The statistical analysis and forecast of the FDI inflow in Kyrgyzstan (based on the data in 2021 - 2023)

© 2023 by Aleksey Lim, Emirlan Torokulov is licensed under CC BY-NC-ND 4.0

The Theory of Probabilities and Mathematical Statistics

Pr. Polina Dolmatova

AUCA, Fall 2023

MAT-307

Contents:

Project topic

Introduction

World economies rely on foreign direct investment which refers to entities that are invested within certain states to other people's business and property outside their location. However, FDI is characterized as a much higher level of involvement and considerable control over foreign enterprise. In FDI many resources are transferred – finance, capital, technology, skills labor, and a good market for products. Generally speaking, this type of transfer is beneficial to both the investor's country and the recipient's country. The long-term gains appeal to investors as they anticipate developments in foreign economies plus the development of long-term relationships.

Formulation of the real-life problem

Therefore, since inflows of FDI influence the development of the economy in the country, it is very important to study this impact within Kyrgyzstan. As a developing country, Kyrgyzstan has a lot to gain from international investment. Examining FDI flows is a means of revealing diversification prospects that could lessen sector dependence. In addition, the introduction of FDI into the economy usually introduces advanced technology and management techniques that improve a country's industrial capacity. Furthermore, the analysis of FDI explains how it affects employment, offering a comprehensive picture of the impact of foreign investments on employment rates and wage levels in different countries. Additionally, FDI helps the country's foreign exchange reserves towards their balances of payment and economic stability. Results of FDI studies have relevance in the policy-making process where policymakers can use these findings to come up with policies that create an attractive environment for foreign investors and minimize the risk while fostering trust from investors. The study of FDI in Kyrgyzstan is a practical undertaking with long-run impacts providing a blueprint of durable economic growth and outreach into the world economy.

Mathematical model of the problem - "Multiple Linear Regression Model for FDI Inflow

$$FDI_t = \beta_0 + \beta_1 \cdot GDP_t + \beta_2 \cdot INF_t + \beta_3 \cdot EX_t + \beta_4 \cdot INT_t + \beta_5 \cdot POL_t + \varepsilon_t$$

Where:

- FDI_t is the FDI inflow in Kyrgyzstan at time t.
- GDP_t represents the Gross Domestic Product of Kyrgyzstan at time t.
- INF_t reflects the inflation rate in Kyrgyzstan at time t.
- EX_t denotes the exchange rate of the local currency to major foreign currencies at time t.
- INT_t is the prevailing interest rate in Kyrgyzstan at time t.
- POL_t is a policy variable accounting for the ease of doing business, regulatory environment, etc.
- β_0 is the intercept term representing the base level of FDI when all other variables are zero.
- $\beta_1, \beta_2, \ldots, \beta_5$ are coefficients indicating the impact of each respective variable on FDI.
- ε_t is the error term, encompassing unobserved factors influencing FDI that are not included in the model.

in Kyrgyzstan.»

The assumption behind this model is that there is a positive correlation between FDI and the specified variables. Historical data will be used for the estimation of coefficients via statistical

methods such as simple or multiple regressions. The objective is to explore the linkages of essential macroeconomic variables with the amount of foreign direct investment for the Kyrgyz Republic. If required, additional specifications might be needed depending on the nature and quality of the information.

Evidence on Population and Sampling Method Decision

In our project, we are focusing on the most significant investments, such as agriculture, mining, manufacturing, construction, transport, information and communication, education, and last but not least healthcare. Generally, we will look

at the statistics of each sector from 2021-2023 and identify

common trends.

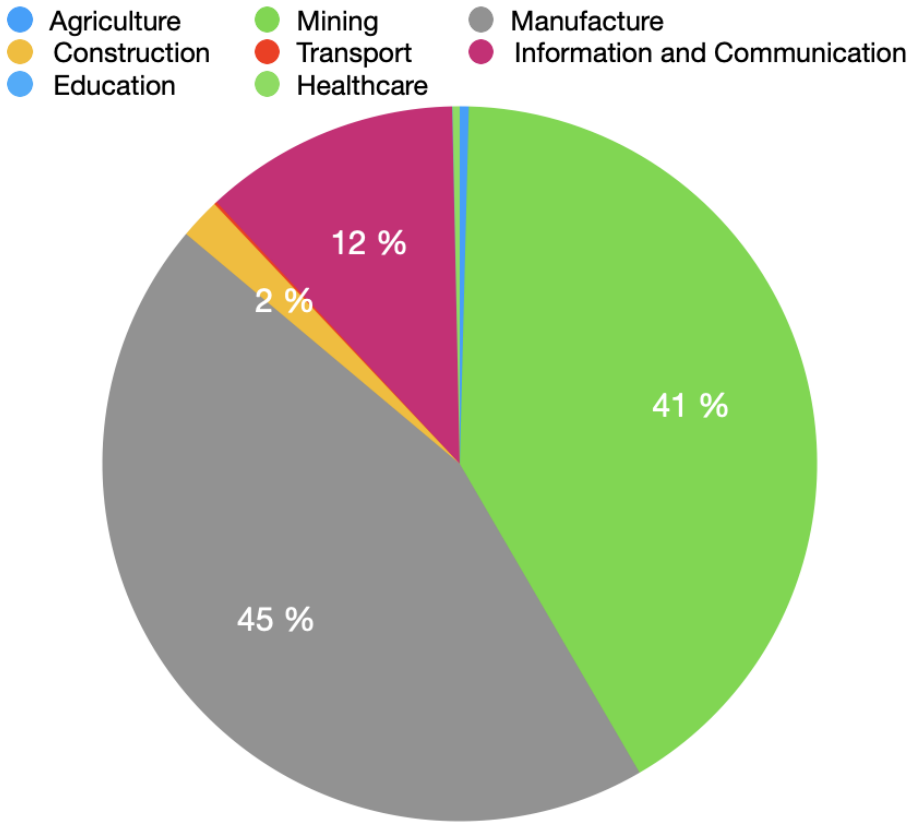
Description of collecting data procedure

When undertaking substantial projects, it is crucial to rely solely on current, official sources of data. For our undertaking, we exclusively sourced all information from the official website of Kyrgyzstan's statistics - stat.kg and statista.com

Representation of data using reasonable graphs, charts

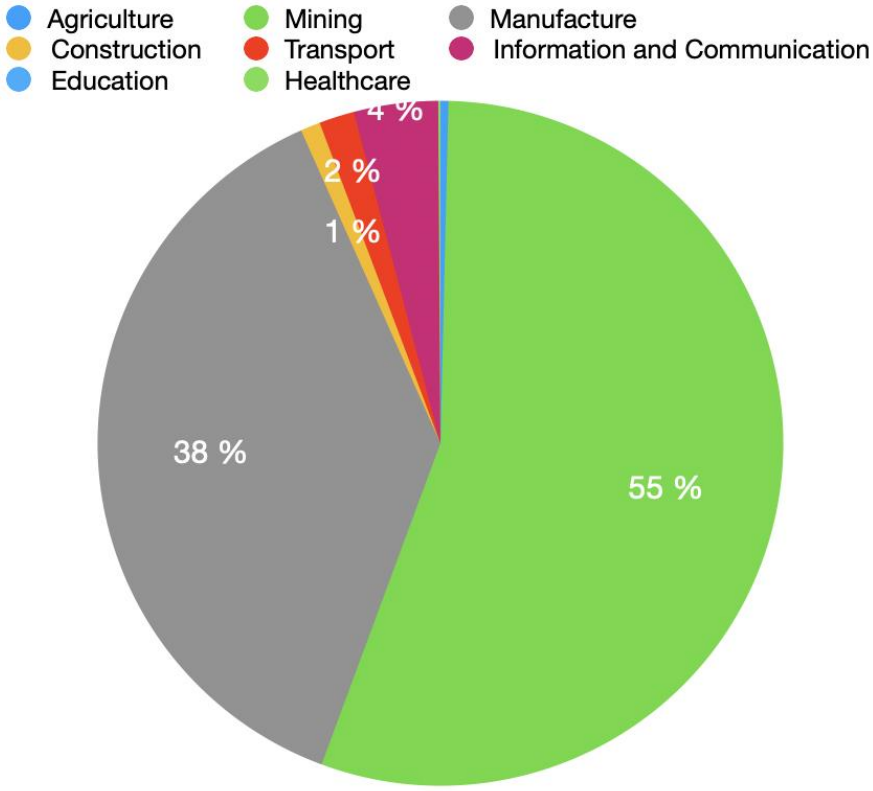
FDI inflow in Kyrgyzstan in 2021

Agriculture	2 092,7
Mining	211 904,6
Manufacture	228945,8
Construction	9456
Transport	656,4
Information and Communication	59 559,1
Education	76,5
Healthcare	1 672,3



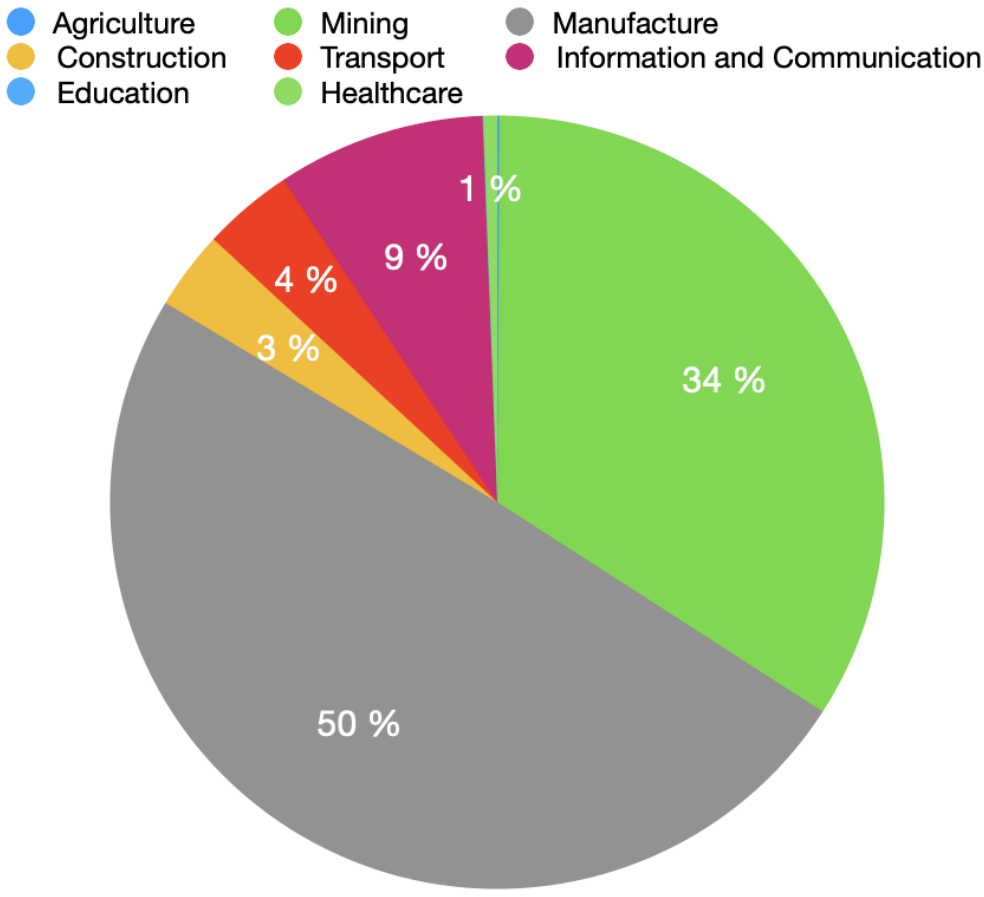
FDI inflow in Kyrgyzstan in 2022

Agriculture	3 358,2
Mining	499 785,2
Manufacture	340 881,9
Construction	8 430,1
Transport	15 068,2
Information and Communication	35 842,2
Education	14,6
Healthcare	923,3



FDI inflow in Kyrgyzstan from jan-jun in 2023

Agriculture	245,1	
Mining	89 489,8	
Manufacture	130 253	
Construction	8 654,7	
Transport	9 955,3	
Information and Communication	22 882,7	
Education	74,4	
Healthcare	1 528,7	



FDI inflow in Kyrgyzstan

	2021	2022	Difference in comparison of both years
Agricultural	2 092,7	3 358	Increase in 1 266
Mining	211 914,6	499 758,2	Increase in 287 844
Manufacture	228 945,8	340 881,9	Increase in 111 936
Construction	9456	8 430,1	Decrease in 1026
Transport	656,4	15 068,2	Increase in 14 412
Information and communication	59 599,1	35 842,2	Decrease in 23 758
Education	76,5	14,6	Decrease in 62
Healthcare	1 672	923,3	Decreas in 749

Solution of the problem

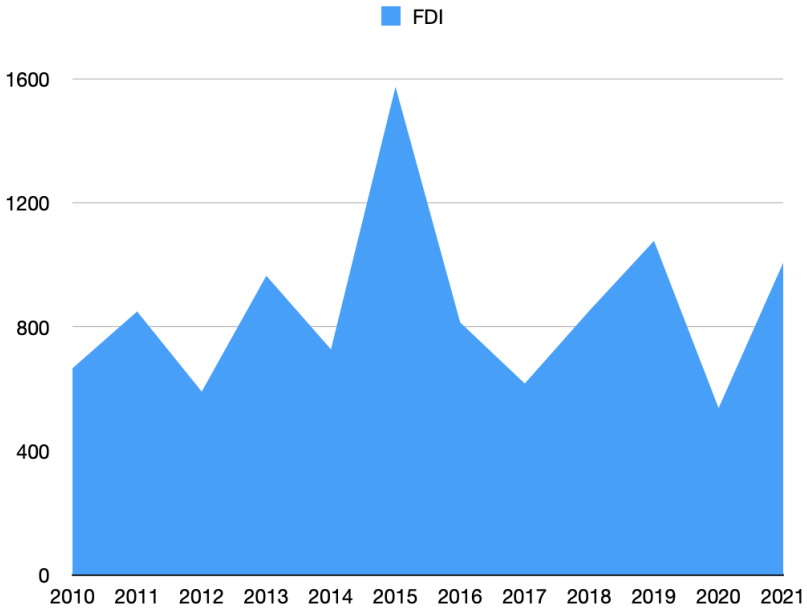
After analyzing all the data, we can conclude that FDI in the mining and manufacturing sectors experienced an increased rate from 2021 until 2022. Manufacture increased from 228,945.8 to 340,881.9 while mining went up by 211,904.6 to 499,785.2. First, a rising trend among foreign investors reflects confidence or faith in the sector in which they have an interest – that is, natural resources and manufacturing. Other trends that were observed in various sectors included agriculture, construction, transport, information and communication, education, and healthcare. Increases were observed in agriculture, construction showed moderate growth while various trends of change were found among transport, information and communication, education, and

healthcare sectors respectively. A change of attention from capital towards information and communication resulted in an increment of FDI increasing from 59,559.1 to 35,842.

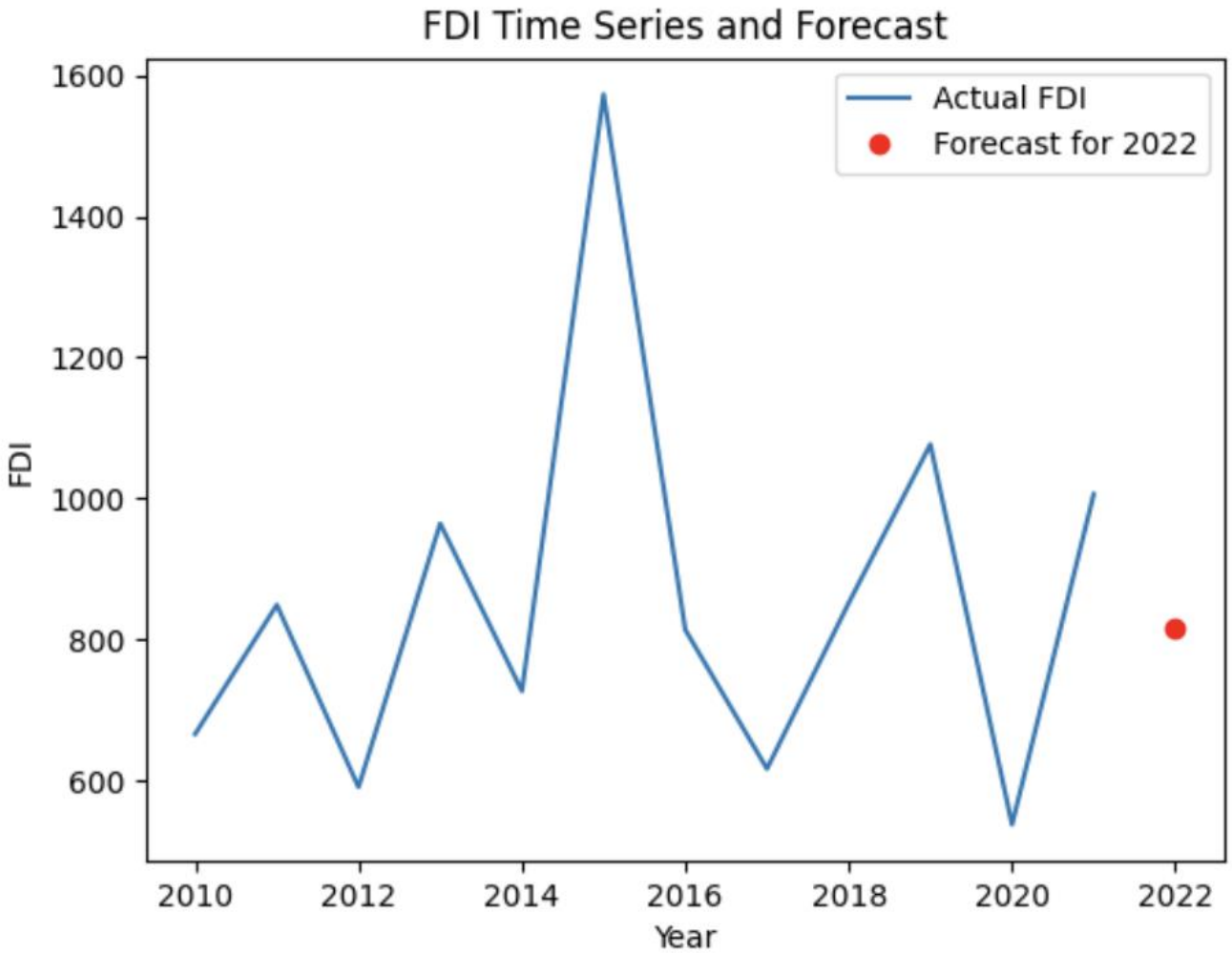
Possible predictions based on the

FDI inflow in Kyrgyzstan

Year	FDI
2010	666,09
2011	849,2
2012	590,73
2013	964,51
2014	727,09
2015	1573,24
2016	813,96
2017	616,79
2018	851,74
2019	1076,92
2020	537,55
2021	1006,09



above solution - ARIMA model



We use Python code to make predictions using ARIMA model:

```
# Import libraries
import warnings
import itertools
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import statsmodels.api as sm

# Suppress warnings
warnings.filterwarnings("ignore",
category=RuntimeWarning)

# Load the data
data = pd.read_csv('fdi_inflow.csv')
data['Year'] = pd.to_datetime(data['Year'], format='%Y')
data.set_index(['Year'], inplace=True)

# Plot the data
data.plot()
plt.ylabel('FDI Inflow (Thousands of Dollars)')
plt.xlabel('Year')
plt.show()

# ARIMA
q = d = range(0, 2)
p = range(0, 4)
pdq = list(itertools.product(p, d, q))
seasonal_pdq = [(x[0], x[1], x[2], 1) for x in
list(itertools.product(p, d, q))]

AIC = []
```

```
SARIMAX_model = []

for param in pdq:
    for param_seasonal in seasonal_pdq:
        try:
            mod = sm.tsa.statespace.SARIMAX(data,
                                            order=param,

seasonal_order=param_seasonal,

enforce_stationarity=False,

enforce_invertibility=False)
            results = mod.fit(disp=0)
            AIC.append(results.aic)
            SARIMAX_model.append([param, param_seasonal])
        except:
            continue

best_model_index = np.argmin(AIC)
best_model_params = SARIMAX_model[best_model_index]
print(f'The best model has AIC={AIC[best_model_index]}
and parameters:
ARIMA{best_model_params[0]}x{best_model_params[1]}')

# Fit the best model
mod = sm.tsa.statespace.SARIMAX(data,

order=best_model_params[0],

seasonal_order=best_model_params[1],

enforce_stationarity=False,
```

```
enforce_invertibility=False)
results = mod.fit(disp=0)

# Forecast
forecast_steps = 5  # You can adjust this based on how
many steps into the future you want to forecast
forecast = results.get_forecast(steps=forecast_steps)
forecast_index = pd.date_range(data.index[-1],
periods=forecast_steps + 1, freq='A')[1:]

# Plot the results
ax = data.plot(figsize=(12, 6))
forecast.predicted_mean.plot(ax=ax, label=f'Forecast
({forecast_steps} steps ahead)')
ax.fill_between(forecast_index,
forecast.conf_int()['lower FDI Inflow'],
forecast.conf_int()['upper FDI Inflow'], color='k',
alpha=.1)
plt.ylabel('FDI Inflow (Thousands of Dollars)')
plt.xlabel('Year')
plt.legend()
plt.show()
```

So our forecast for 2022 was just over 800’000 based on the python code and the actual FDI was around 1’202’599.

Conclusion and practical recommendations

In summary, while the ARIMA model can be a valuable tool for time series forecasting, users must exercise caution and acknowledge its limitations. Employing a holistic approach that considers diverse factors, explores alternative models, and adapts to changing data dynamics is crucial for obtaining reliable and accurate predictions in real-world applications. One limitation of the ARIMA model is its sensitivity to outliers and sudden changes in the data. Abrupt shifts, extreme values, or irregular patterns can significantly impact the model's predictive ability, leading to inaccurate forecasts. Additionally, ARIMA may struggle with non-linear trends or complex dependencies within the data, as it inherently assumes a linear relationship between past observations and future predictions. In practice, it is advisable to complement ARIMA with other forecasting techniques and models, such as machine learning algorithms or hybrid models, to enhance predictive accuracy. Ensemble methods that combine the strengths of different models can mitigate the weaknesses of individual approaches and provide more robust forecasts.

List of references

1) https://www.investopedia.com/terms/f/fdi.asp
2) https://corporatefinanceinstitute.com/resources/data-science/multiple-linear-regression/#
3) https://www.stat.kg/
4) https://www.statista.com/
5) https://loginom.ru/blog/arima-example
6) https://www.linkedin.com/advice/0/what-advantages-disadvantages-using-arima
7) https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/