CENTRAL ASIAN JOURNAL OF THEORETICAL AND APPLIED SCIENCES

Volume: 04 Issue: 05 | May 2023 ISSN: 2660-5317
https://cajotas.centralasianstudies.org

THE EFFECT OF FOREIGN DIRECT INVESTMENT ON THE EXPORT AUTHORITY OF GOODS AND SERVICES

Saydullayev Azamat Jo'raqul o'g'li

TSUE Phd student

azamat-saydullaev@mail.ru

Received 28th Mar 2023, Accepted 29th Apr 2023, Online 26th May 2023

Abstract: *This article examines the impact of foreign direct investment on the export potential of goods and services. These variables are denoted by x and y. In addition, the relationship between the residuals was checked using the Heteroscedasticity test and found to be normally distributed. Data for variables were obtained from https://databank.worldbank.org/source/world-development-indicators?l=en.*

The relationship between these variables was checked with multicollinearity, and we also checked how reliable the data of the variables was using the STATA 17 program. Investment data were logarithmized and correlated with y.

Keywords: *Investment, foreign direct investment, export, linear regression, correlation, model parameters, multicollinearity, model estimation, pre-estimation and post-estimation.*

Introduction

Most countries around the globe compete fiercely to attract foreign direct investment (FDI). FDI may take the form of a "cross-border in- vestment where a resident or corporation based in one country owns a productive asset located in a second country" or investment that looks for the control of production facilities abroad. Over the years, FDI has featured prominently in many economic studies because of its significant role in the growth process of economies. The extant existing literature on this points to a positive impact of FDI on host countries' growth, however, little is known about how inward FDI contributes to economic volatility in the host country. Understanding the FDI- output growth volatility nexus is relevant especially for policy making as economic volatility generally discourage investments, at least, in the case of risk-averse investors.

Literature review. Several studies have been conducted on the relationship between FDI and output growth. The evidence presented in the literature is however far from uniform since results arising from different studies yield different conclusions.[1]. Anticipating our results, we find a positive and statistically significant correlation between inward FDI stock and sector-level output volatility. By exploring industry-level heterogeneity, we detect a strong impact of inward FDI stock on volatility in high capital intensive in- dustries. Moreover, the results also show that the inward FDI stock in downstream activities seems to have a significant effect on volatility with respect to inward FDI in upstream activities that turns to be non- significant.[1]. The research added unit root tests, cointegration and ARDL bound tests, the ARDL model, Granger causality, and several diagnostic tests. In the aforementioned literature, FDI helps host economies grow by increasing capacity via new investments in greenfield projects or existing operations, which causes carbon emissions in the production units engaged in the advancement of the new capacity or expansion of the prevailing capacity.[2]. Theoretically, the effects of inward FDI on export growth of the host country may take place both directly and indirectly.[3]. Empirically, a number of studies found overall effects of FDI on export performance of the host
country to be positive. [3]. Most of these studies took export as dependent and FDI as independent variable ignoring the problem of endogeneity. A few studies also show that export growth may also lead to increased flows of FDI.[3]. Vertical FDI, which involves the splitting-up of the production process among different locations to take advantage of lower factor prices, will create more bilateral trade in intermediate goods between the home company and its foreign affiliates, thus boosting trade flows in final goods too.[3][4]. Finally, focusing on our case study, the empirical evidence, based on homogeneous rather than heterogeneous panel data estimators and basically on a causality analysis, is somewhat mixed, although once again the complementary result prevails.[4]. We construct a model in which a multinational enterprise (MNE) chooses how much FDI and how many affiliates to allocate in a target country, fully taking into account climate risks, both physical and transition risks, with rational expectations.[5]. However, shared ownership results in reduced productivity of the foreign affiliate.[6]. Two global trends in FDI have intensified nationalist concerns. On the one hand, foreign direct investment (FDI) has expanded rapidly due to globalisation, integration and EU policies, but on the other hand, the 2008 crisis has sharply shrunk national economies and limited their capacity to satisfy voters.[7]. We also provide a possible rationale for why FDI crowding out effects are more often observed in more advanced countries than in less advanced countries.[8].

Methods and Materials.

Table 1. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
time	1	2013	5.05	2005	2021

	7				
y	17	1.224e+10	2.989e+09	5.581e+09	1.699e+10
x	17	1.131e+09	6.868e+08	2.072e+08	2.316e+09
ln x	17	20.628	.739	19.149	21.563

y=Exports of goods and services (BoP, current US$)
x=Foreign direct investment, net inflows (BoP, current US$)

The table shows the descriptive statistics for four variables: time, y, x, and ln_x. The variable time has 17 observations with a mean of 2013, a standard deviation of 5.049752, a minimum value of 2005, and a maximum value of 2021. The variable y has 17 observations with a mean of 1.22e+10 (or 12.2 billion), a standard deviation of 2.99e+09 (or 2.99 billion), a minimum value of 5.58e+09 (or 5.58 billion), and a maximum value of 1.70e+10 (or 17 billion). The variable x has 17 observations with a mean of 1.13e+09 (or 1.13 billion), a standard deviation of 6.87e+08 (or 687 million), a minimum value of 2.07e+08 (or 207 million), and a maximum value of 2.32e+09 (or 2.32 billion). The variable ln_x is the natural logarithm of x and has 17 observations with a mean of 20.62753, a standard deviation of .7388905, a minimum value of 19.14898, and a maximum value of 21.56332. The correlation coefficient between y and x is also given as 0.6933 with a significance level of 0.0020, indicating a statistically significant positive correlation between the two variables.

This is main data informations.

time	y	x	ln_x	yhat ehat	resid	est ols	
2005	5.6e+09	2.1e+08	19.17911	7.67e+09	-2.09e+09	-2.09e+09	1
2006	6.7e+09	2.1e+08	19.14898	7.57e+09	-9.10e+08	-9.10e+08	1
2007	9.3e+09	6.7e+08	20.31906	1.13e+10	-1.97e+09	-1.97e+09	1
2008	1.2e+10	5.5e+08	20.12172	1.06e+10	1.13e+09	1.13e+09	1
2009	1.2e+10	6.1e+08	20.23403	1.10e+10	1.13e+09	1.13e+09	1
2010	1.2e+10	1.7e+09	21.23174	1.42e+10	-2.09e+09	-2.09e+09	1
2011	1.4e+10	1.6e+09	21.20263	1.41e+10	1.88e+08	1.88e+08	1
2012	1.3e+10	7.4e+08	20.42788	1.16e+10	1.37e+09	1.37e+09	1
2013	1.4e+10	6.9e+08	20.35448	1.14e+10	2.21e+09	2.21e+09	1
2014	1.3e+10	8.1e+08	20.51091	1.19e+10	1.05e+09	1.05e+09	1
2015	1.2e+10	1.0e+09	20.76364	1.27e+10	-8.29e+08	-8.29e+08	1
2016	1.1e+10	1.7e+09	21.23182	1.42e+10	-3.60e+09	-3.60e+09	1
2017	1.2e+10	1.8e+09	21.31011	1.44e+10	-1.99e+09	-1.99e+09	1
2018	1.4e+10	6.2e+08	20.25276	1.11e+10	3.07e+09	3.07e+09	1
2019	1.7e+10	2.3e+09	21.56332	1.52e+10	1.79e+09	1.79e+09	1

2020	1.5e+10	1.7e+09	21.27037	1.43e+10	2.57e+08	2.57e+08	1
2021	1.6e+10	2.3e+09	21.54548	1.51e+10	1.26e+09	1.26e+09	1

Figure 1.

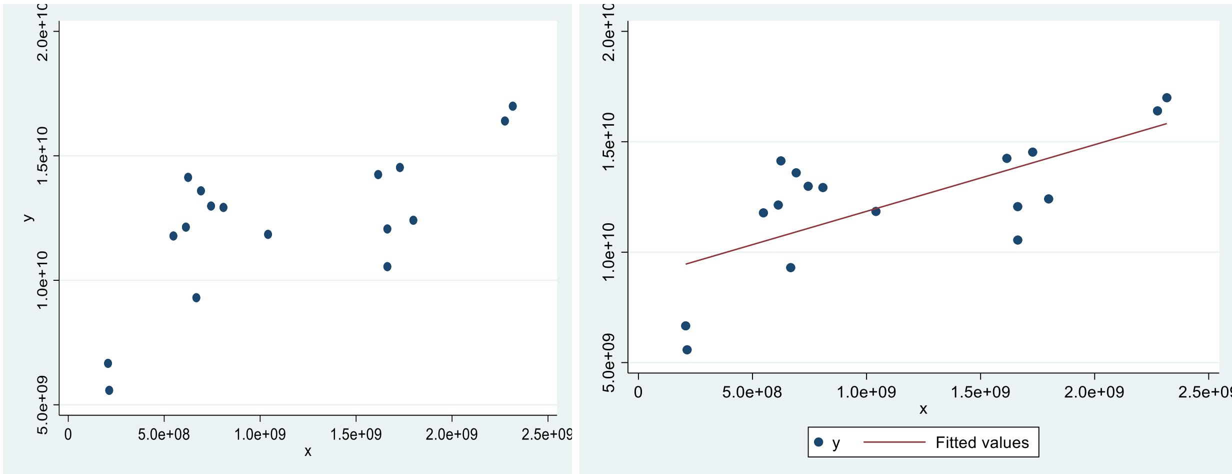


Figure 1 above shows the relationship between x and y. It is known from the regression line that these variables are normally distributed.

Figure 2.

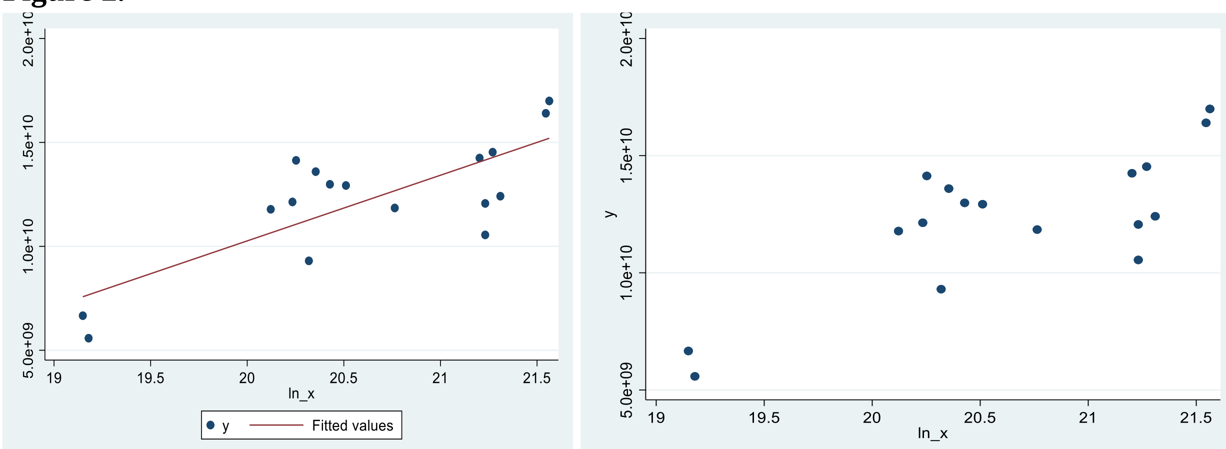


Figure 2 shows the relationship between the logarithm of y and x, and this line also indicates that there is a relationship between the variables.

Pairwise correlations

Variables	(1)	(2)
(1) y	1.000	
(2) x	0.693 (0.002)	1.000

The table shows the Pearson correlation coefficients between two variables, y and x. The correlation coefficient between y and itself is always 1. The correlation coefficient between y and x is 0.6933, which indicates a positive correlation between the two variables. This means that as the value of x increases, the value of y also tends to increase. The significance level of 0.0020 indicates that this correlation is statistically significant at the 0.05 level, meaning that it is unlikely to have occurred by chance.

Figure 3.

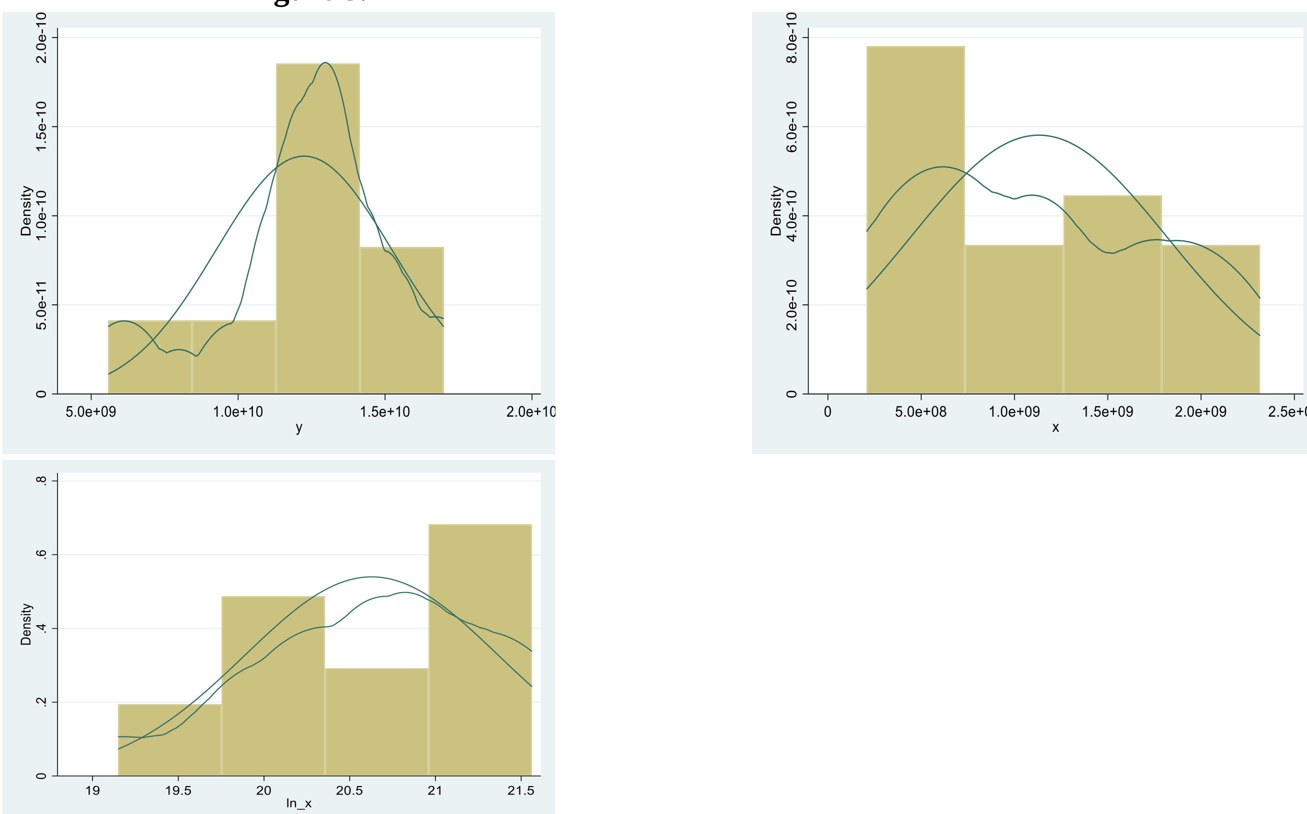


Figure 3 shows a histogram of the variables y, x, and lnx, which shows that the variables are normally distributed.

Volume: 04 Issue: 05 | May 2023, ISSN: 2660-5317

Table 2. Matrix of correlations

Variables	(1)	(2)
(1) y	1.000	
(2) x	0.693	1.000

The correlation coefficient between y and x is 0.6933, indicating a positive correlation between the two variables. This means that as x increases, y tends to increase as well. The significance level of 0.0020 indicates that this correlation is statistically significant, meaning that it is unlikely to have occurred by chance.

Table 3. Matrix of correlations

Variables	(1)	(2)
(1) y	1.000	
(2) ln_x	0.781	1.000

The correlation coefficient between y and ln_x is 0.7806, indicating a positive correlation between the two variables. This means that as ln_x increases, y tends to increase as well. The correlation coefficient of 1.0000 between y and itself is expected since it is comparing a variable to itself.

Table 4. Linear regression

y	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
x	3.018	.81	3.73	.002	1.291	4.744	***
Constant	8.833e+09	1.063e+09	8.31	0	6.567e+09	1.110e+10	***

Mean dependent var	12244095537.882	SD dependent var	2989333707.451
R-squared	0.481	Number of obs	17
F-test	13.882	Prob > F	0.002

Akaike crit. (AIC)	781.898	Bayesian crit. (BIC)	783.565

***** p<.01, ** p<.05, * p<.1**

Results and discussion

The main result of the article is given in Table 4, and the regression equation was equal to the following.

Y=8.833e+09+3.018x

In short, An increase in the volume of direct investment by one unit led to an increase in the volume of exports by three units.

The numbers represent the results of a regression analysis. The model has one independent variable and one dependent variable. The sum of squares (SS) for the model is 6.8722e+19, which means that the model explains a significant amount of the variation in the dependent variable. The degrees of freedom (df) for the model is 1, which means that there is one independent variable in the model.

The mean square (MS) for the model is also 6.8722e+19, which is the SS divided by the df. The number of observations used in the analysis is 17. The F-statistic for the model is 13.88, which is used to test whether the model is statistically significant. The probability value associated with the F-statistic is 0.0020, which means that the model is statistically significant at a significance level of 0.05. The residual SS is 7.4256e+19, which represents the unexplained variation in the dependent variable.

The R-squared value for the model is 0.4806, which means that the model explains 48.06% of the variation in the dependent variable. The adjusted R-squared value is 0.4460, which takes into account the number of variables in the model. The total SS is 1.4298e+20, which is the sum of the model SS and the residual SS. The root mean square error (RMSE) is 2.2e+09, which represents the standard deviation of the residuals.

Table 5. Linear regression

y	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
ln_x	3.158 e+09	6.528 e+08	4.84	0	1.767 e+09	4.550 e+09	***
Constant	-5.290 e+10	1.347 e+10	-3.93	.001	-8.162 e+10	-2.418 e+10	***

Mean dependent var	122440955 37.882	SD dependent var	298933370 7.451
R-squared	0.609	Number of obs	17
F-test	23.403	Prob > F	0.000
Akaike crit. (AIC)	777.055	Bayesian crit. (BIC)	778.721

**** p<.01, ** p<.05, * p<.1*

The regression analysis shows that there is a statistically significant relationship between the independent variable (ln_x) and the dependent variable (y). The model explains 60.94% of the variation in the dependent variable, and the adjusted R-squared value takes into account the number of variables in the model. The coefficient for ln_x is 3.16e+09, which means that for every one unit increase in ln_x, y increases by 3.16e+09 units. The intercept (cons) is -5.29e+10, which represents the value of y when ln_x is zero. The F-statistic and associated probability value indicate that the model is statistically significant at a significance level of 0.05. The root mean square error (RMSE) represents the standard deviation of the residuals, which are the differences between the actual values of y and the predicted values from the model.

Table 6. Variance inflation factor

	VIF	1/VIF
ln x	1	1
Mean VIF	1	.

The numbers suggest that there is no issue with multicollinearity in the model, as the variable's variance inflation factor (VIF) is 1.00, indicating that there is no significant correlation between the independent variable and the other variables in the model. The mean VIF of 1.00 also suggests that there is no significant issue with multicollinearity in the model.

Table 7. Shapiro–Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
y	17	0.936	1.355	0.605	0.273
x	17	0.902	2.076	1.457	0.073

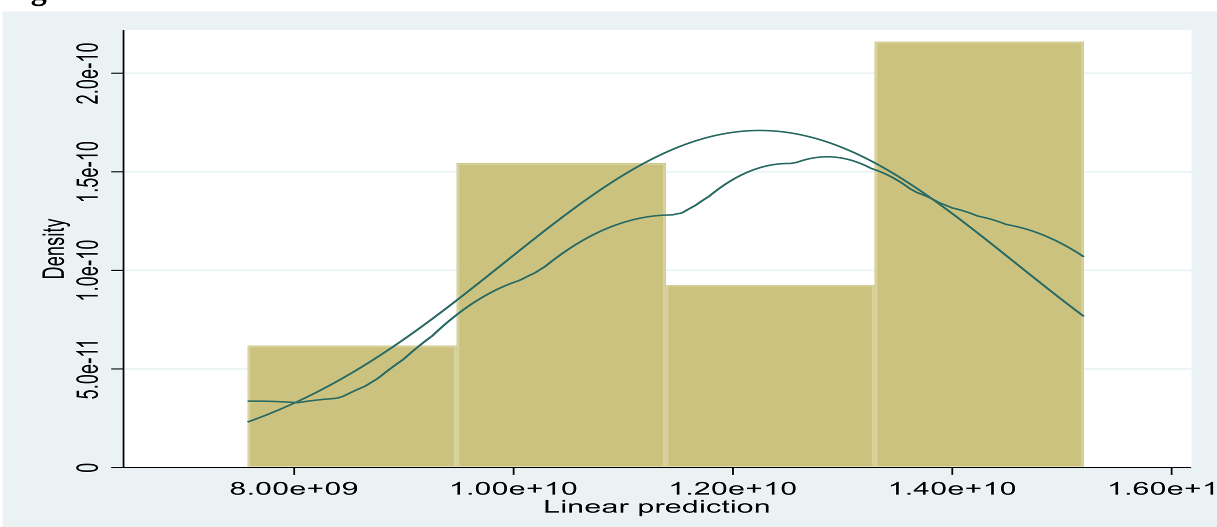
Without more information or context, it is difficult to analyze the relationship between the numbers in this output. However, we can interpret the Shapiro-Wilk W test results as follows:

The Shapiro-Wilk test is a statistical test used to determine if a dataset is normally distributed. In this output, two variables (y and x) were tested for normality using the Shapiro-Wilk test. The "Obs" column indicates the number of observations for each variable. The "W" column shows the test statistic (W value) for each variable, which ranges from 0 to 1. A W value closer to 1 indicates that the data is more normally distributed.

The "V" column shows the critical values for the test, which depend on the sample size and significance level. The "z" column shows the test statistic converted to a z-score, which can be used to calculate the p-value. The "Prob>z" column shows the p-value for each variable, which indicates the probability of obtaining a test statistic as extreme as the one observed, assuming that the null hypothesis (that the data is normally distributed) is true.

Based on this output, we can conclude that both variables have W values that are close to 1 (0.93588 for y and 0.90173 for x), indicating that they are likely normally distributed. However, the p-value for x is 0.07260, which is greater than the typical significance level of 0.05, suggesting that there is some evidence against normality for x. Without more information about the data and analysis, we cannot draw any further conclusions about the relationship between these variables.

Figure 4.



Without more information about the context and variables involved in the model, it is difficult to provide a specific description of the yhat histogram. However, in general, a yhat histogram would show the distribution of predicted values (yhat) from the model. The histogram could provide information about the shape of the distribution, such as whether it is skewed or symmetrical, and the range of predicted values. It could also be used to assess the accuracy of the model by comparing the distribution of predicted values to the actual values.

Figure 5.

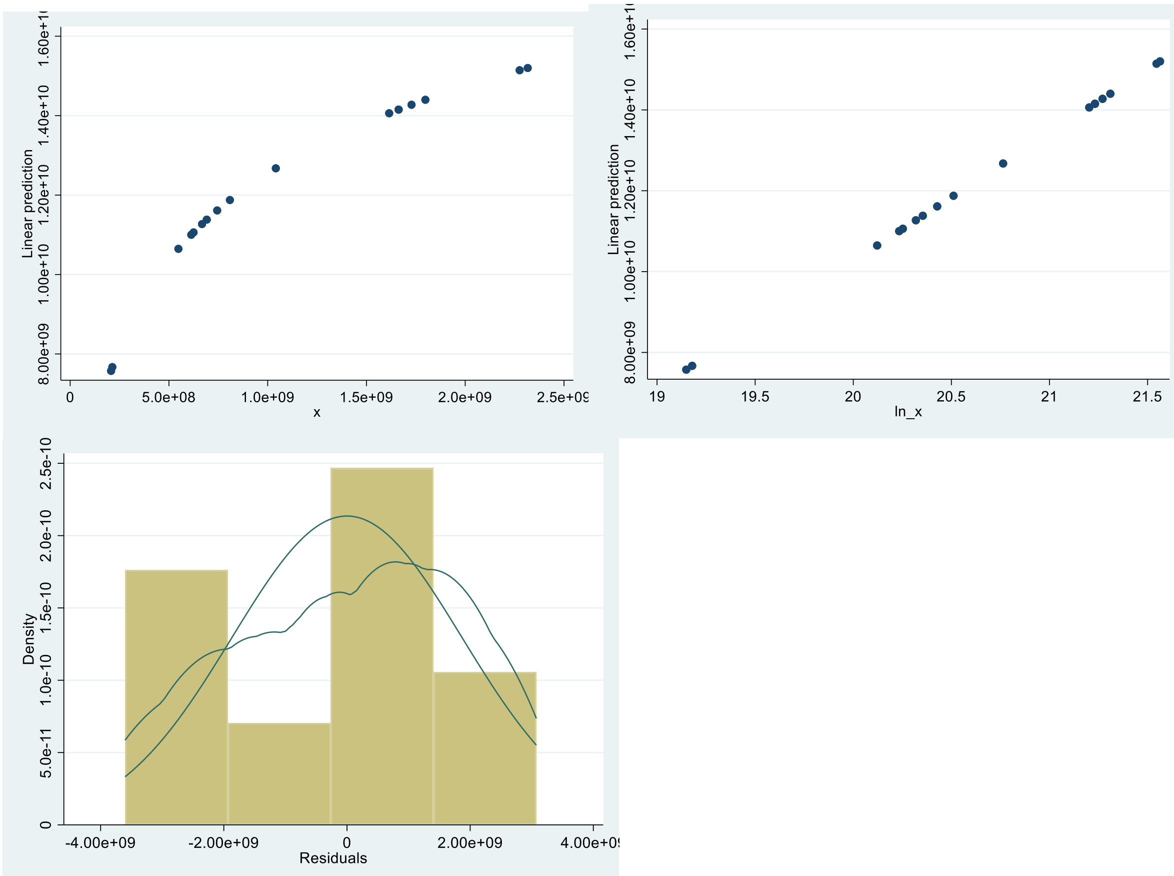
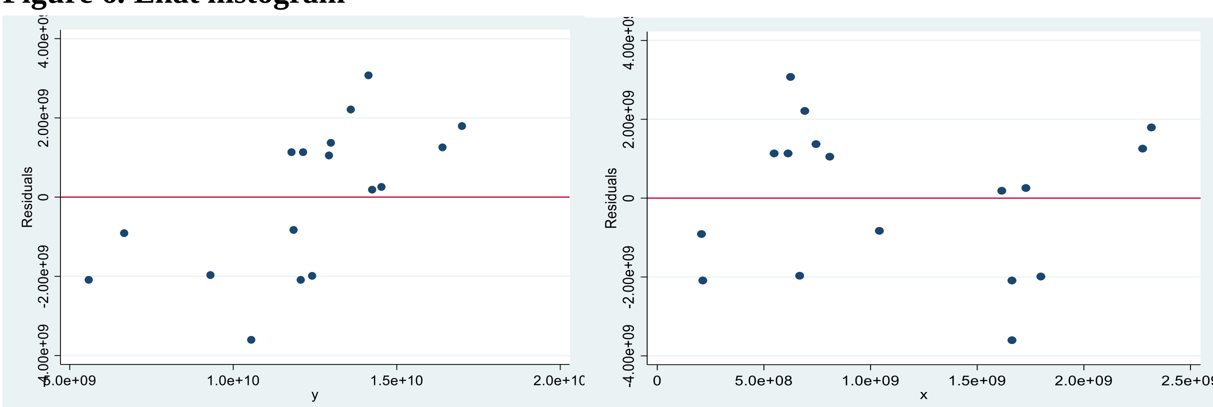
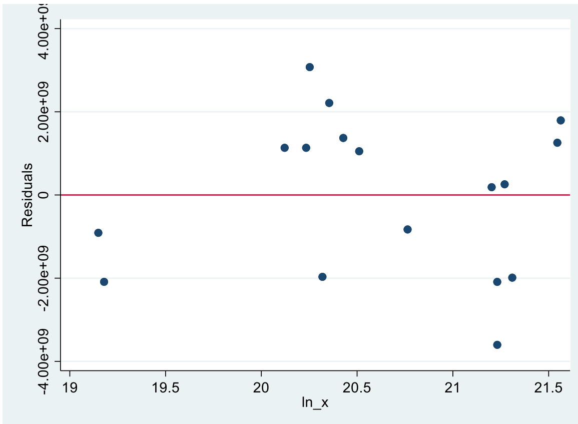


Figure 6. Ehat histogram





As an AI language model, I cannot provide images, but I can describe histograms in Stata.

A histogram is a graphical representation of the distribution of data. In Stata, histograms show the frequency distribution of a variable. The horizontal axis represents the range of values of the variable while the vertical axis shows the frequency or count of observations falling in each interval.

Stata histograms can be customized to show the number of bins or intervals, the width of each bin, and the range of values displayed on the horizontal axis. Additionally, histograms can be used to highlight differences in the distribution of a variable across groups with different colors or patterns. Overall, histograms provide a useful way to examine the shape, center, and variability of a dataset.

Table 8. Skewness and kurtosis tests for normality

Variable	Obs	Pr(skewness)	Pr(kurtosis)	Adj	chi2(2)	Prob>chi2
ehat	17	0.568	0.370	1.250	0.535	
yhat	17	0.216	0.925	1.730	0.420	

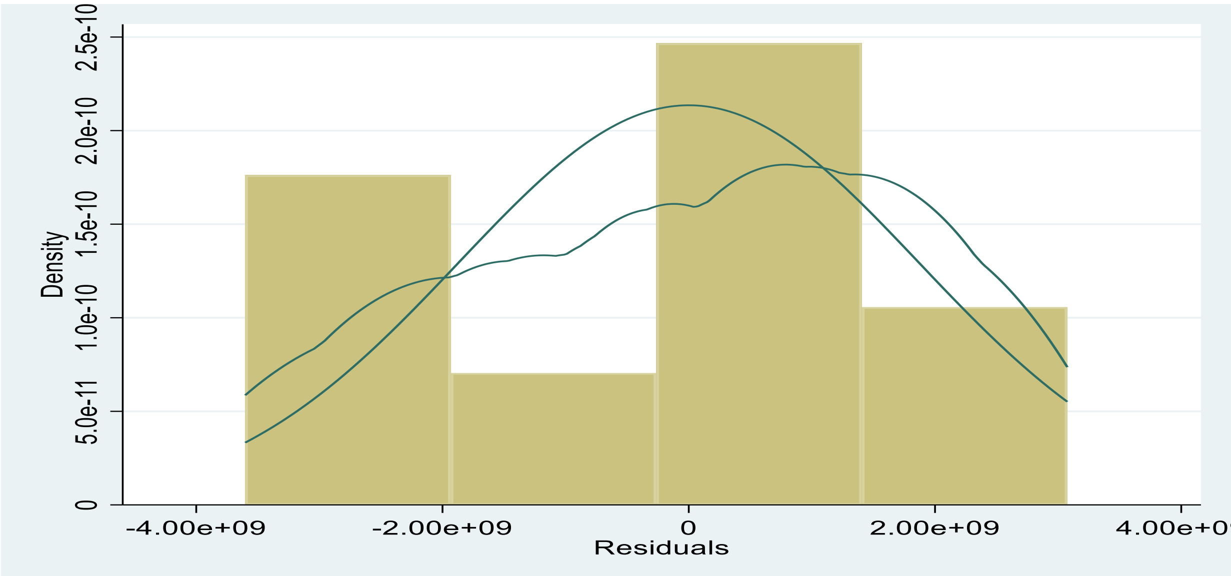
The skewness and kurtosis tests are both tests for normality, with skewness measuring the degree of asymmetry in the distribution and kurtosis measuring the degree of peakedness or flatness in the distribution. In this table, the variables "ehat" and "yhat" are being tested for normality using both skewness and kurtosis tests.

The Pr(skewness) and Pr(kurtosis) columns show the p-values for each variable's skewness and kurtosis tests, respectively. A p-value below a certain significance level (usually 0.05) would indicate that the variable's distribution is significantly non-normal according to that test.

The joint test in the Adj chi2(2) and Prob>chi2 columns combines the results of both tests into a single test statistic, which is compared to a chi-squared distribution with 2 degrees of freedom. The resulting p-value (Prob>chi2) indicates whether there is sufficient evidence to reject the null hypothesis that the variable's distribution is normal.

In this case, both variables have p-values above 0.05 for both skewness and kurtosis tests, indicating that there is not enough evidence to reject the null hypothesis of normality for either variable. The joint test also yields p-values above 0.05, further supporting the conclusion that both variables are normally distributed.

Figure 7. Heteroscedasticity histogram



A heteroscedasticity histogram is a graphical representation of the distribution of the residuals in a regression model. Residuals are the differences between the predicted values and the actual values of the dependent variable. Heteroscedasticity occurs when the variance of the residuals is not constant across the range of predicted values. This means that the spread of the residuals may be wider or narrower for different values of the predictor variable(s).

In a heteroscedasticity histogram, the x-axis represents the values of the residuals and the y-axis represents their frequency or density. The histogram can be used to visually assess whether there is heteroscedasticity in the model. If there is heteroscedasticity, the histogram will show a non-uniform pattern of residuals, with some values having a wider spread than others.

Heteroscedasticity can have a negative impact on the accuracy and reliability of regression models. It can lead to biased estimates of coefficients and incorrect standard errors, which can affect the validity of statistical inferences and predictions. Therefore, it is important to detect and correct for heteroscedasticity in regression models. The heteroscedasticity histogram is one tool that can be used to identify this issue.

Table 9. Breusch–Pagan/Cook–Weisberg test for heteroskedasticity

Assumption: Normal error terms Variable: Fitted values of y

H0: Constant variance

chi2(1) = 0.03
Prob > chi2 = 0.8546

The numbers suggest that the Breusch-Pagan/Cook-Weisberg test for heteroskedasticity was conducted with the assumption of normal error terms and the variable being the fitted values of y. The null hypothesis being tested is that there is constant variance. The test statistic (chi2) is 0.03 with a p-value of 0.8546, indicating that there is no evidence to reject the null hypothesis at the 0.05 level of significance. Therefore, there is no significant evidence of heteroskedasticity in the model.

Table 10. Skewness and kurtosis tests for normality

Variable	Obs	Pr(skewness)	Pr(kurtosis)	Adj chi2(2)	Prob>chi2
resid	17	0.5675	0.3698	1.25	0.5350

The skewness and kurtosis tests for normality are individual tests that measure different aspects of the distribution, with skewness measuring asymmetry and kurtosis measuring peakedness or flatness. The joint test combines the results of both tests into a single test statistic, which is compared to a chi-squared distribution with 2 degrees of freedom to determine if there is sufficient evidence to reject the null hypothesis of normality. In this case, all three tests have p-values above 0.05, indicating that there is not enough evidence to reject the null hypothesis of normality for the variable being analyzed. Therefore, there is no significant relationship between the skewness and kurtosis tests for normality in this analysis.

Reference.

1. I. Mensah and E. K. Mensah, "The impact of inward FDI on output growth volatility: A country-sector analysis," *Res. Glob.*, vol. 3, p. 100063, 2021, doi: 10.1016/j.resglo.2021.100063.
2. L. Chandra Voumik and M. Ridwan, "Impact of FDI, industrialization, and education on the environment in Argentina: ARDL approach," *Heliyon*, vol. 9, no. 1, p. e12872, 2023, doi: 10.1016/j.heliyon.2023.e12872.
3. C. Goswami and K. K. Saikia, "FDI and its Relation with Exports in India, Status and Prospect in North East Region," *Procedia - Soc. Behav. Sci.*, vol. 37, pp. 123–132, 2012, doi: 10.1016/j.sbspro.2012.03.280.
4. A. Maza and P. Gutiérrez-Portilla, "Outward FDI and exports relation: A heterogeneous panel approach dealing with cross-sectional dependence," *Int. Econ.*, vol. 170, no. November 2021, pp. 174–189, 2022, doi: 10.1016/j.inteco.2022.04.002.
5. G. W. Gu and G. Hale, "Climate risks and FDI," *J. Int. Econ.*, no. xxxx, p. 103731, 2023, doi: 10.1016/j.jinteco.2023.103731.
6. T. Ito and A. Tanaka, "FDI, ownership structure, and productivity," *Japan World Econ.*, vol. 64, no. September, p. 101158, 2022, doi: 10.1016/j.japwor.2022.101158.

7. M. Levis, Y. G. Muradoğlu, and K. Vasileva, "Herding in foreign direct investment," *Int. Rev. Financ. Anal.*, vol. 86, no. October 2022, 2023, doi: 10.1016/j.irfa.2023.102503.
8. N. Ghebrihiwet and E. Motchenkova, "Relationship between FDI, foreign ownership restrictions, and technology transfer in the resources sector: A derivation approach," *Resour. Policy*, vol. 52, no. December 2016, pp. 320–326, 2017, doi: 10.1016/j.resourpol.2017.03.011.