DATA SCIENCE PROJECT REPORT

Subject title and code: Global Economic Environment (AF2602)

Student Name:

Student ID:

Student Email Address:

Day and time of the tutorial:

Tutor name:

<https://databank.worldbank.org/reports.aspx?source=world-development-indicators>

Section 1. Choose ONE of the following series from WDI as the dependent variable:

* Select “Current account balance (BoP, current US$)” as dependent variable.

Section 2. Choose ANY FOUR series from WDI as independent variables.

1. 'Exports of goods and services (BoP, current US$)',
2. 'Imports of goods, services and primary income (BoP, current US$)',
3. 'Transport services (% of commercial service exports)',
4. 'Net trade in goods and services (BoP, current US$)'

Section 3. Collect the data on the dependent and independent variables of at least 15 countries over a sample period of 10 years. The sample size should be at least 100 country-year observations.

In a table, provide a list of the sample countries, the sample period, the sample size, the sample mean, and the sample standard deviation. (3 marks)



Section 4. Set a hypothesis for the effect of each independent variable on the dependent variable (i.e., a total of four hypotheses). Explanations / Theories / Intuitions should be provided to justify each hypothesis. Sources of references should be cited. (20 marks)

|  |  |  |  |
| --- | --- | --- | --- |
| **Dependent Variable** | **Independent Variable** | **Hypothesis for the effect of each independent variable on the dependent variable** | **Reference** |
| Current account balance (BoP, current US$) | 1.     'Exports of goods and services (BoP, current US$)', | Positive, when export increase, current account balance will increase. | The Formula for Current Account Balance:  CAB=(X−M)+(NY+NCT) where: X=Exports of goods and services M=Imports of goods and services NY=Net income abroad NCT=Net current transfers ​ <https://www.investopedia.com/insights/exploring-current-account-in-balance-of-payments/> |
| Current account balance (BoP, current US$) | 2.     'Imports of goods, services and primary income (BoP, current US$)', | Negative, when import increase, current account balance will decrease. |
| Current account balance (BoP, current US$) | 3.     'Transport services (% of commercial service exports)', | Positive, when export increase, current account balance will increase. |
| Current account balance (BoP, current US$) | 4.     'Net trade in goods and services (BoP, current US$)' | Positive, when net trade increase, current account balance will increase. |

Section 5. Using the data collected, estimate a fixed-effect model by a Python program. Report the main findings in a table. Provide the Python program in an appendix. (5 marks)

A screenshot of a computer

Description automatically generated with medium confidence

* **R-squared = 0.0157 which is low, it is not a good fit model.**

Section 6. Based on the findings, test for the hypotheses one by one. For each hypothesis, intuitively explain why it is / is not supported by the findings. (12 marks)

|  |  |  |  |
| --- | --- | --- | --- |
| **Independent Variable (Full Name)** | **Independent Variable** | **Hypothesis** | **Hypothesis Conclusions** |
| 1.     'Exports of goods and services (BoP, current US$)', | EGS | has a significantly negative effect on "Current account balance" (Beta1 = -0.0007) | p value (0.5931) > 0.05, one cannot reject the null hypothesis of Beta = 0 (at least at the 5% level of significance) |
| 2.     'Imports of goods, services and primary income (BoP, current US$)', | IGS | has a significantly postive effect on "Current account balance" (Beta2 =5.699e-05) | p value (0.9673) > 0.05, one cannot reject the null hypothesis of Beta = 0 (at least at the 5% level of significance) |
| 3.     'Transport services (% of commercial service exports)', | TS | has a significantly negative effect on "Current account balance" (Beta3 = -3.458e+07) | p value (0.3926) > 0.05, one cannot reject the null hypothesis of Beta = 0 (at least at the 5% level of significance) |
| 4.     'Net trade in goods and services (BoP, current US$)' | NT | has a significantly postive effect on "Current account balance" (Beta4 = 0.0252) | p value (0.1681) > 0.05, one cannot reject the null hypothesis of Beta = 0 (at least at the 5% level of significance) |

**Reference**

Current Account Balance, Heakal, <https://www.investopedia.com/insights/exploring-current-account-in-balance-of-payments/>

**Appendix**

from linearmodels import PanelOLS

import statsmodels.api as sm

def Session4(df):

df = df.copy()

df.rename(columns={'Current account balance (BoP, current US$)': 'CAB',

'Exports of goods and services (BoP, current US$)': 'EGS',

'Imports of goods, services and primary income (BoP, current US$)': 'IGS',

'Transport services (% of commercial service exports)': 'TS',

'Net trade in goods and services (BoP, current US$)': 'NT',},

inplace=True)

print(df.columns)

df['Year'] = df['Year'].str[0:4].astype(int)

print(df['Year'].head())

df = df.set\_index(['Country Code','Year'])

mod = PanelOLS(df['CAB'], df[['EGS','IGS','TS','NT']], entity\_effects=True)

res = mod.fit()

print(res)

return res

s4 = Session4(df)