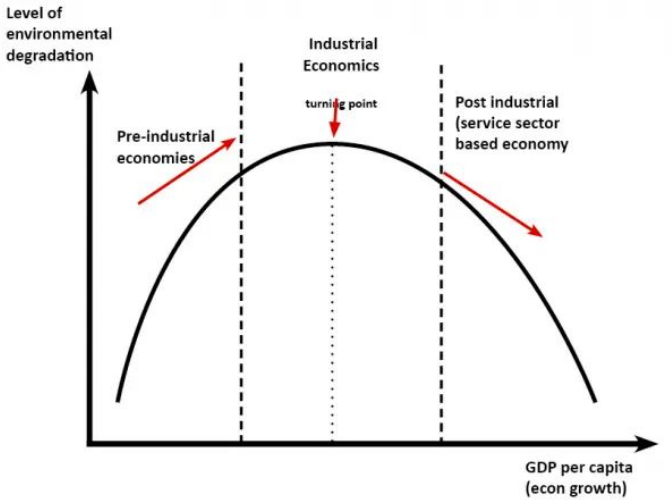
Analysis of the impact of Power Inequality on Ground Water Level

Group-14
Name and Roll no. of Teammates
Animesh Pareek (2021131)
Divyam Khorwal (2021388)
Mohit Bansal (2019257)
Sejal Khurana (2021418)

Introduction and Background

The environmental Kuznets curve (EKC), shows the relationship between economic progress and environmental degradation through time as an economy progress.

Groundwater depletion is a critical environmental problem, particularly in areas with high population growth. While environmental factors like rainfall and geological formations can influence groundwater quality, social and economic factors like power inequality may also play a role. In this analysis, we use the Environmental Kuznets Curve framework to explore the relationship between power inequality and groundwater level degradation. By examining the data through this lens, we can better understand how different social and economic indicators may contribute to this pressing environmental issue.



Variable Definition and Description

Variables	Description	Acronym
SDP	State domestic product - It is the aggregate money value of all goods and services produced within the geographical boundary of the state. Measured as per capita	SDP
SDP^2	Square of State Domestic Product	SDP2
SDP^3	Cube of State Domestic Product	SDP3
Gini	Gini index is the statistical measure used to measure the income distribution among the country's population, i.e., it helps measure the income inequality of the country's population	Gini

Baseline Model Results and Interpretation

$$\text{Ground Water Level}_{i,t} = \alpha_0 + \alpha_1 SDP_{i,t} + \alpha_2 SDP_{i.t}^{\ 2} + \alpha_3 SDP_{i,t}^{\ 3} + \alpha_4 GINI_i + \gamma_{i,t}$$

- The model has four coefficients: α0 represents the average ground water level when SDP and GINI Index are both 0
- α1+α2+α3 represents the average decline in ground water level due to a unit increase in SDP when GINI Index=0, and
- α4 represents the average decline in ground water level due to a unit increase in GINI Index when SDP=0.
- The table shows that the effect of SDP2 and SDP3 on ground water level is negligible, with the intercept value being -2.388 and the coefficients for SDP and GINI being 3.013 x 10^-5 and 24.76, respectively.

Dependent Variable >> Ground Water Level [N = 54,916 ; R^2 = 0.033]	
Explanatory Variables	Coefficients
Intercept	-2.388 *** (0.33)
SDP	3.013×10^{-5} *** (1.102×10^{-6})
SDP^2	-3.529×10^{-11} *** (1.604×10^{-12})
SDP^3	1.075×10^{-17} *** (6.384×10^{-19})
GINI	24.76 *** (1.026)

*** : p-value less than 0.001

Power Inequality and Control Variables Incorporated

Variable	Description	Acronym
Win-Margin(Turn out)	The difference in percentage between the share of votes cast for the winning candidate and the second-place candidate in an election.	win_margin
Contesting Candidates	State-wise number of candidates contesting for elections	candidateno
Literacy Rate	Literate rate per state	literacy_rate
Population	Population of the state	Population
Rainfall	Actual rainfall in the state between 2011-2021	Actual.rainfall
Forest Area %	Percentage Area of Forest Land	perForestArea
Net Sown Area %	Percentage Area of Net Sown Area	perNetSownArea

Enhanced Model

$$\text{Ground Water Level}_{i,,t} = \alpha_0 + \alpha_1 SDP_{i,t} + \alpha_2 SDP_{i.t}^{\ 2} + \alpha_3 SDP_{i,t}^{\ 3} + \alpha_4 GINI_i + \alpha_5 win\_margin + \alpha_6 candidateno + \alpha_7 literacy\_rate + \alpha_8 Population + \alpha_9 Actual.rainfall + \alpha_{10} perNetSownArea + \alpha_{11} perForestArea + \gamma_{i,t}$$

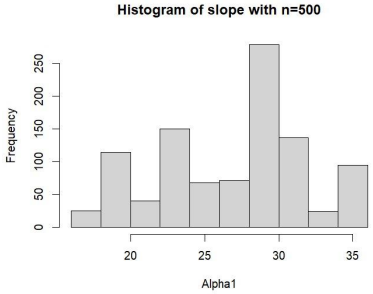
- We have an adjusted R squared of 0.2046.
- The F-statistic of 297.8 with a very low p_value
- Power inequality and gini variable have significant effects on the dependent variable at 0.05 level of significance .

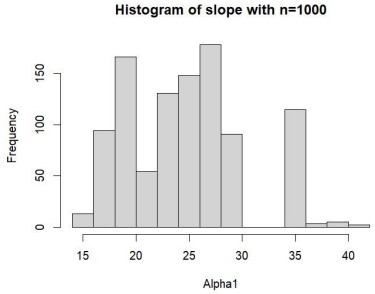
Variables	Coefficients	Standard error
Intercept	1.364e+01	1.100e+00
SDP	-4.108e-06	4.900e-06
SDP2	-8.126e-12	9.479e-12
SDP3	5.876e-18	5.172e-18
GINI	2.074e+01	2.556e+00
Winmargin	-1.798e+01	9.370e-03
Candidateno	8.510e-03	6.487e-04
Literacyrate	-1.907e-01	1.342e-02
Population	1.529e-06	1.211e-07
ActualRainfall	-1.094e-04	3.268e-05
Per net sown area	2.6433-01	7.547e-03
PerForestArea	4.570e-02	9.029e-03

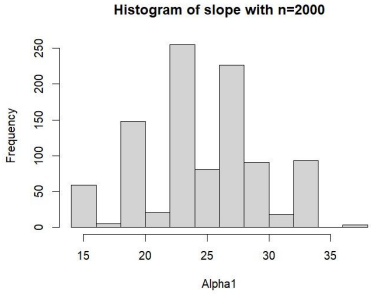
*p<0.1; **p<0.05; ***p<0.01

MONTE-CARLO SIMULATION

- We run Monte- Carlo simulation for sample sizes 500, 1000 and 2000
- With increase in number of sample sizes, standard deviation decreased and estimated value came close to true value.
- With the increase in sample size, histogram gets skinnier and more concentrated around true value.







Chow Test and T-test

- After running the Chow test, we noticed that the p-value(2.2e-16) is less than the 0.05 significance level for all the possible combinations of state groups.
- This shows that the coefficient of regression models are not consistent and differ for the different regions.
- For the t-test, our null hypothesis is mean of EQI of region1 = mean of EQI of region2.
- After running the test, we noticed that the p-value(=5.70e-16) is less than 0.05 and thus the null hypothesis was rejected for all the different combinations of regions.

MLE

Maximum Likelihood Estimation is used to estimate the parameters of a model by maximizing the likelihood function which is a measure of how well the model fits the data . MLE can improve estimation of our model as it provides estimates that are efficient and asymptotically unbiased .

Variance

According to our data, there are considerable differences in the average environmental quality between groupings of states. Further research using the Bartlett test of homogeneity of variances indicated that there is not equal variation in environmental quality across state groupings.

The assumption of homoscedasticity of the residuals in OLS is broken if the variance considerably varies across state groups. Across all levels of the independent variables, OLS makes the assumption that the error variance is constant.

Alternate strategy that could be used to take care of the heteroskedastic nature of the data is to use Weighted Least Squares (WLS) regression , by adjusting the weights of the observation based on their variances . The variances of the errors are allowed to differ across observations . These weights are determined based on the inverse of the variance of each observation . Smaller variance - more weight , than larger variance , which reduces the impact of larger variance on estimated coefficients

What’s new in our research

We have considered many factors such as :-

Control : Population, Rainfall, %Forest, %Net Sown Area

Economic : GINI, SDP, SDP^2, SDP^3

Political : Win-Margin, Contesting candidates

Social : Literacy Rate

To have a better analysis of Ground Water Level we considered all of the above factors.

Higher literacy rates , Rainfall and Forest Cover may lead to an increase in ground water level .

On the other hand , High Population and Net Sown Area may lead to a decrease in ground water level.

Conclusion

- In conclusion, this linear regression model provides evidence of significant relationships between Ground.water.level and some of the independent variables, including SDP, perNetSownArea, and win_margin. However, the model explains only about 21% of the variability in Ground.water.level, indicating that other factors not captured in the model may also be important.
- The relationship between Ground.water.level and SDP is not linear but shows a U-shaped curve. As SDP increases from very low values, Ground.water.level first increases, then decreases, and then increases again. The relationship between Ground.water.level and New_gini is not statistically significant at the 0.05 level.
- The positive, negative, and positive coefficients for SDP, SDP squared, and SDP cubed, respectively, suggest a U-shaped relationship between SDP and Ground.water.level. As SDP increases from very low values, Ground.water.level first increases, then decreases, and then increases again.
- Additionally, the coefficients for perNetSownArea and win_margin are positive, indicating that as these variables increase, Ground.water.level is also expected to increase.
- The significant positive coefficient for perNetSownArea suggests that increasing the net sown area can lead to an increase in Ground.water.level, which may be an important consideration for water management and agricultural planning. The significant negative coefficient for win_margin suggests that closer election results may lead to lower Ground.water.level, which could be an interesting finding for further investigation.

THANK YOU!!

References

- https://eci.gov.in/statistical-report/statistical-reports/
- https://data.gov.in/
- https://www.indiavotes.com
- Data Assignment 2 (google.com)
- Department of Social Sciences and Humanities - IIIT Delhi