is in the results from our interaction terms. Developing countries might

not want to focus on FDI alone, since its effect is amplified by various other factors. The result from our political participation term is especially relevant today, as some less democratic African governments are considering copying the undemocratic Chinese model for development. Our results show that democracy is a large benefit for taking advantage of FDI for development. We also show the vast importance of connectivity. The result of the mobile phone penetration interaction indicates developing governments should focus on improving connectivity before focusing on attracting FDI, in order to optimize development effects.

Our results lay the groundwork for future avenues of research. FDI's effect on tourism seems mysterious. Tourism is an important industry for many African countries, and understanding FDI's effect on it is important for these countries. Thorough case studies could help elucidate this topic. Our results for the university enrollment interaction term go against most economic intuition. More experimentation is warranted to validate the robustness of this result. Our dataset also allows for experimentation with many more development indicators than presented here, but a much larger work is required to present these results thoroughly.

9 References

- Hoekman, B. M., Javorcik, B. K. (2005). Global integration and technology transfer. Washington, DC: World Bank.
- Keller, W., Yeaple, S., (2003). Multinational enterprises, international trade, and productivity growth: Firm-level evidence from the United States. IMF Working Paper/03/248.
- Lehnert, K., Benmamoun, M., Zao, H. (2013). FDI inflow and human development: Analysis of FDI's impact on host countries' social welfare and infrastructure. Thunderbird International Business Review, 55(3), 285-298. doi:10.1002
- Soumaré, I. (2015). Does FDI improve economic development in North African countries? Applied Economics, 47(51), 5510-5533. doi:10.1080