Causal Inference Study of Political Stability on Innovation in Brazil from 2000-2020

Objective: To determine the causal relationship between political stability and the rate of technological and scientific innovation.

Hypothesis: Political stability positively influences the rate of innovation within a country.

```
In [ ]:
```

Causal Pipeline: Step 1 Draw Causal DAG (Directed Acyclic Graph)

- Y: Innovation
- X: Political Stability

Measure of Innovation:

- **Number of Patents Filed:** Annual number of patents filed per capita.
- **Research and Development (R&D) Expenditure:** Total R&D expenditure as a percentage of GDP.
- **Innovation Index:** Composite index measuring innovation capabilities and results (e.g., Global Innovation Index).

Relevant causes for political stability/instability:

- **Political Stability Index:** Quantitative measure of political stability (e.g., World Bank's Political Stability and Absence of Violence/Terrorism Index).
- **Government Effectiveness:** Quality of public services, civil service, and policy implementation.
- **Rule of Law:** Extent to which agents have confidence in and abide by the rules of society.
- **Corruption Perception Index:** Measure of perceived corruption within the public sector.

Control Variables:

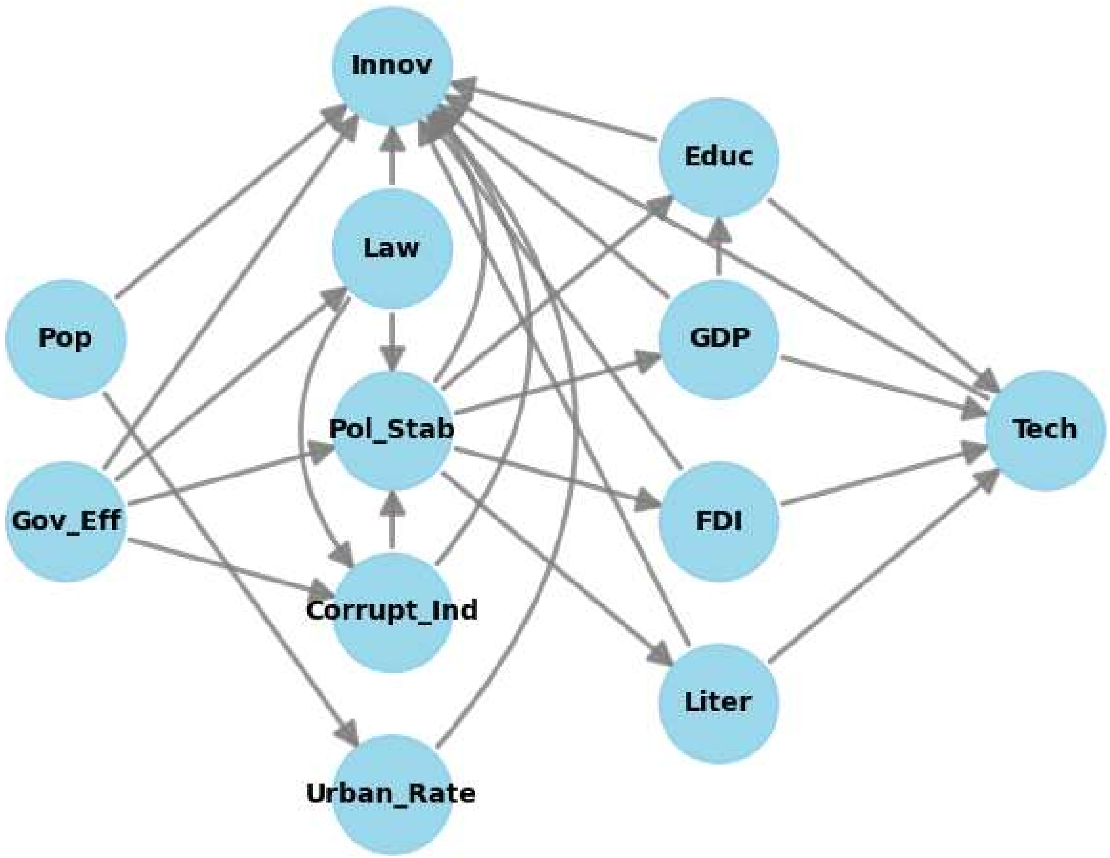
- **GDP per Capita:** Economic performance indicator.

- **Education Level:** Average years of schooling or percentage of population with tertiary education.
- **Foreign Direct Investment (FDI)**: Inflows of foreign investment.
- **Population Size:** Total population to control for market size effects.
- **Urbanization Rate:** Percentage of population living in urban areas.
- **Technological Infrastructure:** Access to technology, such as internet penetration rates.

In [2]:

```
pip install pandas dowhy  matplotlib
```

```
    graph=causal_graph
)
model.view_model()
```



In [4]:
```
import pandas as pd

# Define the legend table as a DataFrame
legend_table = pd.DataFrame({
    'Short Name': ['Pol_Stab', 'Gov_Eff', 'Law', 'Corrupt_Ind', 'GDP', 'Educ','Lite
    'Full Variable Name': [
        'Political Stability',
        'Government Effectiveness',
        'Rule of Law',
        'Corruption Perception Index',
        'GDP per Capita',
        'Education Level',
        'Literacy Rate',
        'Foreign Direct Investment',
        'Population Size',
        'Urbanization Rate',
        'Technological Infrastructure',
        'Innovation'
    ],
    'Description': [
        'The degree to which political processes are stable and predictable',
        'Efficiency of the government in implementing policies and providing servic
        'The extent to which laws are clear, publicized, and stable, and are applie
        'Perception of the level of corruption within the country',
```

```
        'Economic output per person',
        'Average level of education attained by the population',
         'Literacy rate, youth total (% of people 15-24)',
        'Investments made by foreign entities in the country',
        'Total population size',
        'Proportion of people living in urban areas',
        'Infrastructure supporting technological development and use',
        'Capacity to generate and implement new ideas'
    ]
})

# Display the legend table
legend_table
```

Out[4]:

	Short Name	Full Variable Name	Description
0	Pol_Stab	Political Stability	The degree to which political processes are st...
1	Gov_Eff	Government Effectiveness	Efficiency of the government in implementing p...
2	Law	Rule of Law	The extent to which laws are clear, publicized...
3	Corrupt_Ind	Corruption Perception Index	Perception of the level of corruption within t...
4	GDP	GDP per Capita	Economic output per person
5	Educ	Education Level	Average level of education attained by the pop...
6	Liter	Literacy Rate	Literacy rate, youth total (% of people 15-24)
7	FDI	Foreign Direct Investment	Investments made by foreign entities in the co...
8	Pop	Population Size	Total population size
9	Urban_Rate	Urbanization Rate	Proportion of people living in urban areas
10	Tech	Technological Infrastructure	Infrastructure supporting technological develo...
11	Innov	Innovation	Capacity to generate and implement new ideas

Causal Pipeline: Step 2 Identification

In [5]:

```
# Ensure the dataset contains all required columns before proceeding
required_columns = ['Pol_Stab', 'Gov_Eff', 'Law', 'Corrupt_Ind', 'GDP', 'Educ', 'Li
missing_columns = [col for col in required_columns if col not in df_model.columns]

if missing_columns:
    print(f"Missing columns in the dataset: {missing_columns}")
else:
    model = CausalModel(
        data=df_model,
        treatment='Gov_Eff',  # example treatment variable
        outcome='Innov',  # example outcome variable
```

```
        graph=causal_graph
    )

    # Identify the causal effect
    estimands = model.identify_effect(proceed_when_unidentifiable=True)
    print(estimands)
```

```
Estimand type: EstimandType.NONPARAMETRIC_ATE

### Estimand : 1
Estimand name: backdoor
Estimand expression:
    d
──────────(E[Innov])
d[Gov_Eff]
Estimand assumption 1, Unconfoundedness: If U→{Gov_Eff} and U→Innov then P(Innov|Gov
_Eff,,U) = P(Innov|Gov_Eff,)

### Estimand : 2
Estimand name: iv
No such variable(s) found!

### Estimand : 3
Estimand name: frontdoor
No such variable(s) found!
```

This expression represents the expected change in Innovation (Innov) when Government Effectiveness (Gov_Eff) changes.

Assumption: Unconfoundedness

If there are other variables (U) that influence both Government Effectiveness and Innovation, then by controlling for these variables, the relationship between Government Effectiveness and Innovation is not confounded by other factors. Mathematically, this means that the probability of Innovation given Government Effectiveness and these other variables (U) is equal to the probability of Innovation given only Government Effectiveness.

In our study, the goal is to estimate the causal effect of Government Effectiveness (Gov_Eff) on Innovation (Innov). Based on our causal graph and the backdoor criterion, we need to ensure our dataset contains adequate measures for the following variables:

- Political Stability (Pol_Stab): Reflects the degree to which political processes are stable and predictable.
- Government Effectiveness (Gov_Eff): Efficiency of the government in implementing policies and providing services.
- Rule of Law (Law): The extent to which laws are clear, publicized, and stable, and are applied evenly.
- Corruption Perception Index (Corrupt_Ind): Perception of the level of corruption within the country.
- GDP per Capita (GDP): Economic output per person.

- Education Level (Educ): Average level of education attained by the population.
- Foreign Direct Investment (FDI): Investments made by foreign entities in the country.
- Population Size (Pop): Total population size.
- Urbanization Rate (Urban_Rate): Proportion of people living in urban areas.
- Technological Infrastructure (Tech): Infrastructure supporting technological development and use.
- Innovation (Innov): Capacity to generate and implement new ideas.

In [6]:
```
# Create the measures table
measures_table = pd.DataFrame({
    'Variable': [
        'Pol_Stab', 'Gov_Eff', 'Law', 'Corrupt_Ind', 'GDP', 'Educ', 'FDI', 'Pop', '
    ],
    'Description': [
        'An index or score representing political stability',
        'A composite score or index representing government effectiveness',
        'An index representing the rule of law',
        'The Corruption Perception Index score',
        'GDP per capita in constant dollars',
        'Average years of schooling or an education index',
        'Total foreign direct investment inflows',
        'Total population size',
        'Percentage of the population living in urban areas',
        'An index representing technological infrastructure',
        'An innovation index or number of patents filed'
    ]
})

# Display the measures table
measures_table
```

Out[6]:

	Variable	Description
0	Pol_Stab	An index or score representing political stabi...
1	Gov_Eff	A composite score or index representing govern...
2	Law	An index representing the rule of law
3	Corrupt_Ind	The Corruption Perception Index score
4	GDP	GDP per capita in constant dollars
5	Educ	Average years of schooling or an education index
6	FDI	Total foreign direct investment inflows
7	Pop	Total population size
8	Urban_Rate	Percentage of the population living in urban a...
9	Tech	An index representing technological infrastruc...
10	Innov	An innovation index or number of patents filed

Brazil has experienced political instability, economic crises, and corruption scandals. It has comprehensive data available through the Brazilian Institute of Geography and Statistics (IBGE) and other sources.As a result, a country like Brazil is more worthwhile than modern Europe.

Normalization Technique: Standardization (Z-score Normalization)

Methodology

The normalization process employs standardization, also known as Z-score normalization. This technique transforms the data so that each variable has a mean of 0 and a standard deviation of 1. The formula used is:

[z = \frac{(x - \mu)}{\sigma}]

Where:

- (x) is the original data point.
- (\mu) is the mean of the variable.
- (\sigma) is the standard deviation of the variable.
- (z) is the standardized value.

Steps

1. **Calculate the Mean ((\mu))**: Compute the average value of each variable.
2. **Calculate the Standard Deviation ((\sigma))**: Measure the amount of variation or dispersion of the variable.
3. **Transform the Data**: Apply the Z-score formula to each data point, resulting in standardized values.

Purpose and Acceptability

- **Comparability**: Standardization allows variables with different units and scales to be compared directly. This is particularly useful when combining variables like GDP (in dollars) and literacy rates (percentages) into a single analysis.
- **Normalization of Data Range**: By transforming variables to a common scale, standardization ensures that no single variable dominates due to its scale. This is crucial for algorithms that are sensitive to the scale of input data, such as regression models and causal inference techniques.

- **Retention of Data Distribution**: Unlike min-max scaling, which compresses data to a fixed range, Z-score normalization retains the original distribution and outliers, providing a more accurate representation of the data.

In this context, standardization is acceptable and beneficial because it prepares diverse socioeconomic indicators for combined analysis, ensuring each variable contributes equally regardless of its original scale.

Reading in Data

In [8]:
```
!pip install openpyxl
```

```
Collecting openpyxl
  Downloading openpyxl-3.1.5-py2.py3-none-any.whl.metadata (2.5 kB)
Collecting et-xmlfile (from openpyxl)
  Downloading et_xmlfile-1.1.0-py3-none-any.whl.metadata (1.8 kB)
Downloading openpyxl-3.1.5-py2.py3-none-any.whl (250 kB)
   ---------------------------------------- 0.0/250.9 kB ? eta -:--:--
   ------------------- -------------------- 122.9/250.9 kB 3.6 MB/s eta 0:00:01
   ---------------------------------------- 250.9/250.9 kB 5.1 MB/s eta 0:00:00
Downloading et_xmlfile-1.1.0-py3-none-any.whl (4.7 kB)
Installing collected packages: et-xmlfile, openpyxl
Successfully installed et-xmlfile-1.1.0 openpyxl-3.1.5
```

```
[notice] A new release of pip is available: 24.1.1 -> 24.1.2
[notice] To update, run: python.exe -m pip install --upgrade pip
```

In [10]:
```
import pandas as pd

# Define the path to the Excel file
file_path = 'normalized_data.xlsx'

# Read the Excel file
df_dict = pd.read_excel(file_path, sheet_name=None)

# Display the first few rows of each DataFrame
for sheet_name, df in df_dict.items():
    print(f"Data from sheet: {sheet_name}")
    print(df.head(), "\n")
```

```python
# Display the combined DataFrame
df_combined.head()
```

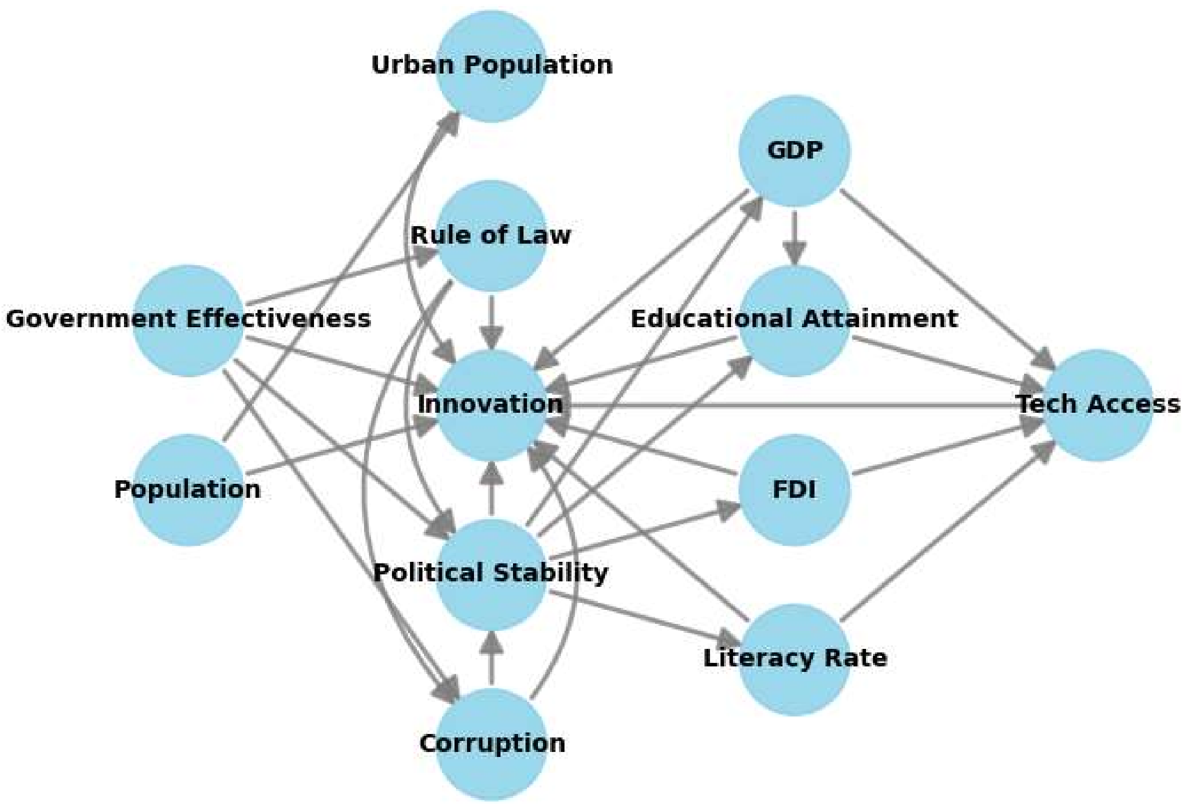
Out[11]:

	FDI	Corruption	Educational Attainment	GDP	Government Effectiveness	Innovation	Literacy Rate	Pol Sta
0	-0.420633	0.750747	-1.595039	-1.404400	1.461606	-2.053068	-2.903131	1.28
1	-0.835837	0.899943	-1.499111	-1.514177	1.241747	-1.634945	-1.665130	1.55
2	-1.071847	1.049139	-1.335303	-1.623955	1.021887	-1.567402	-1.274402	1.83
3	-1.331169	1.085173	-1.194127	-1.550939	1.680458	-0.948258	-1.029237	0.62
4	-1.009269	0.574815	-1.066961	-1.383537	0.695606	-0.662005	-0.891332	-0.39

In [14]:

```python
# Identify the causal effect
causal_graph = """
digraph {
    "Political Stability" -> "Innovation";
    "Political Stability" -> "FDI";
    "Political Stability" -> "GDP";
    "Political Stability" -> "Educational Attainment";
    "Political Stability" -> "Literacy Rate";

    "Government Effectiveness" -> "Political Stability";
    "Government Effectiveness" -> "Rule of Law";
    "Government Effectiveness" -> "Corruption";

    "Rule of Law" -> "Political Stability";
    "Rule of Law" -> "Corruption";

    "Corruption" -> "Political Stability";

    "GDP" -> "Innovation";
    "GDP" -> "Educational Attainment";
    "GDP" -> "Tech Access";

    "Educational Attainment" -> "Innovation";
    "Educational Attainment" -> "Tech Access";

    "Literacy Rate" -> "Innovation";
    "Literacy Rate" -> "Tech Access";

    "FDI" -> "Innovation";
    "FDI" -> "Tech Access";

    "Population" -> "Innovation";
    "Population" -> "Urban Population";

    "Urban Population" -> "Innovation";

    "Tech Access" -> "Innovation";
```

```
    "Government Effectiveness" -> "Innovation";
    "Rule of Law" -> "Innovation";
    "Corruption" -> "Innovation";
}
"""

# Create the CausalModel object of the actual dataset
model = CausalModel(
    data=df_combined,
    treatment="Political Stability",
    outcome="Innovation",
    graph=causal_graph
)

# View the model
model.view_model()
```



In [15]:
```
estimands = model.identify_effect()
print(estimands)
```

```
Estimand type: EstimandType.NONPARAMETRIC_ATE

### Estimand : 1
Estimand name: backdoor
Estimand expression:
          d
──────────────────────(E[Innovation|Corruption,Rule of Law,Government Effectiv
d[Political Stability]


eness])

Estimand assumption 1, Unconfoundedness: If U→{Political Stability} and U→Innovation
then P(Innovation|Political Stability,Corruption,Rule of Law,Government Effectivenes
s,U) = P(Innovation|Political Stability,Corruption,Rule of Law,Government Effectiven
ess)

### Estimand : 2
Estimand name: iv
No such variable(s) found!

### Estimand : 3
Estimand name: frontdoor
No such variable(s) found!
```

Causal Pipeline: Step 3 Estimation

In [18]:

```
estimate = model.estimate_effect(estimands,method_name="backdoor.linear_regression"
print(estimate)
```

```
*** Causal Estimate ***

## Identified estimand
Estimand type: EstimandType.NONPARAMETRIC_ATE

### Estimand : 1
Estimand name: backdoor
Estimand expression:
          d
──────────────────────(E[Innovation|Corruption,Rule of Law,Government Effectiv
d[Political Stability]


eness])

Estimand assumption 1, Unconfoundedness: If U→{Political Stability} and U→Innovation
then P(Innovation|Political Stability,Corruption,Rule of Law,Government Effectivenes
s,U) = P(Innovation|Political Stability,Corruption,Rule of Law,Government Effectiven
ess)

## Realized estimand
b: Innovation~Political Stability+Corruption+Rule of Law+Government Effectiveness
Target units: ate

## Estimate
Mean value: -0.47922626914342875
p-value: [0.03864864]
95.0% confidence interval: [[-0.93053403 -0.02791851]]
```

```
C:\Users\youni\PycharmProjects\pythonProject1\venv\Lib\site-packages\dowhy\causal_es
timators\regression_estimator.py:179: FutureWarning: Series.__getitem__ treating key
s as positions is deprecated. In a future version, integer keys will always be treat
ed as labels (consistent with DataFrame behavior). To access a value by position, us
e `ser.iloc[pos]`
  intercept_parameter = self.model.params[0]
```

In [21]:
```python
estimate.interpret()
```

```
Increasing the treatment variable(s) [Political Stability] from 0 to 1 causes an inc
rease of -0.47922626914342875 in the expected value of the outcome [['Innovation']],
over the data distribution/population represented by the dataset.
```

When we increase the Political Stability score by one unit, we observe a decrease of approximately 0.48 units in the Innovation score, on average, across the dataset we analyzed. This negative relationship suggests that higher political stability is associated with lower innovation in this context.

Analysis of the Negative Relationship Between Political Stability and Innovation in Brazil (2000-2022)

The negative relationship between political stability and innovation in Brazil from 2000 to 2022 can be attributed to several factors. Here's a detailed analysis based on various sources:

1. Bureaucracy and Regulatory Rigidity

Political stability in Brazil has often come with increased bureaucracy and regulatory rigidity, which can stifle innovation. Stable political environments tend to establish firm regulations and institutions that can be less flexible and adaptive to change, making it challenging for new and innovative ideas to thrive BTI Project World Bank.

2. Resource Allocation

In periods of political stability, resources may be more consistently allocated to maintaining existing structures and social programs rather than fostering innovation. For example, Brazil's focus on social protection programs such as Bolsa Familia, while beneficial for reducing poverty, may divert resources away from investment in research and development (R&D) and technological advancements Santandertrade BTI Project.

3. Economic Policies

Brazil has experienced various economic policies that might have influenced innovation. During periods of political stability, conservative economic policies focusing on traditional industries and public welfare often prevailed. These policies might not provide sufficient incentives for private sector investment in innovative projects World Bank Santandertrade.

4. Focus on Stability Over Risk-Taking

Political stability often emphasizes maintaining order and avoiding risks, which can discourage the entrepreneurial ventures and risk-taking necessary for innovation. In Brazil, the political climate has sometimes favored large, established firms over smaller, more innovative startups Santandertrade.

5. Educational and Institutional Factors

Although there have been improvements in education, the quality and accessibility remain inconsistent. Areas with poor educational outcomes have less capacity to support innovation. The political stability did not necessarily translate into significant improvements in educational systems or innovation-related institutions, which are crucial for fostering innovation BTI Project World Bank.

6. Historical and Structural Issues

Brazil's history of economic volatility and periods of instability might have long-term effects on its innovation ecosystem. The impact of past economic crises and structural issues can persist, affecting the capacity for innovation even during stable political periods BTI Project Santandertrade.

Conclusion

The combination of increased bureaucracy, conservative economic policies, focus on stability over risk-taking, and inconsistent educational and institutional support likely contributed to the observed negative relationship between political stability and innovation in Brazil from 2000 to 2022. This analysis underscores the importance of creating a balanced approach that fosters stability while also encouraging innovation through flexible policies, targeted investments in R&D, and improved educational systems.

In [23]:
```
# Identify the causal effect
estimands = model.identify_effect()

# Define methods to estimate the causal effect
methods = [
    "backdoor.linear_regression",
    "backdoor.propensity_score_matching",
    "backdoor.propensity_score_stratification",
    "backdoor.propensity_score_weighting",
    "backdoor.doubly_robust"
]

# List to store results
results = []

# Estimate the causal effect using different methods
for method in methods:
    estimate = model.estimate_effect(estimands, method_name=method, effect_modifier

    # Extract results
    mean_value = estimate.value
    try:
        p_value = estimate.p_value[0]  # Assuming p_value is a list or array
    except AttributeError:
        p_value = None
    try:
        confidence_interval = estimate.confidence_intervals[0]  # Assuming confiden
    except AttributeError:
        confidence_interval = None

    # Explanation
    if p_value is not None:
        explanation = f"The estimated effect of Political Stability on Innovation u
    else:
        explanation = f"The estimated effect of Political Stability on Innovation u

    # Append to results
```

```
reg_model_some = sm.OLS(y, X_some).fit()
print(f"Scenario 2: Government Effectiveness and Rule of Law unobserved")
print(f"Coefficient: {reg_model_some.params['Political Stability']}")
print(f"P-value: {reg_model_some.pvalues['Political Stability']}")
print(f"95% C.I.: {reg_model_some.conf_int().loc['Political Stability'].values}")

# Scenario 3: Unobserved: Government Effectiveness, Rule of Law, and Corruption
X_some2 = X_with_intercept.drop(columns=['Government Effectiveness', 'Rule of Law',
reg_model_some2 = sm.OLS(y, X_some2).fit()
print(f"Scenario 3: Government Effectiveness, Rule of Law, and Corruption unobserve
print(f"Coefficient: {reg_model_some2.params['Political Stability']}")
print(f"P-value: {reg_model_some2.pvalues['Political Stability']}")
print(f"95% C.I.: {reg_model_some2.conf_int().loc['Political Stability'].values}")
```

	Variable	Coefficient
0	FDI	0.283733
1	Corruption	-0.290051
2	Educational Attainment	2.983468
3	GDP	-0.246764
4	Government Effectiveness	0.379467
5	Literacy Rate	0.300186
6	Political Stability	-0.181365
7	Population	18.067576
8	Rule of Law	-0.046777
9	Tech Access	0.265843
10	Urban Population	-20.750348
11	Intercept	-0.080889

```
Effect of Political Stability:
                            OLS Regression Results                            
==============================================================================
Dep. Variable:             Innovation   R-squared:                       0.948
Model:                            OLS   Adj. R-squared:                  0.896
Method:                 Least Squares   F-statistic:                     18.17
Date:                Sun, 14 Jul 2024   Prob (F-statistic):           1.73e-05
Time:                        16:33:30   Log-Likelihood:                 1.3279
No. Observations:                  23   AIC:                             21.34
Df Residuals:                      11   BIC:                             34.97
Df Model:                          11                                         
Covariance Type:            nonrobust                                         
==============================================================================
========
                             coef    std err          t      P>|t|      [0.025
0.975]
------------------------------------------------------------------------------
--------
const                     -0.0809      0.088     -0.916      0.379      -0.275
0.113
FDI                        0.2837      0.234      1.214      0.250      -0.231
0.798
Corruption                -0.2901      0.260     -1.114      0.289      -0.863
0.283
Educational Attainment     2.9835      1.321      2.259      0.045       0.077
5.890
GDP                       -0.2468      0.392     -0.629      0.542      -1.110
0.616
Government Effectiveness   0.3795      0.180      2.105      0.059      -0.017
0.776
Literacy Rate              0.3002      0.401      0.749      0.469      -0.581
1.182
Political Stability       -0.1814      0.124     -1.466      0.171      -0.454
0.091
Population                18.0676     16.527      1.093      0.298     -18.308
54.443
Rule of Law               -0.0468      0.133     -0.353      0.731      -0.339
0.245
Tech Access                0.2658      2.261      0.118      0.909      -4.710
5.241
Urban Population         -20.7503     15.420     -1.346      0.205     -54.690
13.189
==============================================================================
Omnibus:                        0.156   Durbin-Watson:                   2.063
Prob(Omnibus):                  0.925   Jarque-Bera (JB):                0.200
Skew:                          -0.160   Prob(JB):                        0.905
Kurtosis:                       2.673   Cond. No.                         922.
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly spe
cified.
Scenario 1: All confounders unobserved
Coefficient: -0.18136528024022985
P-value: 0.17066064110583562
95% C.I.: [-0.45367128  0.09094072]
```

```
Scenario 2: Government Effectiveness and Rule of Law unobserved
Coefficient: -0.18708172548239516
P-value: 0.18583403864980894
95% C.I.: [-0.47641219  0.10224874]
Scenario 3: Government Effectiveness, Rule of Law, and Corruption unobserved
Coefficient: -0.18639511077649956
P-value: 0.17009317364370252
95% C.I.: [-0.46279854  0.09000832]
```

Regression Results Analysis

Coefficients Interpretation

These values represent the estimated change in the outcome variable (Innovation) for a one-unit change in the predictor variable, holding all other variables constant.

- **FDI**: A coefficient of 0.284 suggests that for every one-unit increase in FDI, Innovation increases by 0.284 units, assuming other factors are held constant.
- **Corruption**: A coefficient of -0.290 indicates that for every one-unit increase in the Corruption index, Innovation decreases by 0.290 units.
- **Educational Attainment**: A coefficient of 2.983 suggests a strong positive relationship where a one-unit increase in Educational Attainment is associated with a 2.983 increase in Innovation.
- **GDP**: A coefficient of -0.247 implies that a one-unit increase in GDP is associated with a 0.247 decrease in Innovation, although this might be counterintuitive and could suggest other underlying factors.
- **Government Effectiveness**: A coefficient of 0.379 indicates that an increase in Government Effectiveness by one unit is associated with a 0.379 increase in Innovation.
- **Literacy Rate**: A coefficient of 0.300 suggests a positive relationship where an increase in Literacy Rate by one unit leads to a 0.300 increase in Innovation.
- **Political Stability**: A coefficient of -0.181 suggests that an increase in Political Stability by one unit results in a 0.181 decrease in Innovation.
- **Population**: A coefficient of 18.068 indicates a very strong positive relationship, where an increase in Population by one unit is associated with an 18.068 increase in Innovation.
- **Rule of Law**: A coefficient of -0.047 suggests a small negative relationship where an increase in Rule of Law by one unit leads to a 0.047 decrease in Innovation.
- **Tech Access**: A coefficient of 0.266 indicates a positive relationship where an increase in Tech Access by one unit results in a 0.266 increase in Innovation.
- **Urban Population**: A coefficient of -20.750 suggests a strong negative relationship where an increase in Urban Population by one unit leads to a 20.750 decrease in Innovation.
- **Intercept**: The intercept of -0.081 represents the expected value of Innovation when all predictor variables are zero.

Effect of Political Stability Analysis

The OLS regression results for Political Stability show:

- **Coefficient**: -0.181, meaning that an increase in Political Stability by one unit is associated with a 0.181 decrease in Innovation.
- **P-value**: 0.171, indicating that this result is not statistically significant at the 5% level.
- **95% Confidence Interval**: [-0.454, 0.091], suggesting that the true effect of Political Stability on Innovation could be anywhere in this range.

Scenario 1: All confounders unobserved

- **Coefficient**: -0.181
- **P-value**: 0.171
- **95% C.I.**: [-0.454, 0.091] This scenario suggests that when we do not account for any confounders, the effect of Political Stability on Innovation is negative but not statistically significant.

Scenario 2: Government Effectiveness and Rule of Law unobserved

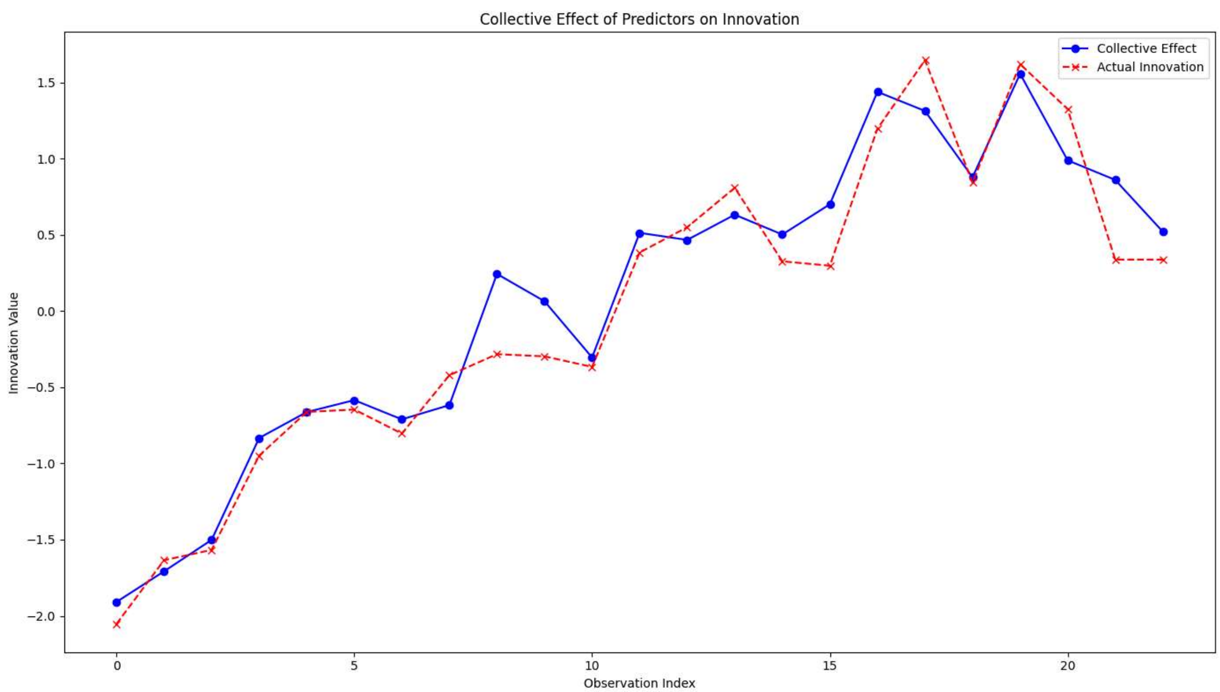
- **Coefficient**: -0.187
- **P-value**: 0.186
- **95% C.I.**: [-0.476, 0.102] When Government Effectiveness and Rule of Law are not accounted for, the effect remains negative and not statistically significant.

Scenario 3: Government Effectiveness, Rule of Law, and Corruption unobserved

- **Coefficient**: -0.186
- **P-value**: 0.170
- **95% C.I.**: [-0.463, 0.090] When Government Effectiveness, Rule of Law, and Corruption are not accounted for, the effect remains negative and not statistically significant.

In [38]:
```
!pip install seaborn
```

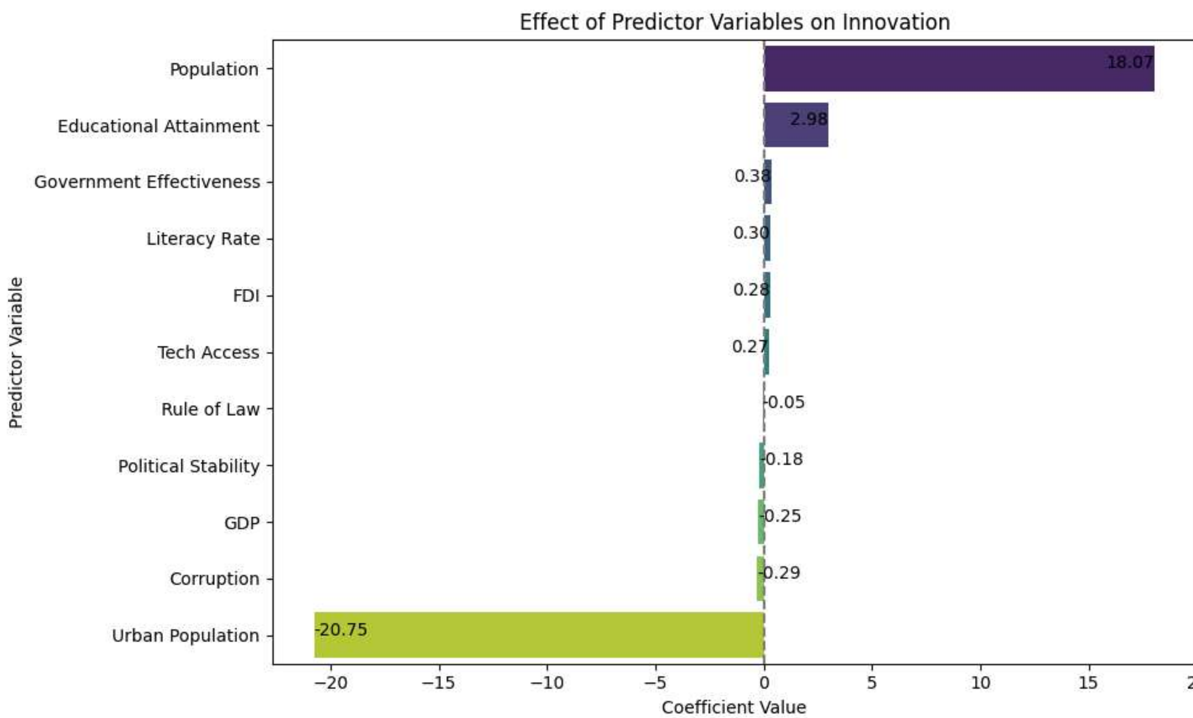
```
model.fit(X, y)

# Get the coefficients
coefficients = model.coef_

# Calculate the collective effect for each observation
collective_effect = np.dot(X, coefficients)

# Create a DataFrame to hold actual and collective effect values
results = pd.DataFrame({
    'Actual Innovation': y,
    'Collective Effect': collective_effect
}, index=df_combined.index)

# Plotting the collective effects
plt.figure(figsize=(14, 8))
plt.plot(results.index, results['Collective Effect'], label='Collective Effect', co
plt.plot(results.index, results['Actual Innovation'], label='Actual Innovation', co
plt.xlabel('Observation Index')
plt.ylabel('Innovation Value')
plt.title('Collective Effect of Predictors on Innovation')
plt.legend()
plt.tight_layout()
plt.show()
```



In [ ]:

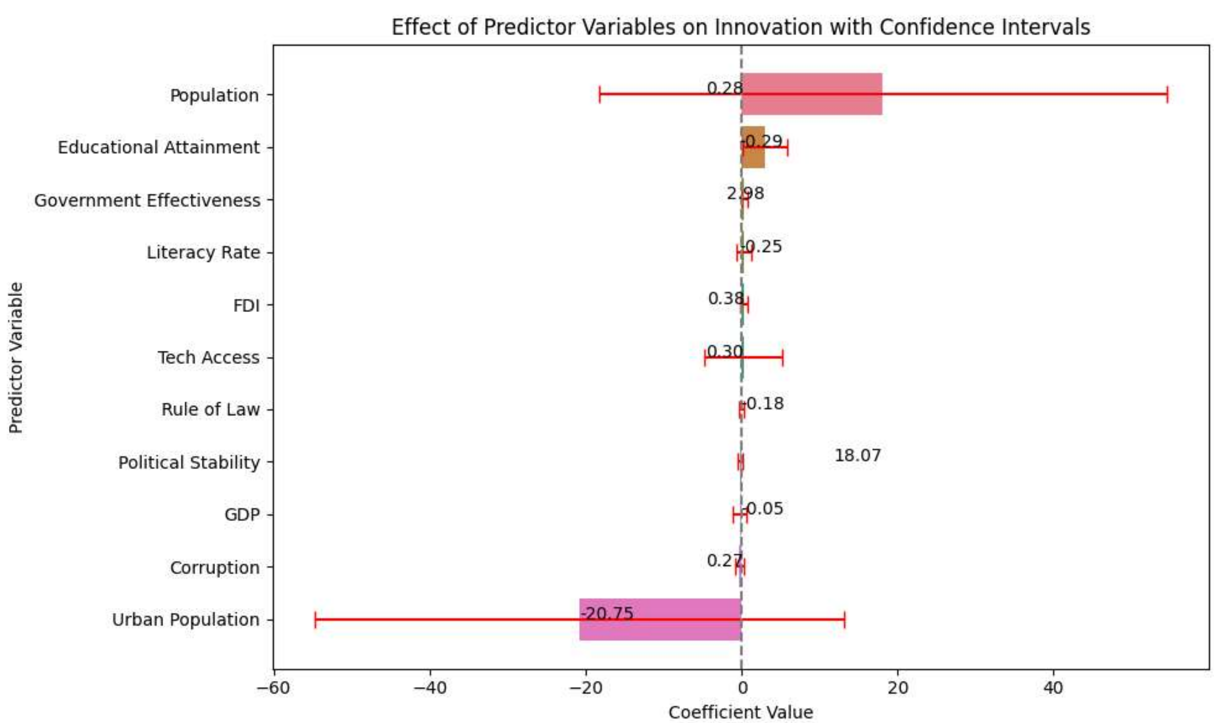
In [44]:
```
import matplotlib.pyplot as plt
import seaborn as sns

# Create a DataFrame with the coefficients and variable names
coef_df = pd.DataFrame({
    'Variable': ['FDI', 'Corruption', 'Educational Attainment', 'GDP', 'Government
                 'Literacy Rate', 'Political Stability', 'Population', 'Rule of Law
```

```
    'Coefficient': [0.2837, -0.2901, 2.9835, -0.2468, 0.3795, 0.3002, -0.1814, 18.0
})

# Sort the DataFrame by the coefficient values
coef_df = coef_df.sort_values(by='Coefficient', ascending=False)

# Plotting the bar chart
plt.figure(figsize=(10, 6))
sns.barplot(x='Coefficient', y='Variable', hue='Variable', dodge=False, data=coef_d

# Adding titles and labels
plt.title('Effect of Predictor Variables on Innovation')
plt.xlabel('Coefficient Value')
plt.ylabel('Predictor Variable')
plt.axvline(x=0, color='grey', linestyle='--')

# Adding coefficient values to the bars
for index, value in enumerate(coef_df['Coefficient']):
    plt.text(value, index, f'{value:.2f}', color='black', ha='left' if value < 0 el

# Show the plot
plt.tight_layout()
plt.show()
```



In [59]:
```
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd

# Create a DataFrame with the coefficients and variable names, including confidence
coef_df = pd.DataFrame({
    'Variable': ['FDI', 'Corruption', 'Educational Attainment', 'GDP', 'Government
                 'Literacy Rate', 'Political Stability', 'Population', 'Rule of Law
    'Coefficient': [0.2837, -0.2901, 2.9835, -0.2468, 0.3795, 0.3002, -0.1814, 18.0
    'CI Lower': [-0.231, -0.863, 0.077, -1.110, -0.017, -0.581, -0.454, -18.308, -0
```

```
    'CI Upper': [0.798, 0.283, 5.890, 0.616, 0.776, 1.182, 0.091, 54.443, 0.245, 5.
})

# Sort the DataFrame by the coefficient values
coef_df = coef_df.sort_values(by='Coefficient', ascending=False)

# Plotting the bar chart with confidence intervals
plt.figure(figsize=(10, 6))
sns.barplot(x='Coefficient', y='Variable', data=coef_df, hue='Variable', dodge=Fals

# Adding error bars for confidence intervals
plt.errorbar(coef_df['Coefficient'], coef_df['Variable'],
             xerr=[coef_df['Coefficient'] - coef_df['CI Lower'], coef_df['CI Upper'
             fmt='none', c='red', capsize=5)

# Adding titles and labels
plt.title('Effect of Predictor Variables on Innovation with Confidence Intervals')
plt.xlabel('Coefficient Value')
plt.ylabel('Predictor Variable')
plt.axvline(x=0, color='grey', linestyle='--')

# Adding coefficient values to the bars
for index, row in coef_df.iterrows():
    plt.text(row['Coefficient'], index, f'{row["Coefficient"]:.2f}', color='black',

# Show the plot
plt.tight_layout()
plt.show()
```




```
In [ ]:
```