Appendix

This is the article's appendix: "Latin America: Does the amount of International Agreements influences the inflow of Foreign Direct Investment? Comparing analyzes in the data."

The following are the (1) exploratory data analysis, (2) the treatment techniques used, (3) the assumptions analysis, (4) the models' adjustment, (5) estimated models and (6) references.

1) Exploratory data analysis

Exploratory data analysis seeks to provide an overview of the data and the relationship between the variables, revealing their trends and patterns, as well as possible anomalies. Performs descriptive statistics of the data.

The article presents as dependent variable the amount of foreign direct investment received as a percentage of Gross Domestic Product by Latin American countries between 1970 and 2009, data provided by UNCTAD and selected in this way to make it possible to compare over time between countries, dispensing deflation (BÜTHE & MILNER, 2008, p. 748), i.e., the need for an index that isolates the real growth of the artificial product caused by the price increase in the economy (IPEA, 2007).

The rationale for choosing UNCTAD data is that the data would be less affected by intentional omissions in their reporting when it comes to developing countries (BÜTHE & MILNER, 2008, p. 748). The Latin America region refers to 20 countries (MAINWARING and PÉREZ – LIÑÁN, 2009, p.532): Argentina, Bolivia, Brazil, Uruguay, Venezuela, Paraguay, Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Haiti, the Dominican Republic, and Cuba.

The summary of the amount of FDI as a percentage received in Latin America ranges from - 10.6749% of a country's GDP to 16.3549% of a country's GDP and has an average value, for the 20 countries, of 1.8754%. In other words, the average percentage of foreign direct investment in Latin America in the period

from 1970 to 2009 is almost 2% of the countries' GDP. Also, the standard deviation of the variable is approximately 2.45 and the coefficient of variation (d.p./average) approaches 1.30 or 130%, demonstrating a large variance between the data that can be observed by the histogram. Most of the FDI values, however, are between 0 and 4% of countries' GDP.

The quantitative independent variables used are: Number of PTAs in force, Number of BITs a country is signatory, Political Constraint Index in a Country - POLCON 2010, Political Instability, Market Size, Economic Development, and Per capita GDP Growth. In addition, the categorical independent variable Participation in GATT or WTO is used, after its creation.

Number of PTAs in force

The number of PTAs in force represents the number of preferential trade agreements in force from 1970 to 2009 for Latin American countries. Importantly, the total number of treaties in force at the time of publication of the database (2013) was 2431 PTAs, compared with 2305 PTAs in the 2008 article. Signed PTAs total are 2659, which of 228 are signed, but not in force.

One hypothesis is that the greater the number of PTAs signed by countries, the more FDI will be received in the region. The minimum value is 0, i.e., no signed agreement and the maximum value is 21, i.e., 21 signed agreements in force by Chile. The average of signed agreements is around 3 and the standard deviation around 3 as well. Plotting the variable on the chart shows that most countries sign up to 3 preferential agreements per year. It is shown that the ratio of the number of PTAs in force to the inflow of FDI in proportion to GDP is positive and concentrated around 0 to 5 agreements.

Number of BITs

The number of bilateral investment treaties to which a country is a signatory, provided by UNCTAD, contains specific provisions on the treatment of foreign investors and the definition of investment (BÜTHE and MILNER, 2008, p.749), which differentiates them into restricted or not. Thus, there is a difference of 119 treaties when they are differentiated into cumulative and restricted cumulative.

The minimum number of agreements is 0 and the maximum number is 61 treaties signed by Cuba. The mean number of treaties is close to 9 and the standard deviation slightly above 13, indicating a large variance in the data, with a coefficient of variation of 1.56.

Most treaties occur up to 5, demonstrating that, like PTAs, the number of BITs signed per year in each country is low, which is natural due to their complexity. The relationship of the dependent variable with BITs is positive and concentrated at the beginning of the chart, indicating a large amount of little firm BITs per year and the impact of these BITs on FDI is in the range of 0 to 5% of countries' GDP.

Political Constriction Index - POLCON

The other variable is the Political Constraint Index (POLCON), produced by Witold J. Henisz to measure political constraints in a country, identifying secondary political structures and the capacity to support credible political commitments (BÜTHE and MILNER, 2008, p. 749). It is considered the weighted measure in the number of players with veto power in the national political system. The index was published in 2000 and this article uses POLCON 2010, which has data up to 2007.

The index ranges from 0 to 1, establishing domestic political constraints on a scale where 0 represents a high level of political risk and instability and 1 a low level of political risk. In this dataset, Cuba has a POLCON of 0, while Brazil in the early 1990s reaches the maximum value for Latin American countries within the time range. The index shows that the average of countries for the index is around 0.3, demonstrating a high level of domestic political constraints in the region. The relationship with the dependent variable is slightly positive, with most data concentrated just above the index average in the FDI receiving range of up to 5% of GDP.

Political instability

The next variable is political instability, a measure composed of Arthur Banks (1999) data for political events that indicate political violence and instability. The database is presented as transnational time series data for more

than 200 countries covering economic, political, population and demographic variables, with emphasis on data from regimes and political systems (BÜTHE and MILNER, 2008, p.749).

The maximum value of the index for countries is 33 and the minimum is 0, averaging around 3.5, i.e., most countries have low political instability. The maximum value was for Argentina coinciding with the death of then President Perón. The standard deviation reaches 4.66 and the coefficient of variation at 1.37, meaning large variation in the data. Plotting the frequency of the variable shows that it is all concentrated to the left of the graph, representing the low political instability of the countries in the region. Its relationship with the dependent variable is slightly negative, that is, the greater the political instability, the lower the amount of FDI received.

The following are the three standard variables of economic control models: market size, represented by the log of the country's population; economic development, represented by the log of GDP - total value of final goods and services produced in the economy, usually over a period of one year (KRUGMAN and WELLS, 2007, p.488) - per capita in dollars in 2000; and GDP growth, representing by the percentage change in the real GDP of the country in relation to the previous year (BÜTHE and MILNER, 2013, p.102). The original database was updated with 2010 data provided by the World Bank.

Market Size

Market size in countries is relatively close. The minimum population log value is 14.23, while the maximum value is 19.08. The average is around 16 and the standard deviation around 1, which is reaffirmed by the coefficient of variation, also below 1. Plotting on the histogram only reaffirms the distribution with the highest population concentration in around the average. The relationship between the dependent and independent variables is almost nonexistent, slightly negative. The FDI range in GDP remains concentrated between 0 and 5% with the population average between 15 and 17.

Economic development

Economic development is also close across countries, with 5.93 as the minimum value of the GDP per capita log and 9.2 as the maximum. The average is around 7.6 and both standard deviation and coefficient are less than 1, with a slight variation in the data. When plotted on the histogram, the concentration around the mean is visible and, again, the relationship between the dependent and independent variable is almost nonexistent, slightly positive, suggesting that the higher the economic development, the greater the inflow of FDI. The data are mostly concentrated in the development range of 7 to 8 log per capita GDP.

GDP growth

GDP growth, as a percentage change, shows great variation between countries, ranging from -26.479 (minimum value), ie negative growth, showing that the country produced less when compared to the previous year. Real GDP serves to determine aggregate output growth based on some base year in contrast to nominal GDP that is calculated at current prices, “the measure of average aggregate output per person” (KRUGMAN and WELLS, 2007, p.497).).

The maximum value is 18.28, with an average of 3.44. The standard deviation is 4.5 and the coefficient of variation is 1.31, that is, there is great variance in GDP growth among countries in the region. Plotting on the histogram shows that most variations are between 0 and 10%. In addition, there would be a positive relationship between FDI inflow and GDP growth, suggesting that the higher the GDP growth, the greater the FDI inflow. The chart shows a massive concentration of data between 0 and 10% for the percentage change in GDP and from 0 upwards in relation to the inflow of FDI.

GATT / WTO Participation

Finally, there is the categorical independent variable GATT / WTO, computing the participation of countries in GATT / WTO as 1 and 0 as not. The idea is that participation in these multilateral organizations will promote greater FDI entry into the countries of the region. Altogether there are 535 measurements, 535 being the number of participating countries per year and 250 the number of non-participating countries. Boxplot analysis shows a greater variation in the number of participating countries, but the median in both is almost

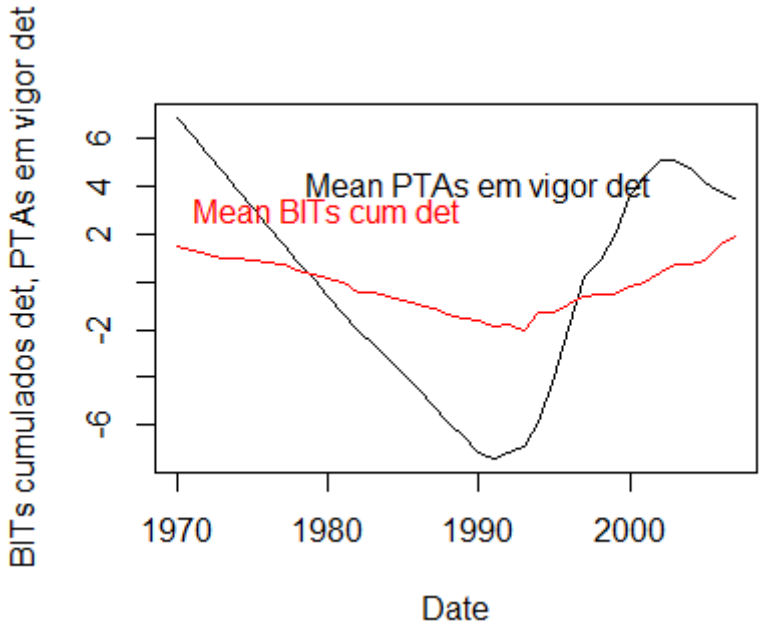
the same, not visually indicating a large influence of FDI in the distribution of participation in GATT / WTO.

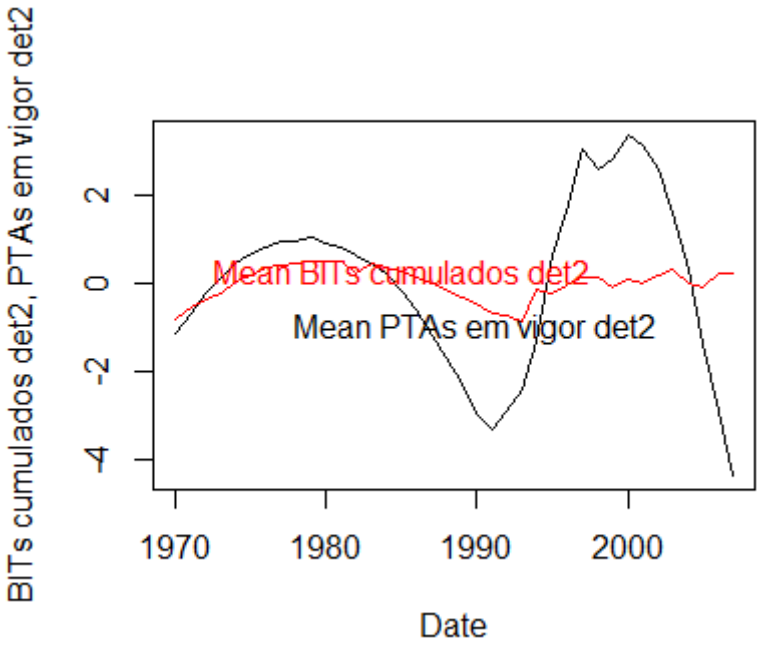
2) Treatment Techniques

The data used in this article are time series for each member in a dataset, in this case countries. Data were collected for the same variables over time and for the same group of countries, Latin American countries. That is then panel data. (WOOLDRIDGE, 2002, p. 6-10).

Time series are series with observations of variables over time. Since past events often influence future events, it is not possible to consider the observations collected as independent over time. When it sets in time and presents a growing trend, this trend can lead to inferences that do not correspond to reality, leading to correlations between positive trends, for example, that are linked to unobserved factors (WOOLDRIDGE, 2002, p. 331).

As in this article, this is exactly what is observed. That is, the inflow of FDI and the number of international treaties grow, therefore, the data processing techniques used were a linear and a quadratic detrend. The following is an example of linear detrend followed by a quadratic:





The comparative analysis of data in the adoption of a linear detrend strategy by Büthe and Milner's (2008) was criticized by King and Roberts (2014) on the grounds that the authors performed linear detrend on quadratic trends and applied the same trend for all countries, contrary to the diversity of developing countries dealt with in the article, which would generate a new spurious pattern in the time series (KING and Roberts, 2014, p. 174).

King and Roberts (2014) solution is the quadratic detrend for each country, which would generate results in the variables closest to a stationary model approximating robust and classic standard errors in the model. The key result is that after the change in technique, the variables of interest are no longer significantly correlated with receiving FDI, ie, the higher number of PTAs, for example, would not significantly influence the amount of FDI received (KING and ROBERTS, 2014, p. 174).

This paper then generates 3 models and compares them without detrend, with linear detrend and quadratic detrend to realize the effect of model techniques in a regional context in which the theoretical framework is close. If the techniques used initially by the authors for a database subset are tested, the results are maintained or due to the proximity of context, the contextually closer

data will allow a significant correlation between the variables independent of the technique.

The linear detrend was performed by the pracma package of R which specifically “computes the least-squares fit of a straight line (or composite line for piecewise linear trends) to the data and subtracts the resulting function from the data” (R DOCUMENTATION, 2019). The main difference between the quadratic and the linear detrend is that for the former, an exponential behavior of the data is assumed, and the time is squared.

Regarding the variables, it is important to highlight the modifications of the original article database, raised by the authors and used by Gary King for the contestation of the model, and the database used in this article. The data used in this article come from an article by Büthe and Milner in which model 4 is the basis for the article and from it, the authors update the variables in time and interpretation, which was also performed here.

The main changes are in the accuracy of the choice of the dependent variable, the update of the POLCON index and the update of the economic development measurement, represented by the log of per capita GDP in dollars. Regarding the dependent variable, the article uses signed and ratified PTAs in force, not just signed PTAs, on the grounds that only the agreements in force would have an impact on the inflow of FDI (BÜTHE and MILNER, 2014, p. 103). Regarding POLCON, its data is updated to 2010 instead of 2002, used in the first model; and for economic development, the 2000 GDP log is used instead of the 1995.

The variables are also lagged under the justification of any effect of the independent variables on the dependent variable may take some time to be perceived and, consequently, measured. The option of the model is then to offset the independent variables by 1 year (BÜTHE and MILNER, 2008, p.748). It is also assumed that the value of FDI is determined by its past value.

Another important concept for the model is the stationary process for time series. This occurs when probability distributions remain stable over time, indicating that a temporal sequence for a given value is equally distributed and

the nature of correlations between terms remains the same across all time periods (WOOLDRIDGE, 2002, p. 348).

A model that produces waste near a stationary process approximates robust standard errors - used to correct the estimation model - from the classical ones, indicating greater accuracy for the model. Distance between them may indicate failures and poor specification in the adoption of the statistical model (KING et ROBERTS, 2014, p.159).

Thus, robust standard errors make it possible to conform to the assumptions of a linear model, for example, to ensure maximum consistency and absence of bias, even though the production of the data does not fit the model a priori (KING et ROBERTS, 2014, p.161).

This exposed, follows by analyzing the assumptions of the 3 models.

3) Assumption Analysis

Multiple regression analysis serves to analyze the relationship of a dependent variable with several independent variables, estimating the degree of association that may exist between them (HAIR et al., 2014). The notation of the model in its simplest form has the following formula:

$$Y = \alpha + \beta 1X1 + \epsilon$$

Y is the dependent variable and what we want to explain in the model, X1 above, represents the explanatory variable of the model that in its multivariate formula would be composed of more than one term, and the influence of each variable in the model is analyzed. Alpha, the intercept, represents the value of Y when X1 is zero, β represents the marginal increase of X1 in Y, and ϵ represents the difference between the observed values and the predicted errors or model residuals.

These have assumptions that must be analyzed in the data in order to adopt the multivariate linear regression, assuming that the smaller the residue, the closer the reality model is. Following are the ones that will be analyzed in this appendix, emphasizing that there is no uniformity about which or how many assumptions are needed among the authors to validate a model (FILHO et al., 2011, p.51):

(1) the expectation of the error term average is zero;

(2) homoscedasticity, that is, the variance of the error term is constant for the different values of the independent variable;

(3) the independent variables do not present a high correlation, the so-called non-multicollinearity assumption;

(4) the error term is assumed to have a normal distribution;

(5) there is an adequate ratio between the number of cases and the number of estimated parameters;

(6) no theoretically relevant variable to explain Y was left out of the model and no irrelevant variable to explain Y was included in the model;

(7) the variables were adequately measured, i.e, it is assumed that there is no systematic measurement error; and

(8) The relationship between the dependent variable and the independent variables must be linear.

The theoretical function of each assumption in the model is explained followed by its application in the article models. The adopted model is:

$$IEDit = \alpha + \gamma1\ (Market\ Size)\ i\ (t - 1) + \gamma2\ (Economic\ Des.)\ I\ (t - 1) + \gamma3\ (GDP\ Growth)\ i\ (t - 1) + (current\ PTAs)\ i\ (t - 1) + (GATT / WTO)\ i\ (t - 1) + (cumulative\ BITs)\ i\ (t - 1) + (POLCON)\ i\ (t - 1) + (install\ inst.)\ i\ (t - 1) + \delta i + \varepsilon it,$$

where δi indicates fixed country effects implemented through a set of n - 1 country dummies.

The first assumption states that the mean error value must be zero, indicating that error values not included in the model did not affect its result, because the values cancel each other out when equidistant. In all three models this assumption passes, with the average waste very close to 0.

The second assumption, homoscedasticity, assumes that the residues vary uniformly in the model; if they do not, they interfere with the model's compatibility range and significance tests. One of the causes of

heteroscedasticity is measurement error in independent variables (TABACHICK AND FIDELL, 2007, in FILHO et al., 2011).

Homoscedasticity is guaranteed by applying the Arellano (1987) function in the generation of models 2 and 3. Since after the application of the function it is not possible to test the behavior of the residuals, the assumptions of models 2 and 3 are evaluated before the application of the model. function and it is verified that before it there is no homoscedasticity in the data.

After application, it is assumed that the function is robust for heteroscedasticity and autocorrelation (the value of a variable gained at a given time does not affect the value of a subsequent measurement). Before the application of the function, however, the tests performed have a p-value less than a level of 0.01, indicating that the variance of the residuals does not occur uniformly, which is confirmed by the graphical analysis.

The third assumption refers to multicollinearity. “Multicollinearity refers to the excessive correlation between predictor variables” (GARSON, 2011 in FILHO et al. 2011). This excessive correlation generates redundant information about the model response. The analysis technique used for multicollinearity measurement is Variance Inflation Factor (VIF), indicating that the higher, the worse, ie, the higher the correlation of the variables under ideal conditions.

The limit value of the VIF for the models varies among authors. Adopting 4 (HAIR et. Al, 2014), we analyze the three models, realizing that only the 3 would present absence of multicollinearity with all VIFs below 2. For the first model, there is a high multicollinearity between Market and Economic Development. For the second model, only for Economic Development. For the third, there is not.

The fourth assumption is the normal distribution of waste. In the three models it is clear that this assumption is met. It is observed by both the graphical analysis and the Shapiro normality test that the models present normal distribution of residues. The Shapiro test on the three models yielded p-values below 0.001, indicating a strong indication of the normal distribution of waste. This is a premise of Gauss-Markov's theorem to ensure that estimators are unbiased and efficient (FILHO et al, 2011, p. 59).

The fifth assumption requires a minimum ratio between the number of cases and the estimators. Using the formula of Tabachnick and Fidell (2007 in FILHO, 2011) where N ≥ 50 + 80X (where X is the number of independent variables included in the analysis), it is calculated that for the model the minimum number of observations would be 690. analysis considers 785 observations.

The sixth assumption requires an adequate specification of the model with the theoretical justification for including the relevant variables to the model and excluding the irrelevant ones, so as not to increase the error in the estimates.

The seventh assumption assumes no measurement error. As the data used are secondary to sources with credible support, such as UNCTAD, World Bank, and index usually adopted in the literature, it is assumed that there was no error in the measurement of variables.

Finally, the eighth assumption states that there is the linearity of the parameters, i.e. that there is a linear relationship between the model variables. Regarding the association of quantitative dependent variables with the independent variable, using graphical analysis and Pearson's correlation test with the indication of p-value, we have that for:

1 - Number of PTAs in force - there is a linear relationship with the data, which is concentrated to the left of the graph, decreasing the compatibility range of the sample in relation to the left side. The p-value of the test is lower than the significance level of 0.05, indicating that there is a correlation between the variables, but the linear relationship is positive and moderate, being 0.47.

2 - Number of BITs - there is a linear relationship with the data, which is concentrated to the left of the graph, decreasing the compatibility range of the sample in relation to the left side. The p value of the test is lower than the significance level of 0.05, indicating that there is a correlation between the variables, but the linear relationship is positive and moderate, being 0.33.

3 - Political Constriction Index - POLCON - there is a linear relationship with the data, which is concentrated more in the middle of the graph, reducing the compatibility range of the sample in relation to the medium. The p value of the test is lower than the significance level of 0.05, indicating that there is a

correlation between the variables, but the linear relationship is positive and weak, being 0.24.

4 - Political Instability - there is a linear relationship with the data, which is concentrated to the left of the graph, decreasing the compatibility range of the sample in relation to the left side. The p value of the test is lower than the significance level of 0.05, indicating that there is a correlation between the variables, but the linear relationship is negative and weak, being - 0.12.

5 - Market size - the linear relationship between the variables is very close to 0, close to the absence of relationship. The p value of the test is lower than the significance level of 0.05, indicating that there is a minimum correlation between the variables, but the linear relationship is negative and very close to 0, being - 0.083.

6 - Economic Development - the linear relationship between the variables is very close to 0, close to the absence of relationship. The p value of the test is lower than the significance level of 0.05, indicating that there is a minimum correlation between the variables, but the linear relationship is positive and very close to 0, being 0.092.

7 - GDP growth - there is a linear relationship with the data, which concentrates more in the middle of the graph, reducing the compatibility range of the sample in relation to the medium. The p value of the test is lower than the significance level of 0.05, indicating that there is a correlation between the variables, but the linear relationship is positive and weak, being 0.19.

Regarding outliers, model 1 shows the points at the beginning and end of the curve outside the compatibility range, indicating the need for a more robust regression to correct possible outlier effects (FOX, 2008 in FILHO, 2011). Model 2, with the linear detrend in the variables, can best fit the curve of the compatibility range, with few values outside it compared to model 1. Model 3, with the quadratic detrend in the variables, also fits straight for the most part, its distribution being very close to model 3.

4) Model adjustment

Evaluating the explanatory capacity of model 1, R-squared reports the percentage value for the explanatory content of the model, so that the higher the better the model fits the data. The explanatory capacity of the model is 34.6%, ie, 34.6% of the percentage change in GDP in inflow of FDI.

The mean square error (RMSE) provides the measure of the average model precision with the dependent variable metric (KELLSTEDT and WHITTEN, 2009, p. 201). It is presented in the same unit as the dependent variable, ie, for model 1, on average, the variation of the predicted values in the model is 1.98 percent in relation to the percentage of FDI receiving GDP in developing countries in America. Latina.

For model 2, there is a decrease of 19.86% in the explanatory capacity of the model, ie, it is now explained, after the linear detrend almost 20% of the percentage change of GDP in inflow of FDI. The RMSE also decreases, on average, the change in the predicted values in model 2 is 1.66 percent over the percentage of FDI receiving GDP in developing countries in Latin America.

For model 3, there is another decrease in the explanatory capacity of the model, here 6.98%, that is, it is now explained, after the quadratic only 7% of the percentage change of GDP in inflow of FDI. However, there is an increase in the quality of the average precision of the model for the variation of the dependent variable, rising to 1.71 percentage in relation to the percentage of FDI receiving GDP in the developing countries of Latin America, being able to explain a larger variation in the FDI. inflow of FDI in relation to 1.

It is interesting to note how the treatment of variables affects the explanatory capacity of the model, suggesting that the trend in data greatly increases the explanatory power of a model by the influence of time in enhancing relationships between variables.

Regarding the compatibility range of the models, showing a 95% probability that the population values are contained in the sample, many cross the value 0, indicating that the p value for the hypothesis test that the parameters will be equal to 0, will be greater than 0.05. The variables that maintain close effects in the three models are: PTAs in force, GATT / WTO. BITs in the third model cross the 0, which does not happen in the first two.

The behavior of the parameters in model 2 and 3 is close, coinciding for most parameters and denoting greater precision. The behavior of the parameters in model 1 has larger range ranges, but also encompass that of other models, indicating that they are not in conflict, although for some there is statistical significance and for others not (NATURE, 2019).

5) Estimated models

Model 1:

```
Call:
lm(formula = fdi_inflow_unctad_gdp ~ lag_pta_force + lag_gattwto +
    lag_bits_cuml_restricted + lag_polconiii_2010 + lag_polinstability
+
    lag_ln_pop + lag_ln_gdp_pc_00d + lag_gdp_gr + factor(country),
    data = fdi)

Residuals:
     Min       1Q   Median       3Q      Max
-14.4076  -0.7973  -0.1847   0.5928  12.7539

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)              -3.19905   11.40555  -0.280 0.779196
lag_pta_force             0.19657    0.06272   3.134 0.001802 **
lag_gattwto               0.42357    0.29195   1.451 0.147303
lag_bits_cuml_restricted  0.04530    0.01247   3.632 0.000303 ***
lag_polconiii_2010        0.23047    0.48457   0.476 0.634506
lag_polinstability       -0.02036    0.01640  -1.241 0.215024
lag_ln_pop                0.33486    0.63459   0.528 0.597902
lag_ln_gdp_pc_00d        -0.33923    0.51563  -0.658 0.510840
lag_gdp_gr                0.07489    0.01796   4.169 3.47e-05 ***
factor(country)2          2.00118    1.46427   1.367 0.172195
factor(country)3         -0.30270    1.17015  -0.259 0.795961
factor(country)4          1.29320    0.89018   1.453 0.146775
factor(country)5          0.04184    0.87915   0.048 0.962060
factor(country)6          0.67395    1.18618   0.568 0.570114
factor(country)7          0.35336    1.54057   0.229 0.818654
factor(country)8          0.19406    0.84050   0.231 0.817479
factor(country)9          0.86603    1.55413   0.557 0.577551
factor(country)10         0.08278    0.66497   0.124 0.900968
factor(country)11         1.92583    1.61521   1.192 0.233570
factor(country)12         0.96730    1.35256   0.715 0.474764
factor(country)13         0.70167    1.21141   0.579 0.562641
factor(country)14         1.06780    1.57389   0.678 0.497727
factor(country)15         0.25921    0.92029   0.282 0.778294
factor(country)16         1.04203    1.75473   0.594 0.552822
factor(country)17         4.14504    1.72166   2.408 0.016334 *
factor(country)18        -0.70279    1.76808  -0.397 0.691139
factor(country)19         1.32009    1.21924   1.083 0.279330
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.023 on 655 degrees of freedom
  (103 observations deleted due to missingness)
Multiple R-squared:  0.346,    Adjusted R-squared:   0.32
F-statistic: 13.33 on 26 and 655 DF,  p-value: < 2.2e-16
```

Model 2:

```
t test of coefficients:

                              Estimate Std. Error  t value  Pr(>|t|)
(Intercept)                   0.450324   0.247650   1.8184 0.0696704 .
lag_pta_force_det             0.163945   0.115116   1.4242 0.1550934
lag_gattwto_det               0.994567   0.456922   2.1767 0.0300259 *
lag_bits_cuml_restricted_det  0.067051   0.024562   2.7299 0.0065843 *
*
lag_polconiii_2010_det        0.823389   0.966460   0.8520 0.3946885
lag_polinstability_det       -0.042139   0.019238  -2.1905 0.0290051 *
lag_ln_pop_det               -0.297094  12.467785  -0.0238 0.9809996
lag_ln_gdp_pc_00d_det        -0.651725   1.149122  -0.5672 0.5708952
lag_gdp_gr_det                0.057957   0.018577   3.1199 0.0019260 *
*
factor(country)2              0.297167   0.062300   4.7699 2.496e-06 *
**
factor(country)3             -0.457498   0.082163  -5.5682 4.438e-08 *
**
factor(country)4              0.050243   0.163649   0.3070 0.7589715
factor(country)5             -0.606043   0.098800  -6.1340 1.882e-09 *
**
factor(country)6             -0.596545   0.188756  -3.1604 0.0016821 *
*
factor(country)7              0.197158   0.145435   1.3556 0.1758942
factor(country)8             -0.197233   0.122418  -1.6111 0.1078502
factor(country)9             -0.906327   0.112083  -8.0862 5.778e-15 *
**
factor(country)10            -0.021891   0.128181  -0.1708 0.8644751
factor(country)11            -0.830202   0.152714  -5.4363 8.944e-08 *
**
factor(country)12            -0.596262   0.128033  -4.6571 4.231e-06 *
**
factor(country)13            -0.194000   0.079306  -2.4462 0.0148188 *
factor(country)14            -0.800020   0.118193  -6.7688 4.073e-11 *
**
factor(country)15            -0.379589   0.193240  -1.9643 0.0501065 .
factor(country)16            -0.640420   0.172980  -3.7023 0.0002403 *
**
factor(country)17            -2.396916   0.109052 -21.9795 < 2.2e-16 *
**
factor(country)18             3.434957   6.921825   0.4963 0.6199607
factor(country)19            -0.678059   0.070225  -9.6556 < 2.2e-16 *
**
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Model 3:

```
t test of coefficients:

                      Estimate  Std. Error t value  Pr(>|t|)
(Intercept)         0.23104520  0.15411785  1.4991  0.134326
pta_det2            0.29368017  0.09313216  3.1534  0.001689 **
gattwto_det2       -0.02158015  0.50063503 -0.0431  0.965631
bits_det2           0.01927869  0.02303007  0.8371  0.402842
pol_det2            0.18769249  0.67763846  0.2770  0.781884
lag_polinstability -0.02735114  0.01824448 -1.4991  0.134326
pop_det2            0.84902951  1.87878641  0.4519  0.651491
gdppc_det2          0.01229045  0.67963908  0.0181  0.985578
gdpgr_det2          0.06851225  0.01702592  4.0240 6.403e-05 ***
factor(country)2   -0.08763649  0.05845760 -1.4991  0.134326
factor(country)3   -0.11804178  0.07873934 -1.4991  0.134326
factor(country)4   -0.10076738  0.06721651 -1.4991  0.134326
factor(country)5   -0.04174648  0.02784684 -1.4991  0.134326
factor(country)6   -0.14529566  0.09691894 -1.4991  0.134326
factor(country)7   -0.19505685  0.13011196 -1.4991  0.134326
factor(country)8   -0.10242089  0.06831948 -1.4991  0.134326
factor(country)9   -0.17994174  0.12002948 -1.4991  0.134326
```

```
factor(country)10  -0.15618943  0.10418559 -1.4991  0.134326
factor(country)11  -0.21736962  0.14499561 -1.4991  0.134326
factor(country)12  -0.11732202  0.07825922 -1.4991  0.134326
factor(country)13  -0.07285209  0.04859572 -1.4991  0.134326
factor(country)14  -0.17490337  0.11666866 -1.4991  0.134326
factor(country)15  -0.18887885  0.12599095 -1.4991  0.134326
factor(country)16  -0.15978827  0.10658618 -1.4991  0.134326
factor(country)17  -0.17431690  0.11627745 -1.4991  0.134326
factor(country)18   0.00052783  0.00035209  1.4991  0.134326
factor(country)19  -0.16482664  0.10994701 -1.4991  0.134326
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
```

6) References

ARELLANO, M. **Computing Robust Standard Errors for Within-groups Estimators**. Oxford Bulletin of Economics and Statistics 49 (4): 1987, p. 431–34. Available at: https://www.cemfi.es/~arellano/OBES_1987.pdf. Accessed on 02.09.2019.

BANKS, ARTHUR S. **Cross-National Time-Series Data Archive**. 1999. Binghamton, NY: Banner Software. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/K2STK. Accessed on 02.09.2019.

BÜTHE, TIM AND MILNER, HELEN V. **The Politics of Foreign Direct Investment in Developing Countries: Increasing FDI through International Trade Agreements?** American Journal of Political Science, Vol. 52, No. 4, October 2008, p. 741–762. Available at: https://people.duke.edu/~buthe/downloads/ButheMilner_AJPS_Oct2008.pdf. Accessed on 02.09.2019.

_____**Foreign Direct Investment and Institutional Diversity in Trade Agreements: Credibility, Commitment, and Economic Flows in the Developing World, 1971–2007.** World Politics, 66, 2014, pp 88-122. doi: 10.1017 / S0043887113000336. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1796506. Accessed on 02.09.2019.

SON, DF, NUNES, F., ROCK, EC DA, SANTOS, ML, BAPTIST, M., & JR, JAS **What to Do and What Not to Do with Regression: Assumptions and Applications of the Ordinary Least Squares Linear Model** (COD). 2011.

Political Journal Today - ISSN: 0104-7094, 20 (1). Available at: https://periodicos.ufpe.br/revistas/politicahoje/article/view/3808. Accessed on 02.09.2019.

HAIR JR., JOSEPH; ANDERSON, RALPH; TATHAM, RONALD; BLACK, WILLIAN. **Multivariate data analysis**. 17th Edition. Pearson Education Limited. 2014. Available at: https://is.muni.cz/el/1423/podzim2017/PSY028/um/_Hair_-_Multivariate_data_analysis_7th_revised.pdf. Accessed on 30.08.2019.

IPEA (Institute of Applied Economic Research) by SOUZA, JORGE LUIZ DE. Which is? Implied deflator. Development Challenges. Year 4. Issue 37. 2007. Available at: http://desafios.ipea.gov.br/index.php?option=com_content&view=article&id=2140:catid=28&Itemid=23. Accessed on 02.09.2019.

KELLSTEDT, PAUL M. and WHITTEN, GUY D. **Fundamentals of Political Science Research**. (2nd ed.) Blucher, 2009.

KING, GARY AND ROBERTS, MARGARET E. **How Robust Standard Errors Expose Methodological Problems They Do Not Fix, and What to Do About It.** Political Analysis 23, 2015, p.159–179. doi: 10.1093 / pan / mpu015. Available at: https://gking.harvard.edu/files/gking/files/robust_0.pdf. Accessed on 20.08.2019.

KRUGMAN, PAUL and WELLS, ROBERTS. **Introduction to economy**. Rio de Janeiro, Elsevier: 2007.

NATURE. AMRHEIN, VALENTIN; GREENLAND, SANDER; MCSHANE, BLAKE. Scients rise up against statistical significance. 20 March 2019. Available at: https://www.nature.com/articles/d41586-019-00857-9. Accessed on 02.09.2019.

MAINWARING, SCOTT AND PÉREZ – LIÑÁN, ANIBAL. **Why are the regions important? Regional specificities and diffusion of democracy on a regional scale**. SAAP Magazine (ISSN 1666-7883) Vol. 3, No. 3, August 2009, p. 523-566. Available at: https://dialnet.unirioja.es/servlet/articulo?codigo=5238129. Accessed on 22.08.2019.

R DOCUMENTATION. Detrend From pracma v1.9.9 by HwB. Available at: https://www.rdocumentation.org/packages/pracma/versions/1.9.9/topics/detrend Access on 18.08.2019.

WOOLDRIDGE, J. **Introductory Econometrics: A Modern Approach** (2nd ed). South-Western College Pub. 2002.