Testing the Environmental Kuznets Curve(EKC) Hypothesis for Groundwater Level in India

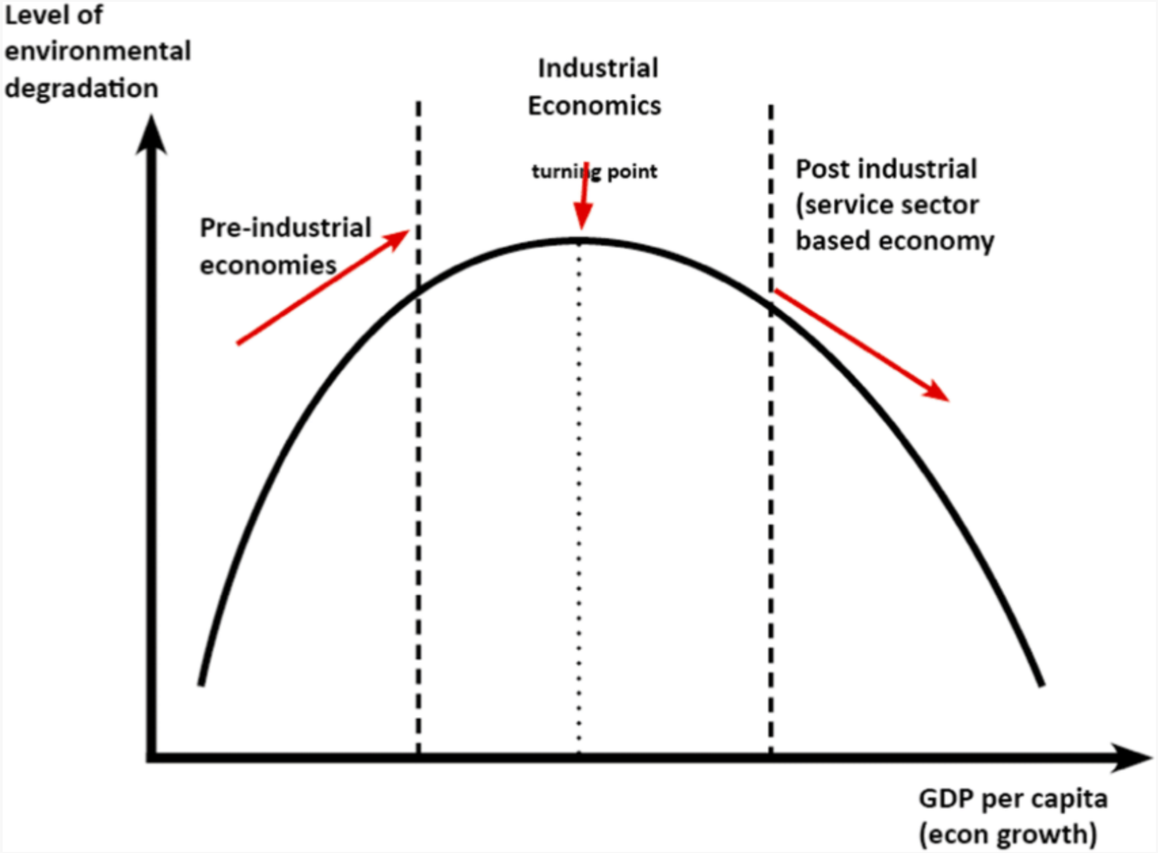
Bhawna Kapoor (2019356)
Nikhil Kumar (2021174)
Hardik Singh (2021390)
Shlok V. Mehrolya (2021421)

I Research Background & Motivation

- The Environmental Kuznets Curve (EKC) is a hypothesized relationship between **environmental quality** and **economic development.**

- It is named after Simon Kuznets who proposed that income inequality **first rises and then falls** as economic development proceeds.

- This project aims to study the validity of the EKC in the presence of power and income inequality



- We will analyze this relationship for the states and union territories in India for the years 2014-2017.

II Variable descriptions

Variable	Description	Acrononym
Groundwater Level	Average annual District wise ground water level. Measured as the distance from ground to water level in meters.	Ground.water.level
State Domestic Product (SDP)	Net State Domestic Product per capita(in 10,000 rupees)	SDP, SDP2, SDP3
GINI Index	Statistical measure of economic inequality in a population. 0 represents perfect equality (everyone has an equal share of the wealth or income) and 1 represents perfect inequality (one individual has all the wealth or income and the rest have none).	GI
Rainfall	Average annual District wise Precipitation. Measured annually in mm.	rainfall
Voter turnout	Percentage of eligible voters that cast their vote in the respective state assembly elections turnout	Turnout
Agriculture sown area	total land under cultivation (per 10,000Ha)	Net.sown.land.area
Gender parity index for high school scheduled tribe students	Ratio of the number of female scheduled tribe students enrolled in a high school to the number of male scheduled tribe students enrolled in the high school .	gphs.st

III Results from Data Assignment 1

Regression Model:
GroundWaterLevel~SDP+SDP^2+SDP^3+GI

Varibles	Coeficients	Std Error
(Intercept)	-1.514*	0.835
SDP	0.215***	0.028
SDP2	-0.002***	0.0004
SDP3	3.883e-06**	1.651e-06
GI	20.099***	20.099

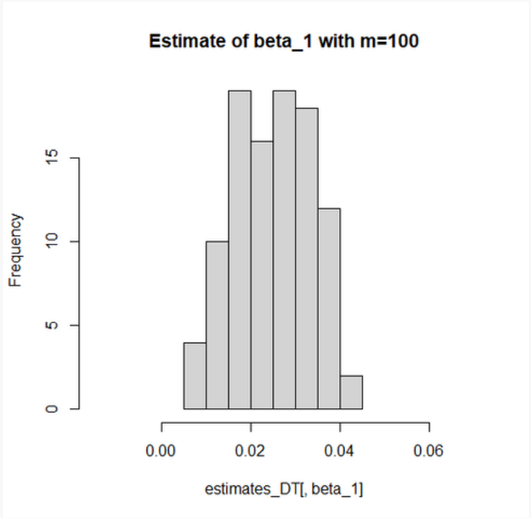
*p<0.1; **p<0.05; ***p<0.01

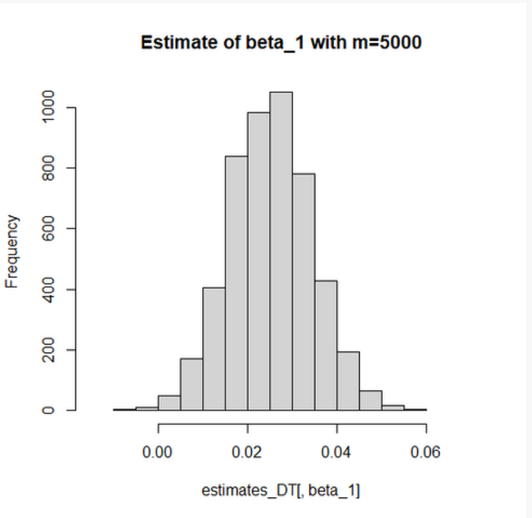
- We did find a U-Shaped Relation that after around 6,00,000 SDP the curve saw a downward trend for a given level of economic inequality.
- We also found out that as we move from a perfectly equal to a perfectly unequal district with the same SDP the Ground water level increased by 20m. Showing that an unequal distribution of wealth causes the people to be exploit the ground water level more.
- However this does not imply causation. As, it's important to note that the R-squared value was only 0.046 which indicates that the model explains only a small amount of the variance in the dependent variable, suggesting that other factors not included in the model may also be influencing ground water levels.

IV Consistency

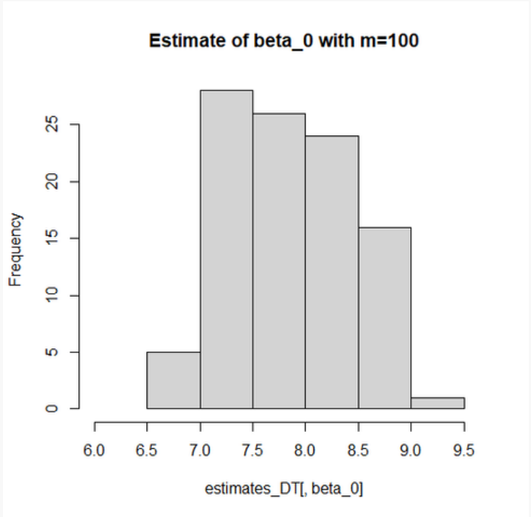
Monte-Carlo Simulation

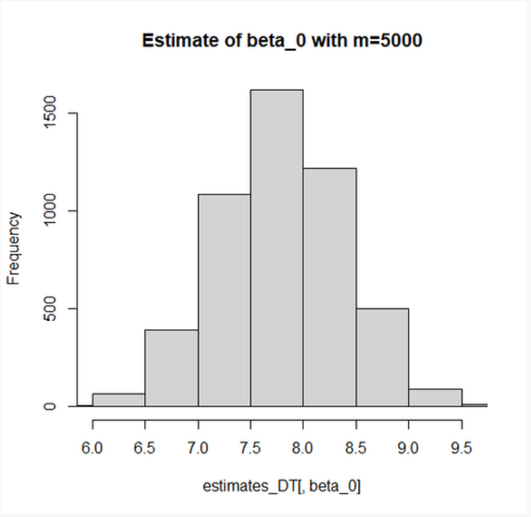
Model:
Groundwaterlevel~beta_0 +
beta_1*SDP + ui





Both the Intercept and the Slope were found Consistent i.e. the model consistently estimated the true relationship between the independent variables and the dependent variable, and the estimated coefficients become more accurate as the sample size grows.





V Results from Data Assignment 2

Variables	Coefficients	Std. Error
(Intercept)	18.658***	2.554
SDP	0.023	0.023
SDP2	0.0001	0.0002
GI	11.668***	3.229
log(Rainfall)	-1.523***	0.419
turnout	-0.103***	0.025
Net.sown.land.area	0.184***	0.010
GHHS.ST	-5.268***	1.767

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

- The regression output shows the relationship between Ground.water.level and several independent variables, including state domestic product, income inequality, rainfall, voter turnout, net sown land area and gender parity in education among Scheduled tribe students in high school.

- The results reveal that SDP and GI have a positive relationship with ground water level, whereas log(Rainfall) and Turnout have a negative relationship. Net.sown.land.area and GPHS.ST also have a positive and negative relationship, respectively.

- The R-squared value was 0.488 respectively which suggest that the model is a good fit, and the F-statistic indicates that the overall regression model is significant.

VI Observations

- When we divided the country into two regions i.e northern and southern region we found structural break in the mean environmental quality across different state-groups. This was verified by the pvalue of both the chow test and the t test. And the difference was statistically different even at the 0.01 significance level. Meaning that the mean of ground water level is statistically different in the two state groups.

- This could lead to omitted variable bias in our model. As a result, the estimated coefficients of the included variables may be biased and inefficient.

- Therefore we realized that it would be better if we include a dummy variable for the state groups(east, west, north, south).

Test	P-value
Chow-test	2.9e-20
t-test	8.7e-130

VII Final Model

Variables	Coefficients	Std Error
SDP	-0.020	0.034
SDP2	0.0002	0.0003
north	6.630***	0.773
east	1.694**	0.706
south	3.508***	0.831
north_east	2.210**	0.930
west	1.416	3.613
GI	5.606*	3.260
Turnout	-0.010	0.033

Variables	Coefficients	Std Error
log(Rainfall)	-1.539***	0.426
Net.sown.land.area	0.169***	0.009
GPHS.ST	-1.387	2.090
(Intercept)	8.320**	3.489

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

- Adding the state group variable to the regression model resulted in an increase in the R-squared value from 0.488 to 0.545. This indicates that the state group variable is an important factor in explaining the variation in the dependent variable, ground water level.
- The increase in R-squared suggests that the state group variable accounts for a significant amount of the variation in ground water level that was previously unexplained by the other independent variables in the model. Therefore, including the state group variable in the model improves its overall explanatory power and makes it more accurate in predicting ground water levels.

VIII Conclusion

- The coefficient for SDP is negative and SDP2 is positive this does show an inverted U-shaped relation, however according to the ekc hypothesis we expected the opposite effect. But both the variables had a high standard error and were not statistically significant even at the 0.1 significance level. Therefore we were not able to get the Kuznets Curve for ground water level.

- The coefficient for **GI is positive** and statistically significant, indicating that areas with a higher level of income inequality tend to have higher ground water levels. This can be interpreted as unequal distribution of wealth causes the people to be exploit the ground water level more.

- The coefficient for **log(Rainfall) is negative** and statistically significant, meaning that an increase in rainfall is associated with a decrease in ground water level. This is obvious and alligns with our expectation as regions with higher rainfall have better ground water level.

- The coefficient for **Turnout is negative**. Suggesting Regions with high electoral tourout have more motivated people showing high awareness and as we expect more aware people care more about the enviroment. However the variable is not statistically significant therefore we cannot make the claim with certainity.

- The coefficient for **Net.sown.land.area is positive** and statistically significant, indicating that an increase in net sown land area is associated with an increase(depletion) in ground water level. This is obvious as we expect the areas with more land under cultivation to have higher ground water level as agriculture is a water demanding activity.

IX Conclusion

- Coefficients for four state group dummy variables (north, east, south, and northeast) are statistically significant and positive, indicating that compared to the reference group (central states), these four state groups tend to have higher ground water levels. For example, we can see that northern sates have 6.630m higher ground water levels as compared to central state group holding other factors fixed.

- The coefficient for the west group is not statistically significant (as std error is very high), meaning that there is no evidence to suggest that this group has a different ground water level compared to the reference group.

- Standard Errors for some of the state group dummy variables (west) are relatively large, indicating some uncertainty around the estimated coefficients.

- The coefficient for GPHS.ST is negative indicating that areas with a higher level of gender equality is associated with higher ground water level this can be expected as areas with less inequality between girls and boys are more considerate towards environment. However the variable is not statistically significant therefore we cannot make the claim with certainity.

- It should be taken into consideration that lower ground water level is an indication of better environment quality.

x

THANK YOU!