P.G.D.A.V. College

University of Delhi

Nehru Nagar, Ring Road, New Delhi – 110065

Website: http://pgdavcollege.in
Email: pgdavcollege.edu@gmail.com

Supporting document

for

Annual Quality Assurance Report, 2023-24
Criteria 1.3.2

Number of courses that include experiential learning through project work/field work/internship during the year

Table of Content

Commercium's Mother Dairy Plant Trip

An Official Site Visit

On the 25th of January, 2024, Commercium planned a visit to the Patparganj plant of Mother Dairy in New Delhi. Hoping to learn how the company manages the supply chain of their operation, Commercium took about 45 students of the society and 3 faculty members on this enriching and insightful trip.

PGDAV College

PGDAV College is one of the oldest constituents of the University of Delhi. It was established in 1957 with the vision of the Dayanand Anglo Vedic movement (DAV) for imparting world class education rooted in the Indian milieu.

The College presents a bouquet of thirteen vibrant undergraduate and four postgraduate courses across twelve disciplines where it offers both the Honours and Programme degrees at the undergraduate level. We have a team of committed and highly qualified teaching faculty and supporting staff. Along with rigorous academic and intellectual training, the College also provides co-curricular and cultural activities for the students. Our wards excel in sports, fine arts, music, dance, and theatre. The College consistently receives numerous accolades in sports and has nurtured and groomed students to achieve distinctions at national and international levels.

P.G.D.A.V. College is committed to facilitating the learning of differently-abled students by making the College a more sensitive space through a very active Enabling Cell and supportive teaching and non-teaching staff.

Commercium: The Commerce Society of PGDAV College

Ever heard of a massive college society that provides opportunities to host and organize events where you can not only learn but continually improve on your marketing, promotional and blogging skills?

COMMERCIUM- The Commerce Society of PGDAV College (DU) is one such society with the motto of **"WHERE IGNITED MINDS THRIVE."** It stands as the biggest society in the college and one of the top 10 biggest departmental societies of Delhi University. With years of legacy, Commercium takes pride in organising two flagship events i.e. YOUTH CONCLAVE and COMMQUEST, a unique platform that brings together dynamic young minds, industry leaders and thought-provoking speakers.

In addition to the Youth Conclave, our society organises COMMQUEST, the annual commerce fest, another stellar event that emphasizes teamwork, creativity, and strategic thinking. Beyond events, we nurture minds through engaging orientations, insightful seminars and much more.

Teachers who accompanied Students in the industrial visit :

1. Dr. Mini Walia
2. Prof. Rakesh Kumar
3. Prof. Surender Singh

Mother Dairy: A Brief History

Mother Dairy was commissioned in 1974 and is a wholly owned subsidiary of the National Dairy Development Board

(NDDB). It was an initiative under Operation Flood, the world's biggest dairy development program launched to make India a milk sufficient nation. Over the years, Mother Dairy has contributed significantly in achieving this objective through a series of innovations and programs. Today, Mother Dairy manufactures markets & sells milk and milk products including cultured products, ice creams, paneer, and ghee under the Mother Dairy brand. The Company also has a diversified portfolio with products in edible oils, fruits & vegetables, frozen vegetables, pulses, processed food like fruit juices, jams, etc. to meet the daily requirements of every household.

The Company over the last many years has created a market leadership position for itself in branded milk segment in Delhi & NCR through a robust network of its booth and retail channels. It has also expanded its reach to other regions in North, South, East and West with its offering of Milk and Milk products pegging it among the few companies to own such a vast channel of distribution in India.

Description of the Event:

Commercium’s educational visit to Mother Dairy on the 25th of January 2024 was a remarkable experience that left a lasting impact on the minds of the members as well as the teachers. Our journey into the heart of Mother Dairy's operations in Patparganj, Delhi, unfolded as a perfect blend of education and enjoyment, providing us with a comprehensive understanding of the dairy industry and its intricacies.

Miss Neha, the Senior Marketing Executive, played a pivotal role in shaping our visit. Throughout the entire visit, she welcomed and guided us with enthusiasm. Her initial introduction to the history of Mother Dairy and its subsidiaries, namely Safal and Dhara, set the stage for a day filled with learning and exploration. The subsequent video presentation served as a virtual gateway into the world of Mother Dairy, offering insights into their growth, values, and their commitment to excellence.

The real adventure began as we embarked on a guided tour of the Mother Dairy factory, immersing ourselves in the various stages of the supply chain. From milk collection to pasteurization, we witnessed firsthand the meticulous processes that ensure the production of high-quality dairy products. Miss Neha's expertise guided us through the complexities of the operation, shedding light on the company's dedication to maintaining stringent quality control measures.

An exceptional aspect of our visit was the opportunity for some members to actively participate in 2-3 milk adulteration tests, part of the 29 tests conducted by Mother Dairy. This hands-on experience provided a deeper understanding of the efforts involved in preserving the purity of the products, reinforcing the importance of quality assurance in the dairy industry.

The factory tour also unveiled the impressive infrastructure supporting Mother Dairy's operations. The giant silos storing collected milk, the solar plants contributing to sustainable energy practices, and the rainwater harvesting system showcased the company's commitment not only to quality but also to environmental responsibility. These elements underscored Mother Dairy's holistic approach to its role in society.

Amidst our educational exploration, the visit took a delectable turn with the chance to savour the diverse products of Mother Dairy. Flavoured milk and an array of delicious ice cream flavours became not just a treat for our taste buds but a celebration of the company's commitment to providing wholesome and enjoyable dairy experiences. The gesture of gifting each participant a keychain bottle opener added a personal touch, serving as a tangible reminder of the day's adventure.

As the trip reached its conclusion, Miss Mini Walia, our co-convenor, extended a sincere token of thanks to Miss Neha. The gratitude expressed not only acknowledged her role in facilitating our visit but also recognized the efforts of the entire Mother Dairy team. The exchange of appreciation was captured in several group photos, immortalizing the joy and camaraderie that defined our day at Mother Dairy.

In retrospect, the visit emerged as a perfect amalgamation of education and enjoyment, providing us with a holistic understanding of the dairy industry. The firsthand exposure to the supply chain intricacies offered valuable knowledge and insights that extended beyond the confines of textbooks. Witnessing the dedication and efforts invested in ensuring the production of high-quality dairy products left a profound impact on our understanding of the industry's dynamics.

Beyond the academic facets, the visit fostered a deep sense of appreciation for the individuals behind the scenes at Mother Dairy. The commitment and hard work of the people working tirelessly to uphold the company's core values were truly inspiring. It became evident that Mother Dairy, as a well-established brand, continues to evolve and adapt to the changing demands of the market while maintaining its unwavering commitment to quality and sustainability.

In conclusion, the visit to Mother Dairy was an enriching experience that transcended traditional learning boundaries. It was a journey that allowed us to bridge the gap between theoretical knowledge and practical applications, bringing the subject matter to life. As we departed the premises, each of us carried not only the keychain bottle opener but also a sense of respect for the dairy industry and an awareness of the dedication required to uphold the highest standards of quality and sustainability.

The memories of this visit will undoubtedly serve as a source of inspiration and enlightenment for years to come, shaping our perspectives and contributing to our personal and academic growth.

The day spent at Mother Dairy stands as a testament to the importance of experiential learning, demonstrating how real-world encounters can significantly enhance our understanding of complex industries like dairy production. The impact of this visit extends far beyond the confines of our classroom, preparing us to face the challenges and opportunities that lie ahead in our academic and professional journeys.

Appendix 1 : Few Images of the Visit


DOCK LABORATORY
डाक प्रयोगशाला

Appendix 2 : Details of the Students

S.N.	Name	Course	Section	Year
1	Mohit Tanwar	B.Com (P)	B	2nd
2	Antisha Karpatne	B.Com (H)	C	3rd
3	Shweta Mehta	B.Com (H)	D	3rd
4	Shivam Gupta	B .Com (P)	C	2nd
5	Samyak Jain	B.Com (H)	A	1st
6	Atishay Jain	B.Com(P)	D	2nd
7	Muskan	B.Com(P)	B	2nd
8	Dishakha	B.Com (H)	D	3rd
9	Vanshita Garg	B.Com (H)	D	2nd
10	Akshay goel	B.Com (H)	A	1st
11	Kanishka Negi	B.Com (P)	C	2nd
12	Harshit Chauhan	B.Com (P)	C	2nd
13	Divyanshi Singh	B.Com	D	3rd
14	Shiv kumar	B.Com (H)	E	1st

15	Bhavik Arora	B.Com (H)	B	1st
16	Deepanshu	B.Com(P)	B	2nd
17	Sneha sharma	B.Com (H)	A	3rd
18	Raghav Bindra	B.Com(P)	D	2nd
19	Pratyaksh Mishra	B.Com (H)	A	3rd
20	Sunder kumar	B.Com(P)	B	2nd
21	Aryan Maan	B.Com	B	3rd
22	Shivani Bansal	B.Com (H)	D	3rd
23	Divya Agarwal	B.Com(P)	A	2nd
24	Sanchit Dilwani	B.Com (H)	C	3rd
25	Ananya	B.Com (P)	A	1st
26	Shreya Bhogal	B.Com (H)	A	3rd
27	Rivaa jain	B.Com(P)	D	1st
28	Apoorva Dhondiyal	B.Com (P)	A	1st
29	Riya arya	B.Com (H)	D	2nd
30	Susrutt Tiwari	B.Com(P)	A	1st
31	Pujit Goyal	B.Com (H)	D	1st
32	Somya Jain	B.Com (H)	B	2nd
33	Tanmay Phabrani	B.Com (H)	D	3rd
34	Yashika Sharma	B.Com (H)	D	2nd
35	Aditya Sahu	B.Com (P)	B	3rd
36	Maanya Palta	B.Com (H)	D	2nd
37	Aanchal Nagar	B.Com (H)	D	2nd
38	Dipit Madan	B. Com (P)	B	1st
39	Divyaanga Arun	B. Com (P)	C	2nd
40	Surabhi Tiwari	B .Com (H)	A	3rd
41	Subham Batu Singh	B . Com (P)	A	3rd
42	Parth Bahl	B.Com (H)	D	2nd
43	Apoorv Jain	B.Com (H)	B	3rd
44	Sahil Anand	B . Com (P)	D	3rd
45	Rohan Kathuria	B.com (P)	C	3rd

Prof . Rakesh Kumar
IQAC, Coordinator

Sh. Surendra Kumar
Convenor - Commercium

Prof. Surender Singh
Teacher In-charge

P.G.D.A.V. (M) COLLEGE, UNIVERSITY OF DELHI
Viksit Bharat Sankalp Club @2047: Sankalp se Siddhi

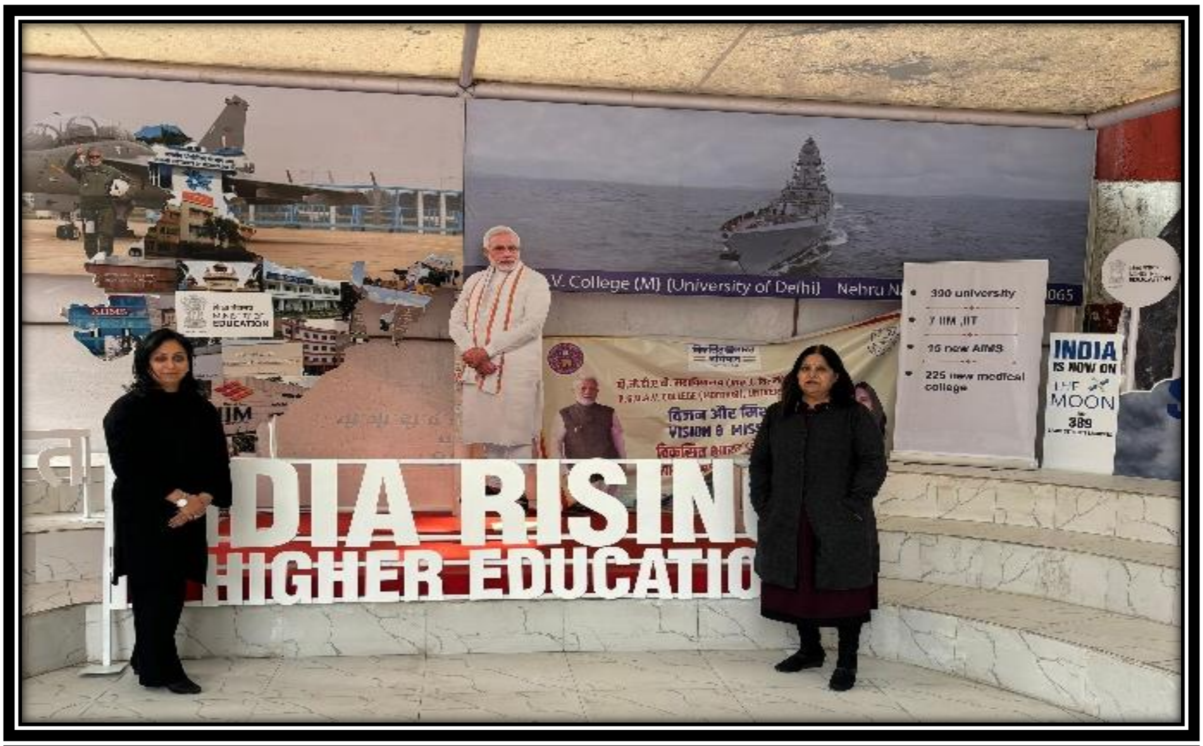
Viksit Bharat Sankalp Club stands as a beacon of progressive thought and action, dedicated to shaping a brighter future for India. With a steadfast commitment to fostering intellectual discourse, promoting socio-cultural awareness, and nurturing talent, the club has been at the forefront of igniting positive change within the community. Through a diverse array of events spanning various themes and disciplines, the club endeavours to inspire, educate, and empower individuals to contribute meaningfully to the nation's development journey. Each event serves as a testament to the club's unwavering dedication to realizing the vision of a prosperous and enlightened India by 2047.

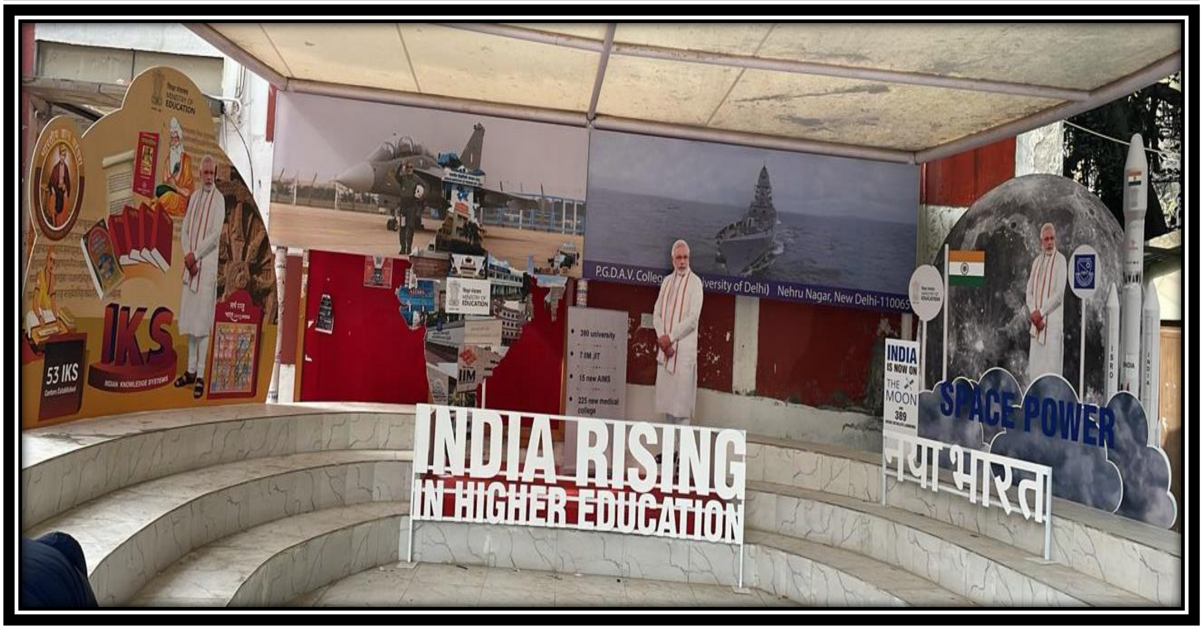
Team Members:

	Faculty Members:	Student Coordinators:
Convenor/Nodal Officer: Dr Gitu Nijhawan **Co-Convenor:** Dr Prakash Chandra	Dr Indu Dutta Dr Dharmendra Singh Ms Anshita Jain Ms Rajni	Khushbu Malhotra Khushi Singh

A Glimpse into our Recent Initiatives:
Dedicated Selfie Point- Capturing Moments, Creating Memories

In 2024, our college embarked on a series of dynamic initiatives that aligned closely with the vision of Viksit Bharat, showcasing our commitment to fostering national development and community engagement. We began the year by introducing a dedicated Selfie Point, which quickly became a popular spot-on campus, enhancing engagement and excitement among students, faculty, and visitors, and providing a fun way to capture and share moments of their academic journey.





EVENT 1
University and College Level Speech Competition
"विकसित भारत @2047: संकल्प और सिद्धि"

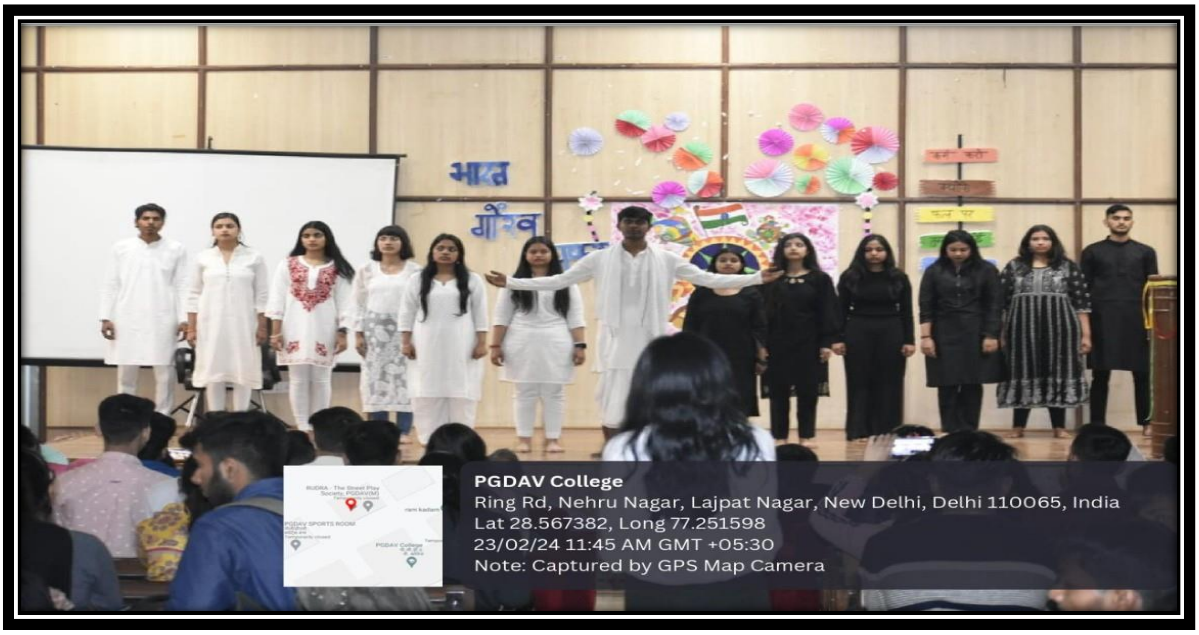
Our efforts to promote intellectual discourse were highlighted by the highly successful Inter College/University Speech Competition themed 'विकसित भारत @2047: संकल्प और सिद्धि' . This event attracted over 300 registrations from 56 esteemed colleges and universities, fostering a robust exchange of ideas and perspectives on the future of India, and underscoring the shared commitment to the nation's development.

EVENT 2:
PADHANTU- Shikshit Bharat
DARPAN- Book Discussion Series on Hindus in Hindu Rashtra

The "Darpan" event, organized by Padhantu, featured a thought-provoking lecture by Dr. Anand Ranganathan on "Hindus in Hindu Rashtra". This event aimed to provide enriching experiences aligned with the vision of Viksit Bharat, encouraging deeper understanding and dialogue on important socio-political topics.

EVENT 3:
UDAAN
Unfolding Drama and Acts to Awaken the Nation

UDAAN, another marquee event, showcased vibrant art, engaging debates, and inspiring stage plays that resonated with PM Modi's vision, leaving a lasting impression on all participants and highlighting the cultural and intellectual vibrancy essential for national progress.

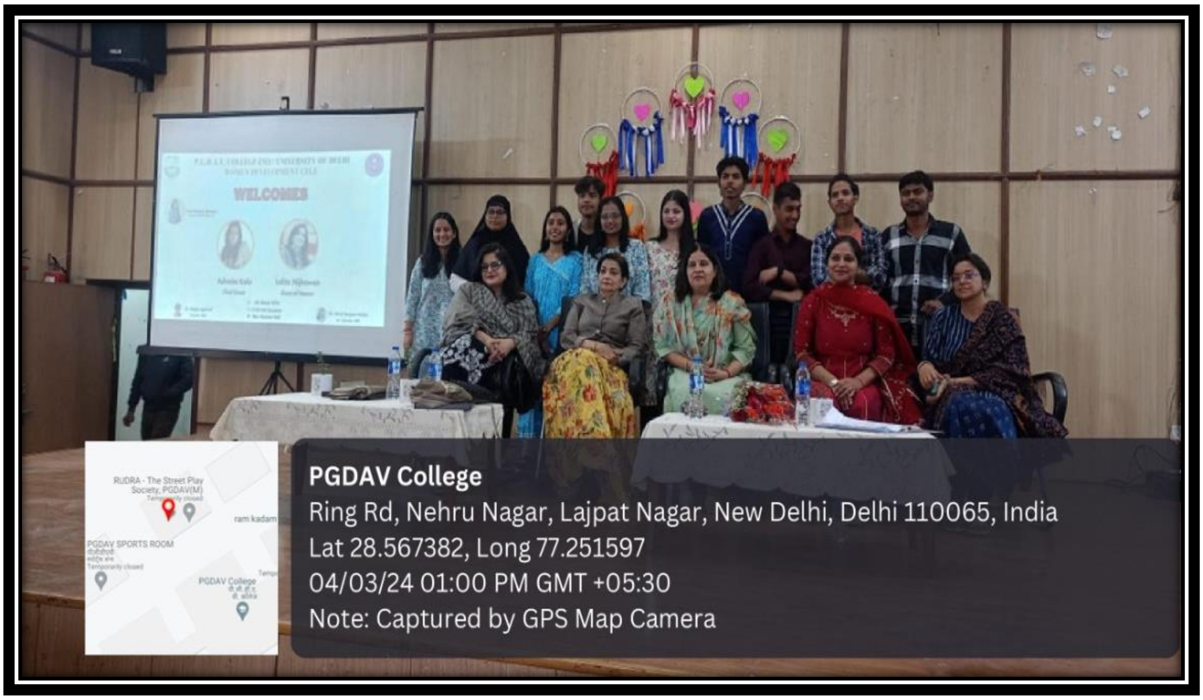


EVENT 4:
KAIZEN: THE CAREER COUNSELLING CLUB
Navigating the Future: Career Pathways in Artificial Intelligence & Robotics

Our career counselling sessions by Kaizen, led by Dr. Paraminder, provided students with invaluable insights into artificial intelligence and robotics, empowering them with the knowledge and skills critical for future careers. This initiative was part of our broader effort to align educational opportunities with the technological advancements and employment trends of the future.

EVENT 5:
Women Development Cell- International Women's Day
"विकसित भारत में महिलाओं की भागीदारी"

In March, we celebrated International Women's Day through a series of events hosted by the Women's Development Cell. These events highlighted the significant contributions of women to Viksit Bharat, featuring inspiring guest speakers from various fields who reinforced the essential role of women in our society's development and progress.



EVENT 6:
Live Screening of Honourable PM Modi
Inauguration of Projects on Semiconductor

A key highlight of our year was hosting a live screening of the Prime Minister's address, which inaugurated transformative projects. This event underscored our commitment to embracing innovation and progressive initiatives in line with the vision of Viksit Bharat.

EVENT 7:
CADEC: CAREER DEVELOPMENT CENTRE
BizBlitz1.0: Intra College B-Plan Competition

BizBlitz 1.0, an Intra College B-Plan Competition organized by the Career Development Centre (CADEC), was another resounding success, fostering entrepreneurial spirit among students through engaging discussions and compelling business pitches.

Launch of CADEC's, Annual Magazine

EVENT 8:
Ecolligence 2024
Viksit Bharat@2047: Challenges & Roadmap

Ecolligence 2024 provided an insightful exploration of economic challenges and opportunities framed within the vision of Viksit Bharat. This event, marked by engaging discussions and spirited competitions, equipped participants with a comprehensive understanding of the economic landscape and the strategies needed to address future challenges.

EVENT 9:

Dr B.R. Ambedkar Society

विकसित भारत की संकल्पना में डॉ. अम्बेडकर के विचारों की प्रासंगिकता

We paid tribute to Dr B.R. Ambedkar’s contributions by hosting a lecture series that highlighted the relevance of his thoughts in the context of Viksit Bharat. This initiative aimed to educate students on the importance of Dr Ambedkar's ideas in nation-building and social justice.

EVENT 10:
Lecture Series

विकसित भारत में मीडिया : वर्तमान परिदृश्य और चुनौतियाँ

A dedicated event examined the role of media in a digital age that addressed issues like global influence, social media impact, and commercialization, highlighting the rapid development of media and the ethical dilemmas it presents.

EVENT 11:
Viksit Bharat@2047: Sankalp se Siddhi
Policy Hackathon Competition

The Policy Hackathon Competition held on April 18th, 2024, showcased the collective ingenuity and dedication of over 400 attendees in proposing innovative solutions to address pressing social issues. With esteemed guests Advocate Sarita Choudhary and Shri Shyam Jaju imparting visionary insights on 'Viksit Bharat @2047: Sankalp se Siddhi', and Prof Rajni Abbi's inspiring video message, the event served as a beacon of hope for a brighter future. The organizers extend heartfelt gratitude to all participants, guests, and collaborators for their invaluable contributions, marking the event as an unforgettable success and fuelling the momentum towards a more prosperous tomorrow.


P. G. D. A. V. COLLEGE (M)
UNIVERSITY OF DELHI
VIKSIT BHARAT@2047: SANKALP SE SIDDHI
GPS Map Camera
New Delhi, Delhi, India
16/101 Nurses Residential Complex, Block G, Sri Niwaspuri, New Delhi, Delhi 110065, India
Lat 28.567449°
Long 77.253566°
18/04/24 12:09 PM GMT +05:30
Google


GPS Map Camera
New Delhi, Delhi, India
88, Vinoba Puri, Lajpat Nagar II, Lajpat Nagar, New Delhi, Delhi 110024, India
Lat 28.567432°
Long 77.251067°
18/04/24 12:23 PM GMT +05:30
Google


P. G. D. A. V. COLLEGE (M)


P. G. D. A. V. COLLEGE (M)


P. G. D. A. V. COLLEGE (M)

EVENT 12:
University of Delhi & Viksit Bharat@2047 Sankalp Club
RUN FOR VIKSIT BHARAT

Our commitment to fostering a sense of community and shared purpose was further demonstrated through the "Run for Viksit Bharat" event. The success of this event was due to the unified participation of community members, including our college, showcasing a collective dedication to the cause of Viksit Bharat. Our college's active involvement in promoting the event significantly contributed to its success, driving participation and awareness among a wider audience.

These diverse and impactful events have significantly influenced our community, fostering unity, intellectual growth, and a shared commitment to the vision of Viksit Bharat. Through these initiatives, we have not only advanced our educational mission but also contributed meaningfully to the broader national agenda of building a developed and progressive India.

Dr Gitu Nijhawan
Nodal Officer/Convenor
Viksit Bharat@2047 Sankalp Club: Sankalp se Siddhi

AN ANALYSIS ON STATE WISE INFLATION IN INDIA

INTRODUCTION

"Inflation poses a serious threat to the growth momentum. Whatever be the cause, the fact remains that inflation is something which needs to be tackled with great urgency …"

- [Dr. Manmohan Singh, Prime Minister of India, February 4, 2011, New Delhi]
Title: "Quarterly Inflation in India"

Aim:

The study aims to understand the factors influencing inflation in India for different states.

Methodology:

Our comprehensive analysis leverages the robust capabilities of R software to provide a nuanced understanding of various measures of tendencies and dispersions. Furthermore, our approach incorporates regression analysis to offer a detailed insight into state-wise inflation dynamics. This meticulous methodology ensures clarity and precision, facilitating a thorough comprehension of the underlying trends and patterns.

Overall, the research contributes to the understanding of inflation behavior in India and provides valuable implications for policymakers and economists studying inflation dynamics in emerging market economies.

Inflation, or a rise in the general price level of goods and services, has long been a subject of concern and debate around the world. It affects everything from the cost of groceries and housing to the economy's overall health. As the nation gears up for the upcoming general elections, it becomes crucial to delve into inflation in India to better understand the country's economic situation.

According to the data released by the National Statistics Office (NSO), India's retail inflation rate dropped to 5.09 percent in February 2024, the lowest in four months. The retail inflation rate registered a slight drop of 0.01 percent in one month, as it stood at 5.10 percent in January 2024. The current inflation remains in the Reserve Bank of India (RBI) tolerance band, which is set at 2 to 6 percent.

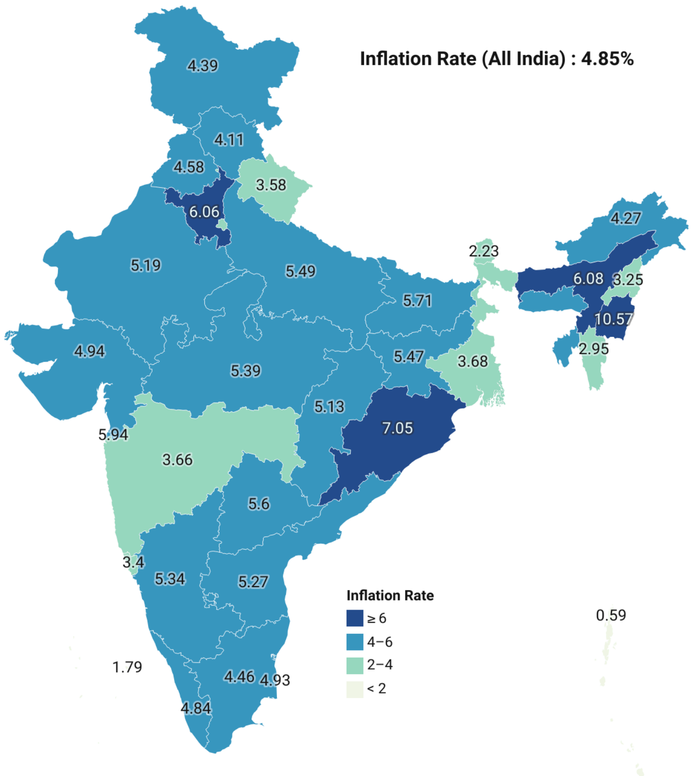
Inflation has detrimental effects on the economy as it diminishes the purchasing power of individuals over time. It leads to a sustained increase in the prices of goods and services, reducing the affordability of essential items.

The Government of India establishes the inflation target in India in accordance with the Reserve Bank of India (RBI). While the government sets an inflation target once every five years, the authority responsible for controlling inflation through monetary policies lies with the Reserve Bank of India.

The inflation rate in the rural areas, at 5.34 percent, remains 0.56 percent higher than the urban areas (4.78 percent), as inflation in the rural areas stood at 4.34 percent in December. The food inflation in February stood at 8.66 percent, a rise from 8.3 percent previous month.

Over the last decade, inflation has emerged as a leading concern of India's economic policymakers and citizens. Worries grew as the inflation rate (measured as the twelve-month change in the consumer price index) rose from 3.7% to 12.1% over 2001-2010.

The inflation rate has since fallen to 5.2% in early 2015, leading to a debate about whether this moderation is likely to endure or inflation will rise again.



Inflation rate shown against Jammu & Kashmir pertains to combined Union Territories of Jammu & Kashmir and Ladakh; Inflation rate pertains to Arunachal Pradesh (Rural) ; Map not to scale

Source: MoSPI, India. Prepared by Computer Centre @GoIStats • Created with Datawrapper

INDIA INFLATION RATE (CPI) - Historical data

The Consumer Price Index (CPI) measures changes in the average price level of goods and services purchased by households over time. It is used to measure inflation and indicates the cost of living for consumers. CPI is calculated by selecting a basket of goods and services that represent typical consumer purchases and tracking the changes in their prices over time. The index is often used to adjust wages, pensions, and government benefits to account for changes in purchasing power.

The Wholesale Price Index measures changes in the average price level of goods traded in bulk or at the wholesale level. It is primarily used as an indicator of inflation in the production and distribution stages of the economy. WPI tracks the price changes of goods before they reach the retail level and includes commodities such as raw materials, intermediate goods, and finished goods. It is often used by policymakers, businesses, and analysts to monitor inflationary pressures within the economy and make decisions based on the price trends in the wholesale market

Calculating the inflation rate involves the use of a formula:

((B - A) / A) x 100

where A represents the initial value, and B represents the final value.
To use this formula, you need the initial and final values of the consumer price index (CPI) for a specific good or service. By subtracting the initial value from the final value, you determine the difference between the two numbers.
This difference indicates the increase in the CPI for that specific good or service. To find the inflation rate, divide the difference by the initial value (the value recorded for the past date) to obtain a decimal figure.
To express this decimal as a percentage, multiply it by 100. The resulting number represents the inflation rate.

Since 2013, India's state-wise inflation rates have exhibited a kaleidoscope of patterns, reflecting the complex economic dynamics prevailing across different regions. States endowed with flourishing industrial and service sectors, such as Maharashtra and Karnataka, often confront higher inflation rates due to heightened demand pressures spurred by urbanization and economic activities. Conversely, agrarian states like Punjab and Uttar Pradesh may contend with inflationary challenges stemming from supply-side constraints, such as erratic weather conditions and infrastructural deficiencies.

States	Inflation Rate
Andhra Pradesh	5.27
Arunachal Prades	4.27
Assam	6.08
Bihar	5.71
Chhattisgarh	5.13
Goa	3.4
Gujarat	4.94
Haryana	6.06
Rajasthan	5.19
Himachal Pradesh	4.11
Jharkhand	5.47
Karnataka	5.34
Kerala	4.84
Madhya Pradesh	5.39
Maharashtra	3.66
Manipur	10.57
Mizoram	2.95
Nagaland	3.25
Orissa	7.05
Punjab	4.58
Sikkim	2.23
Tamil Nadu	4.46
Telangana	5.6
Uttar Pradesh	5.49
West Bengal	3.68
Jammu and Kash	4.39
Uttrakhand	3.58
Puducherry	4.93

Moreover, the impact of government policies looms large over state-level inflation trajectories. Subsidy regimes, tax structures, and regulatory measures adopted by state administrations wield significant influence, shaping the cost structures and price movements within their respective territories.

For instance, states offering substantial agricultural subsidies may witness tempered inflation rates in food items, while those with high tax burdens on essential commodities might experience inflationary spikes.

Understanding these nuanced dynamics is imperative for policymakers, enabling them to tailor interventions that address localized inflationary pressures effectively.

By crafting region-specific strategies, governments can mitigate inflation volatility, foster equitable economic development, and enhance the overall welfare of their constituents. Additionally, businesses and investors can leverage this granular understanding to devise targeted market strategies, harnessing opportunities and mitigating risks inherent in diverse state-level economic landscapes.

II METHODOLOGY

Our comprehensive analysis leverages the robust capabilities of R software to provide a nuanced understanding of various measures of tendencies and dispersions. Furthermore, our approach incorporates regression analysis to offer a detailed insight into state-wise inflation dynamics. This meticulous methodology ensures clarity and precision, facilitating a thorough comprehension of the underlying trends and patterns.

In order to investigate the relationship between inflation rates and geographical variations across states within a single year, a regression analysis was conducted using RStudio. Regression analysis was chosen as it allows for the examination of how changes in one variable, such as state-specific economic indicators, relate to changes in another variable, namely inflation rates. By employing this statistical technique, we aimed to identify any significant predictors or factors influencing inflation variability among states, providing valuable insights into regional economic dynamics

The following codes were formed to run the regression:

First:

To install necessary packages- In R in order to run our further libraries.

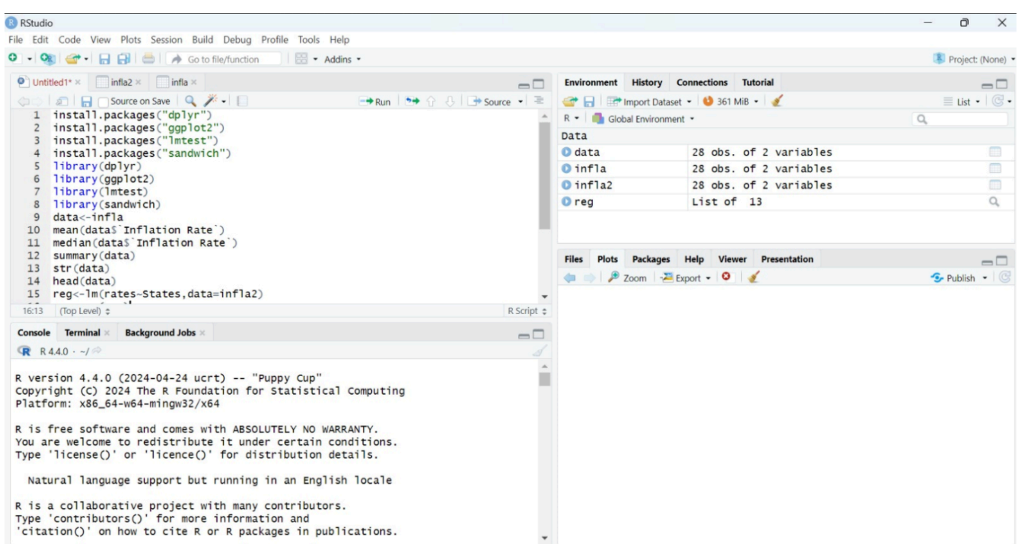
To load necessary libraries- In R in order to run our further commands.

Then giving codes:

For data manipulation

For data visualization

For robust standard errors

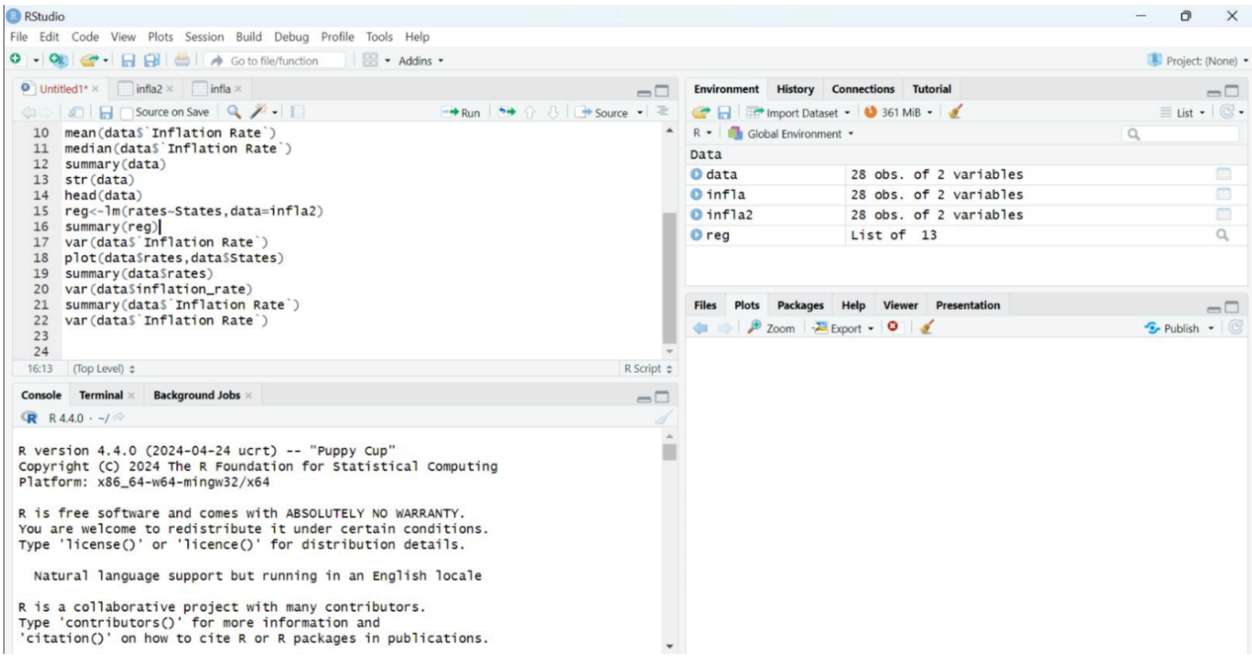


Then giving codes:

To load the dataset- for further research.

To check the structure of the dataset- to understand the dataset better using “str” command.

To explore the first few rows of the dataset- using “head” command.



Finding mean, median, variance to understand the economic situation better:

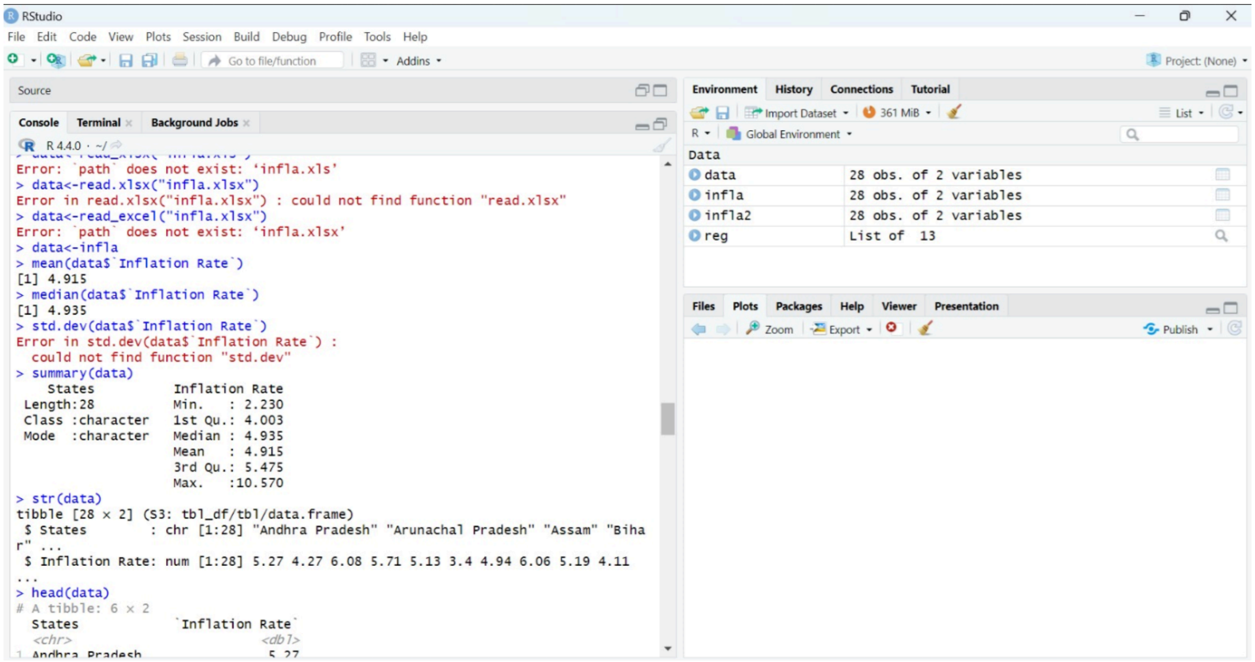
To find the mean of the inflation rates

To find the median of the Inflation rates

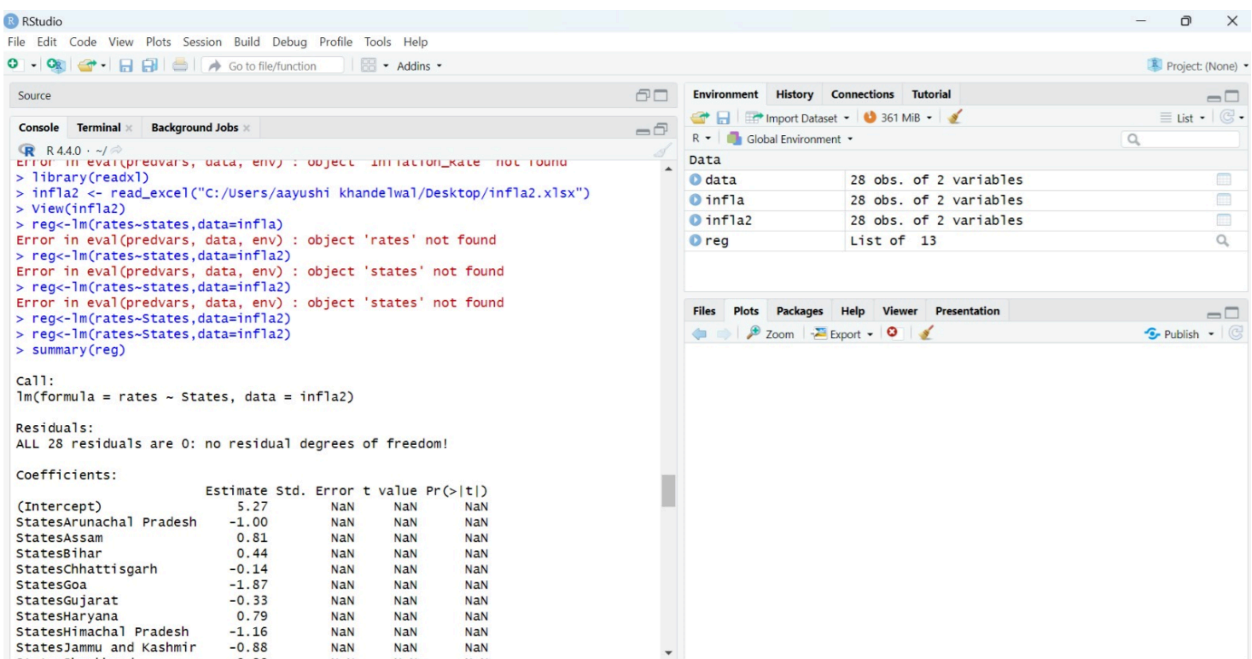
To find the variance of the data

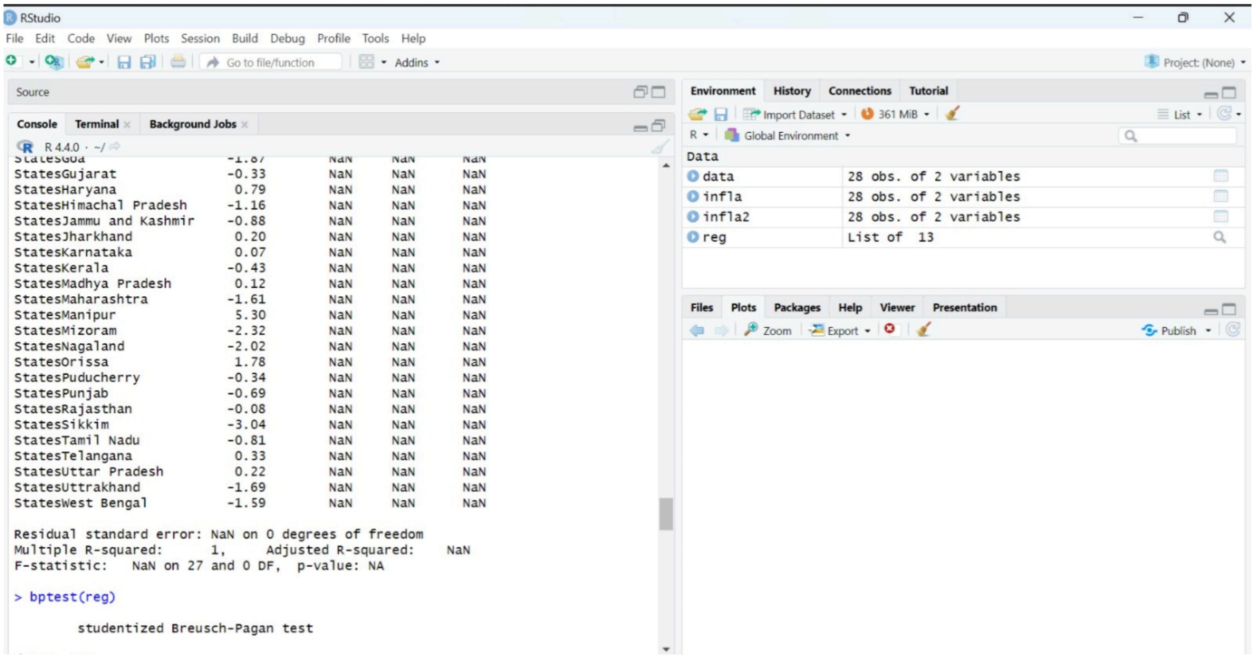
Then performing the:

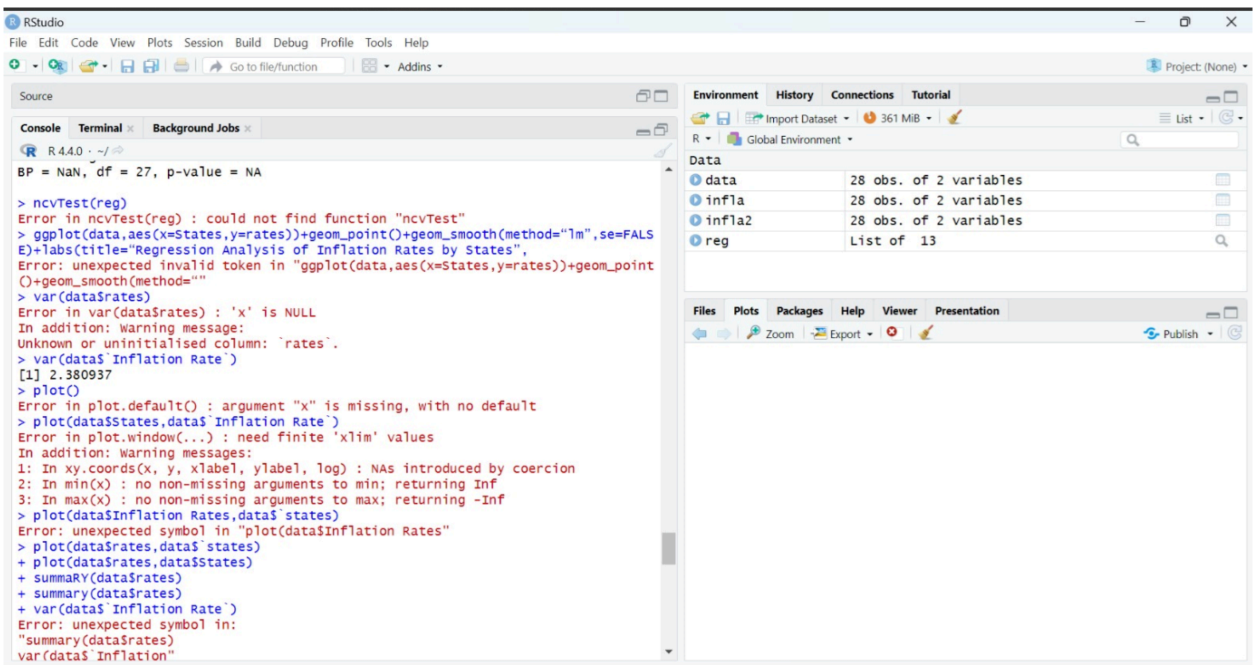
#linear regression- These codes will fit a linear regression model with “Inflation Rate” as the dependent variable and “States” as the independent variable.

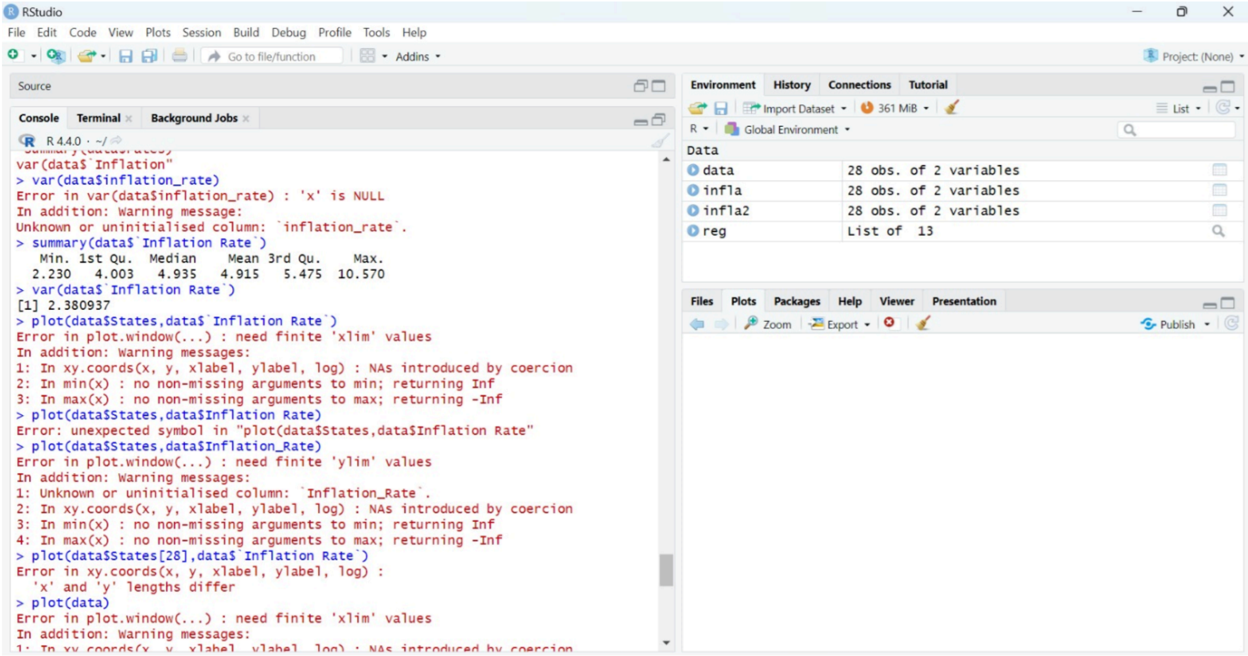


To print the summary of the regression model- In order to find out the minimum, maximum, mean, median, 1st Qu and 3rd Qu.
To plot the regression line




RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Source
Console Terminal Background Jobs
R 4.4.0 · ~/
StatesGujarat -0.33 NaN NaN NaN
StatesHaryana 0.79 NaN NaN NaN
StatesHimachal Pradesh -1.16 NaN NaN NaN
StatesJammu and Kashmir -0.88 NaN NaN NaN
StatesJharkhand 0.20 NaN NaN NaN
StatesKarnataka 0.07 NaN NaN NaN
StatesKerala -0.43 NaN NaN NaN
StatesMadhya Pradesh 0.12 NaN NaN NaN
StatesMaharashtra -1.61 NaN NaN NaN
StatesManipur 5.30 NaN NaN NaN
StatesMizoram -2.32 NaN NaN NaN
StatesNagaland -2.02 NaN NaN NaN
StatesOrissa 1.78 NaN NaN NaN
StatesPuducherry -0.34 NaN NaN NaN
StatesPunjab -0.69 NaN NaN NaN
StatesRajasthan -0.08 NaN NaN NaN
StatesSikkim -3.04 NaN NaN NaN
StatesTamil Nadu -0.81 NaN NaN NaN
StatesTelangana 0.33 NaN NaN NaN
StatesUttar Pradesh 0.22 NaN NaN NaN
StatesUttrakhand -1.69 NaN NaN NaN
StatesWest Bengal -1.59 NaN NaN NaN
Residual standard error: NaN on 0 degrees of freedom
Multiple R-squared: 1, Adjusted R-squared: NaN
F-statistic: NaN on 27 and 0 DF, p-value: NA
> bptest(reg)
studentized Breusch-Pagan test
Environment History Connections Tutorial
Import Dataset 361 MiB List
Global Environment
Data
data 28 obs. of 2 variables
infla 28 obs. of 2 variables
infla2 28 obs. of 2 variables
reg List of 13
Files Plots Packages Help Viewer Presentation
Zoom Export Publish


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Source
Console Terminal Background Jobs
R 4.4.0 · ~/
BP = NaN, df = 27, p-value = NA
> ncvTest(reg)
Error in ncvTest(reg) : could not find function "ncvTest"
> ggplot(data,aes(x=States,y=rates))+geom_point()+geom_smooth(method="lm",se=FALS
E)+labs(title="Regression Analysis of Inflation Rates by States",
Error: unexpected invalid token in "ggplot(data,aes(x=States,y=rates))+geom_point
()+geom_smooth(method=""
> var(data$rates)
Error in var(data$rates) : 'x' is NULL
In addition: Warning message:
Unknown or uninitialised column: `rates`.
> var(data$`Inflation Rate`)
[1] 2.380937
> plot()
Error in plot.default() : argument "x" is missing, with no default
> plot(data$States,data$`Inflation Rate`)
Error in plot.window(...) : need finite 'xlim' values
In addition: Warning messages:
1: In xy.coords(x, y, xlabel, ylabel, log) : NAs introduced by coercion
2: In min(x) : no non-missing arguments to min; returning Inf
3: In max(x) : no non-missing arguments to max; returning -Inf
> plot(data$Inflation Rates,data$`states)
Error: unexpected symbol in "plot(data$Inflation Rates"
> plot(data$rates,data$`states)
+ plot(data$rates,data$States)
+ summaRY(data$rates)
+ summary(data$rates)
+ var(data$`Inflation Rate`)
Error: unexpected symbol in:
"summary(data$rates)
var(data$`Inflation"
Environment History Connections Tutorial
Import Dataset 361 MiB List
Global Environment
Data
data 28 obs. of 2 variables
infla 28 obs. of 2 variables
infla2 28 obs. of 2 variables
reg List of 13
Files Plots Packages Help Viewer Presentation
Zoom Export Publish


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Source
Console
Terminal
Background Jobs
R 4.4.0 · ~/
var(data$`Inflation"
> var(data$inflation_rate)
Error in var(data$inflation_rate) : 'x' is NULL
In addition: Warning message:
Unknown or uninitialised column: `inflation_rate`.
> summary(data$`Inflation Rate`)
Min. 1st Qu. Median Mean 3rd Qu. Max.
2.230 4.003 4.935 4.915 5.475 10.570
> var(data$`Inflation Rate`)
[1] 2.380937
> plot(data$States,data$`Inflation Rate`)
Error in plot.window(...) : need finite 'xlim' values
In addition: Warning messages:
1: In xy.coords(x, y, xlabel, ylabel, log) : NAs introduced by coercion
2: In min(x) : no non-missing arguments to min; returning Inf
3: In max(x) : no non-missing arguments to max; returning -Inf
> plot(data$States,data$Inflation Rate)
Error: unexpected symbol in "plot(data$States,data$Inflation Rate"
> plot(data$States,data$Inflation_Rate)
Error in plot.window(...) : need finite 'ylim' values
In addition: Warning messages:
1: Unknown or uninitialised column: `Inflation_Rate`.
2: In xy.coords(x, y, xlabel, ylabel, log) : NAs introduced by coercion
3: In min(x) : no non-missing arguments to min; returning Inf
4: In max(x) : no non-missing arguments to max; returning -Inf
> plot(data$States[28],data$`Inflation Rate`)
Error in xy.coords(x, y, xlabel, ylabel, log) :
'x' and 'y' lengths differ
> plot(data)
Error in plot.window(...) : need finite 'xlim' values
In addition: Warning messages:
Environment
History
Connections
Tutorial
Import Dataset
361 MiB
List
Global Environment
Data
data 28 obs. of 2 variables
infla 28 obs. of 2 variables
infla2 28 obs. of 2 variables
reg List of 13
Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export
Publish

INTERPRETATION

State Level Inflation Rates

The data shows that inflation rates vary across Indian states. Manipur has the highest inflation rate at 10.57%, followed by Odisha at 7.05% and Assam at 6.08%.

On the other hand, Mizoram has the lowest inflation rate at 2.95%, followed by Sikkim at 2.23% and Delhi at 2.23%.

There are a number of factors that can contribute to these variations, including differences in state economies, agricultural production, and government policies.

For example, states with a high proportion of agricultural production may be more vulnerable to fluctuations in food prices. States with a large manufacturing sector may be more affected by changes in global commodity prices.

Government policies, such as taxes and subsidies, can also affect inflation rates.

National Inflation Rate

According to the MoSPI, the inflation rate in India was 4.85% in March 2024 https://www.mospi.gov.in/cpi.

This suggests that some states in the data set have inflation rates that are higher than the national average, while others have inflation rates that are lower than the national average.

Possible Causes of Inflation in India:

There are a number of factors that can contribute to inflation in India, including:

<u>Supply chain disruptions</u>:The COVID-19 pandemic has caused disruptions to global supply chains, which has led to higher prices for goods and services.

<u>Rising global commodity prices</u>:The prices of oil, food, and other commodities have been rising in recent months, due to factors such as the war in Ukraine and increased demand from China.

Weaker rupee: The Indian rupee has weakened against the US dollar in recent months, which makes imports more expensive.

Domestic factors:Domestic factors such as rising transportation costs, higher input costs for businesses, and increased government spending can also contribute to inflation.

Impacts of Inflation

Inflation can have a number of negative impacts on the Indian economy, including:

Reduced purchasing power:Inflation can erode the purchasing power of households, as their money does not go as far as it used to. This can lead to a decline in living standards.

Uncertainty for businesses: Inflation can create uncertainty for businesses, as it makes it difficult to plan for the future. This can lead to businesses delaying investment and hiring.

Lower economic growth:High inflation can lead to lower economic growth, as it discourages investment and consumption.

Government Policies to Control Inflation

The Indian government has a number of policies in place to control inflation, including:

Monetary policy:The Reserve Bank of India (RBI) is the central bank of India and is responsible for monetary policy. The RBI can use monetary policy tools such as interest rates to influence inflation.

Fiscal policy: The government can also use fiscal policy tools such as taxes and spending to influence inflation.

Supply-side measures: The government can also take steps to increase the supply of goods and services in the economy, which can help to reduce inflation.

III CONCLUSION

Inflation, often considered an indicator of economic health, has been a persistent concern in India's economic landscape. This analysis offers a comprehensive examination of inflation trends across different states in India, identifying key factors contributing to these variations and outlining the broader implications for policymakers.

State-Wise Variations in Inflation

The state-wise analysis revealed significant disparities in inflation rates across India. While some states like Manipur (10.57%) and Odisha (7.05%) recorded higher inflation rates, others like Mizoram (2.95%) and Sikkim (2.23%) experienced much lower rates. These disparities suggest a complex interplay of factors such as regional economic dynamics, agricultural productivity, government policies, and external influences.

States with robust industrial and service sectors often encounter higher inflation due to increased demand, driven by urbanization and economic activities. On the other hand, states with a high agricultural base are more susceptible to fluctuations in food prices due to factors like erratic weather and infrastructural deficiencies. Government policies also play a significant role in shaping state-level inflation trajectories, with tax structures, subsidies, and regulatory measures impacting cost structures and price movements.

National Inflation Trends

At the national level, India's inflation rate has shown a downward trend in recent years, dropping to 4.85% in March 2024. However, the variations between states highlight the need for a nuanced approach to inflation management. While some states consistently maintain rates below the national average, others struggle with higher inflation, indicating the complexity of addressing inflation at a macro level.

The causes of inflation in India are multifaceted, including supply chain disruptions, rising global commodity prices, a weakening rupee, and domestic factors such as transportation costs and increased government spending. These factors contribute to a challenging environment for policymakers seeking to maintain stability while fostering economic growth.

Impacts of Inflation

Inflation's impacts on the Indian economy are profound. It reduces the purchasing power of households, leading to a decline in living standards and creating uncertainty for businesses, which can deter investment and hiring. High inflation can also slow economic growth, further exacerbating the challenges faced by the government in fostering a robust economy.

The research highlights the need for targeted policy interventions to address these impacts. Policymakers must strike a balance between controlling inflation and encouraging economic growth. The risk of high inflation discouraging investment and consumption must be managed through careful monetary and fiscal policies.

POLICY IMPLICATIONS

Given the complex factors contributing to inflation, policymakers must adopt a multi-faceted approach to manage inflation effectively. The following policy implications emerge from this analysis:

1. Monetary Policy: The Reserve Bank of India (RBI) plays a crucial role in controlling inflation through monetary policy. The central bank should continue to use tools like interest rates and reserve requirements to influence money supply and inflation.
2. Fiscal Policy: The government can also influence inflation through fiscal measures such as taxation and government spending. By adjusting these levers, policymakers can manage inflationary pressures while supporting economic growth.
3. Supply-Side Measures: Addressing inflation requires an increase in the supply of goods and services. Policymakers should focus on reducing supply-side constraints by investing in infrastructure, encouraging agricultural productivity, and facilitating industrial growth.
4. Targeted Subsidies: The government should consider targeted subsidies for essential goods to manage inflation in key sectors. However, indiscriminate subsidies can lead to distortions, so careful targeting is essential.
5. Regional Strategies: Since inflation varies across states, a one-size-fits-all approach may not be effective. Policymakers should develop region-specific strategies that address local inflationary pressures, promoting equitable economic development.

In conclusion, India's inflation trends reflect a complex landscape influenced by regional variations, global factors, and domestic policies. The analysis underscores the need for a balanced approach to managing inflation, focusing on both monetary and fiscal measures while addressing supply-side constraints. Policymakers must navigate these challenges to foster economic stability, encourage growth, and enhance the overall welfare of the Indian population. The study's insights provide a valuable roadmap for crafting effective policies to tackle inflation and ensure sustainable economic development in India.

REFERENCES

- Ministry of Statistics and Programme Implementation (MoSPI) https://www.mospi.gov.in/cpi
- Reserve Bank of India (RBI) https://www.rbi.org.in/
- Economic Survey, 2013, Ministry of Finance, Government of India.
- forbesindia.com
- https://www.rbi.org.in
- https://www.mospi.gov.in/
- https://economictimes.indiatimes.com
- https://www.livemint.com/
- https://www.worldbank.org/
- https://www.imf.org/

Research paper

Title: Impact of urbanization on economic growth and development

Authors: Rahimeen Siddiqui(12741)

Inaya shahid(12709)

Astha Srivastava (12743)

Submitted to: Ms Rimpy kaushal

Date: April25,2024

Index

Abstract

Urbanization has been a prevailing global trend, profoundly influencing economic growth and development. This paper examines the impact of urbanization on economic growth and development, employing a comprehensive review of existing literature and empirical evidence.

Firstly, urbanization is a catalyst for economic growth due to the concentration of resources, infrastructure, and markets. Cities offer agglomeration economies, fostering innovation, productivity, and efficiency. The proximity of businesses, suppliers, and customers enhances specialization and economies of scale, driving economic growth. Moreover, urban areas attract a skilled workforce, promoting human capital development, which is vital for sustained economic advancement.

Secondly, urbanization plays a crucial role in poverty alleviation and the improvement of living standards. Cities act as centers for

employment opportunities, providing income-generating activities, and facilitating access to essential services such as healthcare, education, and sanitation. However, the benefits of urbanization are not uniformly distributed, and it can exacerbate inequality if not managed properly.

Thirdly, the impact of urbanization on economic growth and development is contingent upon effective urban planning and governance. Proper infrastructure investment, including transportation, housing, and sanitation, is crucial for optimizing the benefits of urbanization.

The present paper attempts to examine the relationship between urbanization and economic growth in India at the state level from 1971 to 2020 by employing a bootstrap panel Granger causality test. It is found that in India the majority of the states display a unidirectional Granger causality from economic growth to urbanization. This finding indicates not only the lower propulsive power of urban centers but also an unbalanced development of SOC between urban centers and rural areas, hence causing a migration of people to cities with a rise in their income to take advantage of urban facilities. This study also proposes some policy implications for future research.

Keywords: urbanization,economic development,Resources,infrastructure

Introduction

Objective:

The objective of examining the impact of urbanization on economic growth and development is multifaceted. Urbanization, the process by which people move from rural areas to cities, can significantly influence a country's economic trajectory. Cities often serve as hubs for innovation, entrepreneurship, and economic activity, driving productivity and attracting investment. The concentration of people and resources in urban areas can lead to economies of scale and scope, fostering specialization and diversification. Furthermore, urban centers typically offer better access to education, healthcare, and infrastructure, which can enhance human capital and labor productivity.

However, urbanization also presents challenges that can hinder economic growth and development. Rapid urbanization can strain infrastructure, leading to congestion, inadequate housing, and environmental degradation. Moreover, unequal distribution of urban benefits can exacerbate social and economic inequalities, potentially leading to social unrest.

Therefore, understanding the impact of urbanization requires a balanced approach that recognizes both its opportunities and challenges. Policymakers need to implement strategies that promote sustainable urban development, invest in infrastructure, and address social inequalities to harness the economic potential of cities effectively. By doing so, countries can ensure that urbanization contributes positively to economic growth, development, and overall well-being.

Urbanization is a complex phenomenon according to urban economists and is defined as the demographic process resulting in more expansion of the population in the urban area. According to the United Nations projection, in 2050, two-thirds of the global population will be present in urban areas. Many economists welcomed this urbanization as they believe that this process of urbanization improves the wealth of the nation and also improves economic growth. Hence,

the urbanization is correlated positively with the economic evolution (Friedberg

and Hunt (1995)). The last few years witnessed that an outbreak of " urbanization" had affected the economic growth of a country. Urbanization has an impact on the realization of census data, whole population is categorized into two forms Rural and Urban. Both forms have specific identities, Agriculture is primary in rural whereas Industrialization and modernization are most prominent in the urban setup, and nonagricultural setups (Manufacturing units or service structures) are more income-oriented or revenue-generating. Urbanization is the transformation of the workforce from the agriculture sector to non-agriculture sectors like manufacturing units or service structures. In other words, a huge number of the workforce is shifting from an agricultural economy to an industrial economy in urban areas.

Unemployment, overpopulation, and the backwardness of rural areas are the main reasons for urbanization. The labor force shifts from rural to urban setup due to the availability of good-wage work opportunities, better education and health care facilities, and good transportation service. In Pakistan, job opportunities, higher wages, the best education facilities, better quality health care, and the availability of other services are the main pulling factors for the migration of the population from rural to urban areas (Pakistan Bureau of Statistics 2017). The key determinants of large-scale migration in India are higher productivity in an urban area, more employment facilities, and a better lifestyle (Tripathi (2013)). Urbanization in developed countries is more than in developing countries due to advancements in technology and optimal use of resources. Australia, Kuwait, Israel, Qatar, Argentina, Venezuela, Singapore, Belgium, the United Kingdom, and Uruguay are the top ten countries presenting the concept of urbanization.

The South Asian region of the world is witnessing rapid urbanization. According to the World Bank (2016) report the concentration of rapid urbanization is concentrated in the South Asian part of the globe, especially Pakistan, India, and Bangladesh. It is projected that by 2050 the share of urbanization in the Asian region will be expected to increase by 64%. No doubt the increase in the population in the world become more challenging, especially in the Asian region, but it is also believed that the migration of people from rural to urban sites is due to the search for better job opportunities. Hence, people are involved in economic practices thereby contributing to economic growth.

A review of existing literature

Historically, the impact of urbanization on economic growth and development has been profound. Here are some key points based on historical data:

Industrial Revolution: In the 18th and 19th centuries, the Industrial Revolution led to rapid urbanization in Europe and North America. Cities became centers of manufacturing and trade, driving economic growth through increased production and commerce.

Urban Growth Post-WWII: After World War II, many countries experienced significant urbanization as people migrated to cities in search of employment opportunities. This urban migration fueled economic growth by expanding the labor force and consumer markets.

Asian Tigers: Countries like South Korea, Singapore, and Taiwan experienced rapid economic growth in the late 20th century, partly due to urbanization. These countries invested heavily in urban infrastructure and attracted foreign investments, leading to industrialization and economic development.

China's Urbanization: China's economic growth over the past few decades has been closely tied to its urbanization process. The government's urbanization policies, infrastructure investments, and industrialization efforts have transformed many cities, driving economic expansion and lifting millions out of poverty.

Challenges and Inequalities: Despite the economic benefits, urbanization has also led to challenges such as income inequalities, housing shortages, and environmental degradation. Managing these issues has been crucial for sustainable development.

Developing Countries: In many developing countries, urbanization is currently accelerating, with cities becoming engines of economic growth. However, infrastructure gaps, informal settlements, and urban poverty remain significant challenges that need to be addressed for sustainable urban development.

Overall, historical data shows that urbanization has played a critical role in driving economic growth and development. However, managing its impact effectively is essential to ensure inclusive and sustainable development for all segments of society.

Existing evidence

Urbanization's impact on economic growth and development in India has been the subject of extensive research.

Economic Growth: Urbanization often leads to increased productivity due to better infrastructure, access to markets, and economies of scale. According to World Bank data, India's GDP growth rate has often been higher in urban areas compared to rural areas.

Source: World Bank, "World Development Indicators"

Infrastructure Development: Urban areas attract more investment in infrastructure like roads, electricity, and sanitation. Government of India reports indicate increased investment in urban infrastructure projects.

Source: Ministry of Housing and Urban Affairs, Government of India

Employment Opportunities: Urbanization can lead to job creation in various sectors. As per the National Sample Survey Office (NSSO), urban areas have seen growth in service sector jobs.

Source: NSSO reports

Income Disparities: Urbanization can exacerbate income inequalities. The Gini coefficient, a measure of income inequality, has shown increasing trends in India.

Source: World Bank, "World Development Indicators"

Environmental Challenges: Rapid urbanization can lead to environmental degradation. Air and water pollution levels have risen in major Indian cities.

Source: Central Pollution Control Board (CPCB), Government of India

Housing and Slums: Urban growth often results in inadequate housing and the growth of slums. Census data indicates a significant proportion of the urban population living in slums.

Source: Census of India

Migration: Rural-to-urban migration contributes to urbanization. Census data shows increasing rural-to-urban migration rates.

Source: Census of India

These findings suggest that while urbanization can boost economic growth and development, it also brings challenges like income inequality, environmental degradation, and housing issues.

Policymakers need to address these challenges to ensure sustainable urban development in India.

Research gap

Urbanization is the process of shifting the rural population to the urban locality and the urban area becomes the center of economic activities. Hence due to urbanization, the labor force shifted from agriculture to the industrial and services sector in the urban area. This shift in population changes the social as well as the economic status of the urban area. More importantly, the rapid shift causes several other problems. Several problems due to the process of urbanization are identified by (Kuddus et al. (2020), Dodman (2016), and many more).

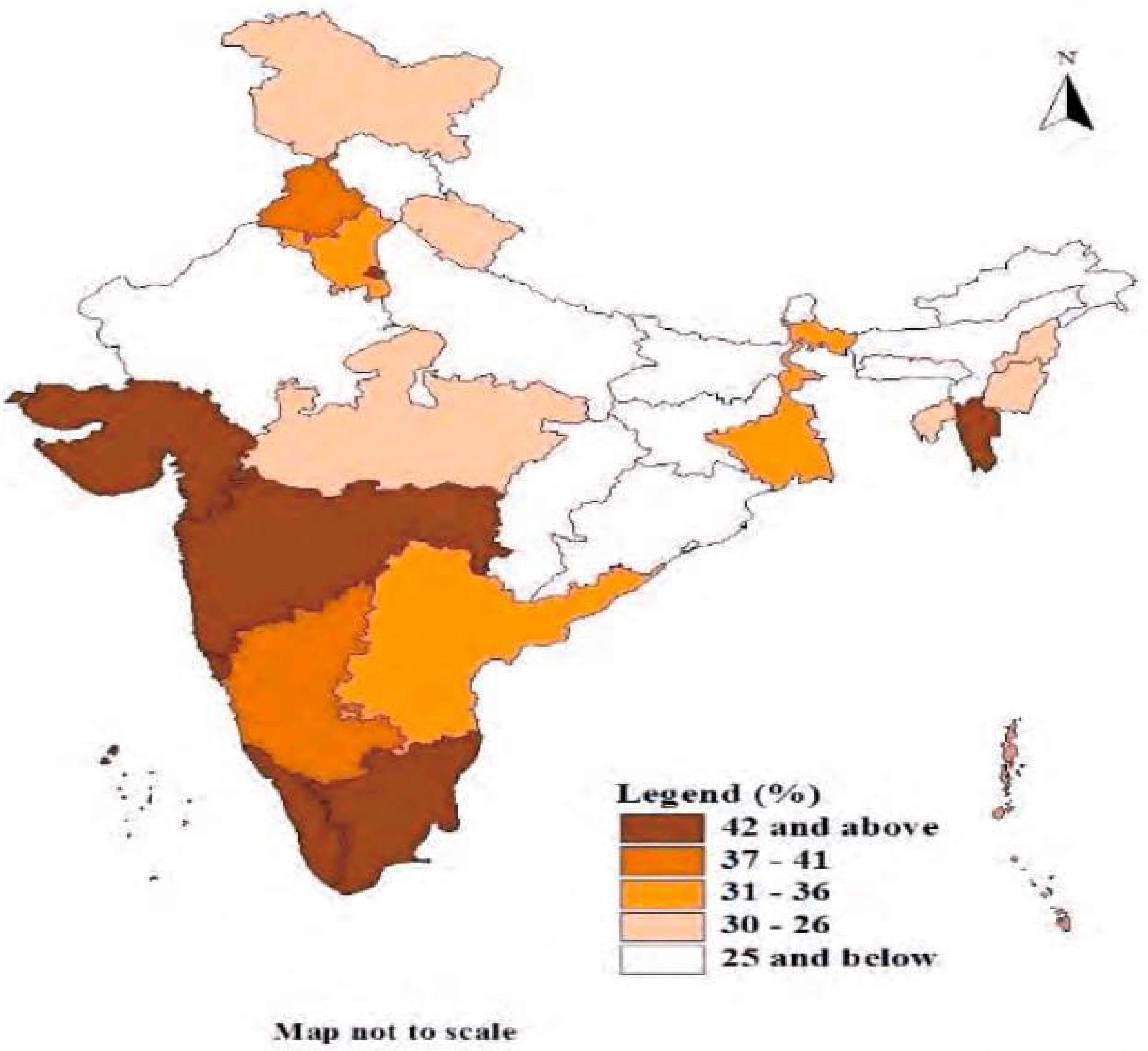
Several studies in the literature claimed that the process of urbanization is directly linked to economic growth (Mills et al (1986), Sarker et al. (2016)). However, there exist

many studies in the literature where the problems due to rapid urbanization are also highlighted (Jalil and Iqbal (2010)). Along with the connection between urbanization and economic stability, the identification of the factors responsible for urbanization is also important. In literature, several studies have been carried out to determine the underline factors, for, example, see (Tripathi and Mahey (2017)). It is noted and highlighted by many authors that in developing countries process of urbanization is high as compared to developed countries. Moreover in south Asian region of the world contains more developing countries i.e. Pakistan, Afghanistan, India, Bangladesh, Sri Lanka, Bhutan, and Nepal having highly urbanizing populations. Hence it is very logical to study the underlying factors of urbanization in those developing countries to study the effect of urbanization on economic growing is much needed one.

Analysis

At the state level, the pattern of urbanisation is very diverse, but economically advanced states show higher level of urbanisation. The emerging regional pattern is evident from Fig. 1 which shows that most parts of central, eastern and northeastern India has very low level of urbanisation. This region is also the economically less developed part of India. On the other hand, all southern states along with states of northern and western India such as Punjab, Haryana, Gujarat, and Maharashtra have higher urbanisation level than the national average, but the small states like Goa continues to top the list among states with 62 percent urban followed by Mizoram (51.5 percent).

Fig 1: Levels of Urbanisation, India



Among the major states, Tamil Nadu continues to be ahead of other states with level of urbanisation at 48.4 percent in 2011. The states which are lagging behind are Himachal Pradesh at the bottom with level of urbanisation at 10 percent followed by Bihar (11.3), Assam (14 percent) and Orissa (16.6). Other states like UP, Rajasthan, MP, Chhattisgarh and Jharkhand also continued to have lower urbanisation than the national level.

Although reversal in the declining trend in urban population growth rate at the national level is a major feature of urbanisation revealed by 2011 Census, there are only 15 states and UTs which show increased urban population growth rate during 2001-2011 compared to 1991-2001. Among them Kerala, Andhra Pradesh, Karnataka, Gujarat, West Bengal, Bihar, Jharkhand, Chhattisgarh and Uttarakhand are the major states. A very high urban population growth has occurred in the states of Kerala and Andhra Pradesh where urban population growth rate has increased to 6.5 percent per annum in Kerala and 3 percent per annum in Andhra Pradesh during 2001-11 compared to just about 1 percent per annum during 1991-2001. In both Kerala and Andhra Pradesh along with West Bengal and Gujarat, a large number of new towns have emerged as a result of rural urban classification in 2011.

Table 1:trends in urbanization in india,1195-2011

Census year	Urban Population (in million)	Percent urban	Annual exponential urban growth rate (%)
1961	78.94	17.97	-
1971	109.11	19.91	3.23
1981	159.46	23.34	3.79
1991	217.18	25.72	3.09
2001	286.12	27.86	2.75
2011	377.1	31.16	2.76

Table-1 shows that India has about 79 million urban population in 1961 which constituted about 18 percent of the total population. The average growth rate of urban population was 2.32 percent during 1951-61 which accelerated up to 3.79 percent during 1971-81 i.e. the highest urban growth since independence. After 1981, the urban growth rate decelerated to 3.09 percent during 1981-91 and further declined to 2.75 during 1991-2001. However, the declining growth rate was slightly reversed during 2001-2011. The total addition to urban population was 91 million during 2001-2011 - the highest ever and for the first time urban population increment was higher than rural increment (90.5 million) since a uniform definition was followed since 1961.

Table-2: Level of urbanization and urban growth in india and states, 2011

State/India	Urban Population (in million)	% Urban	Average Annual Urban Growth Rate*	Average annual rural growth rate*	Urban-rural growth differentials *
Andhra Pradesh	28.35	33.4	3.09	0.19	2.90
Arunachal Pradesh	0.31	22.6	3.18	2.07	1.01
Assam	4.38	14.0	2.43	1.41	1.02
Bihar	11.72	11.30	3.01	2.15	0.86
Chhattisgarh	5.93	23.2	3.49	1.65	2.84
Goa	0.90	62.1	3.01	-2.02	5.03
Gujarat	25.71	42.5	3.06	0.89	2.17
Haryana	8.82	34.7	3.66	0.99	2.67

Himachal Pradesh	0.68	10.0	1.45	1.17	0.28
Jammu & Kashmir	3.41	27.2	3.04	1.88	1.16
Jharkhand	7.92	24.0	2.79	1.79	1.00
Karnataka	23.57	38.5	2.72	0.75	1.97
Kerala	15.93	47.7	6.56	-3.00	9.56
Madhya Pradesh	20.05	27.6	2.28	1.70	0.58
Maharashtra	50.82	45.2	2.12	0.99	1.13
Manipur	0.82	30.2	3.55	0.43	3.12
Meghalaya	0.59	20.0	2.7	2.45	0.25
Mizoram	0.56	51.5	2.42	1.61	0.81
Nagaland	0.57	28.9	5.15	-1.5	6.65
Orissa	6.99	16.6	2.37	1.13	1.24
Punjab	10.38	37.4	2.28	0.76	0.52
Rajasthan	17.08	22.8	2.56	1.74	0.82
Sikkim	0.15	24.9	9.29	-0.52	9.81
Tamil Nadu	34.94	48.4	2.4	0.64	1.76
Tripura	0.96	26.1	5.65	0.23	5.42
Uttar Pradesh	44.47	22.2	2.52	1.64	0.88
Uttarakhand	3.09	30.5	3.49	1.07	2.42
West Bengal	29.13	31.8	2.61	0.74	1.87
Andaman & Nicobar Islands *	0.13	35.6	1.53	0.18	1.35
Chandigarh	1.02	97.2	1.3	-11.55	12.85
Dadra & Nagar	0.15	46.6	11.52	0.73	10.79
Daman & Diu	0.18	75.1	11.58	-5.12	16.70
Delhi	16.33	97.5	2.35	-8.31	10.66
Lakshadweep	0.05	78.0	6.23	-8.68	14.91
Pondicherry	0.85	68.3	2.71	1.91	0.80
India	**377.10**	**31.1**	**2.76**	**1.16**	**1.60**

As stated earlier, urbanisation is the product of urban-rural growth differentials.

Table-1 presents urban-rural growth differentials along with urban growth rate and

level of urbanisation (% urban). There exists a positive relationship between urban

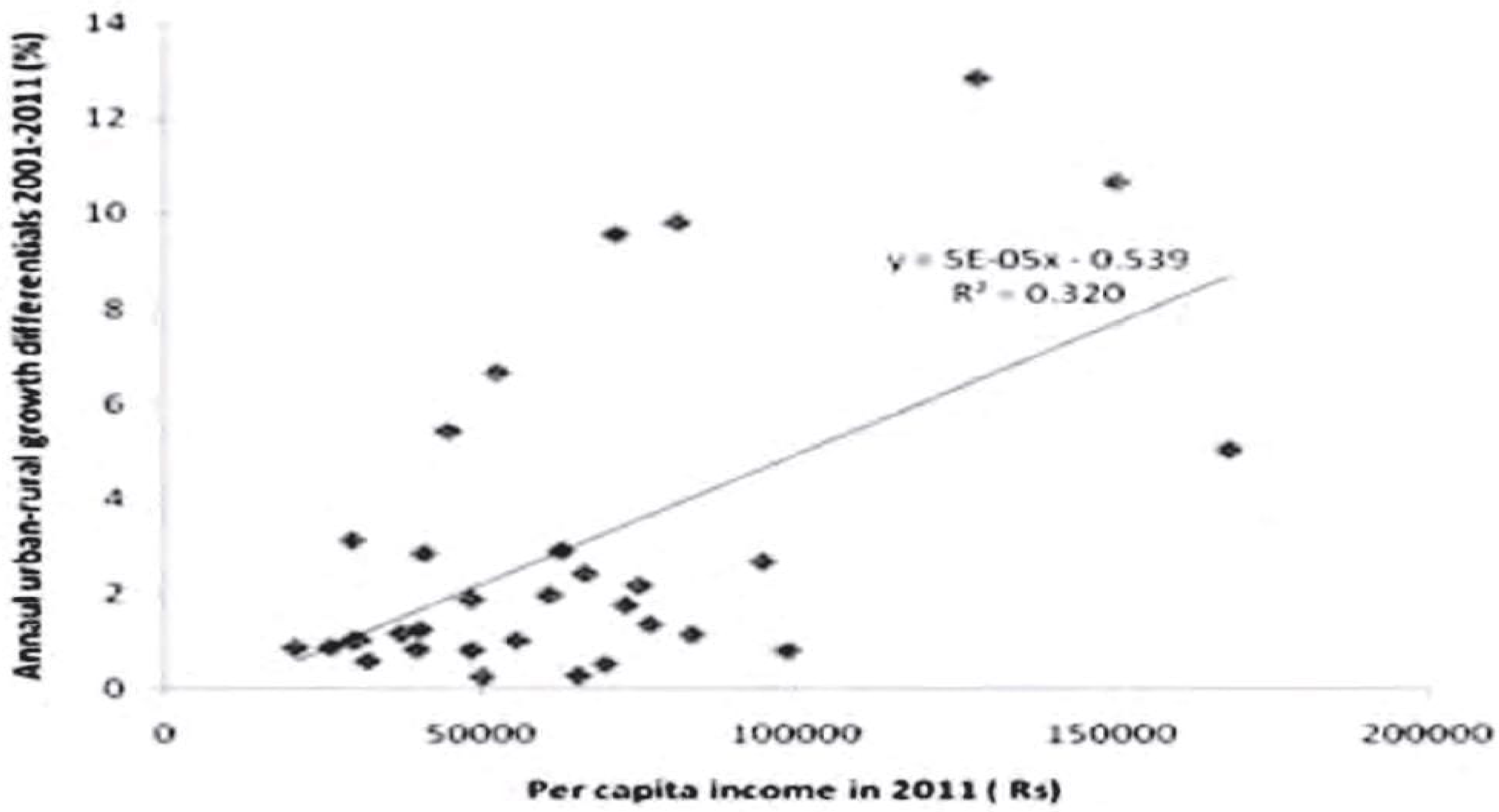
population growth rate and urban-rural growth differentials at the state level. further shows that the level of economic development contributes

positively in widening urban-rural growth differentials and thus contributing to the speed of urbanisation.

Relationship between urbanization and economic growth

Urbanisation in India: Trend, Pattem and Policy Issues

Per capita income and urban rural growth differentials at state level India



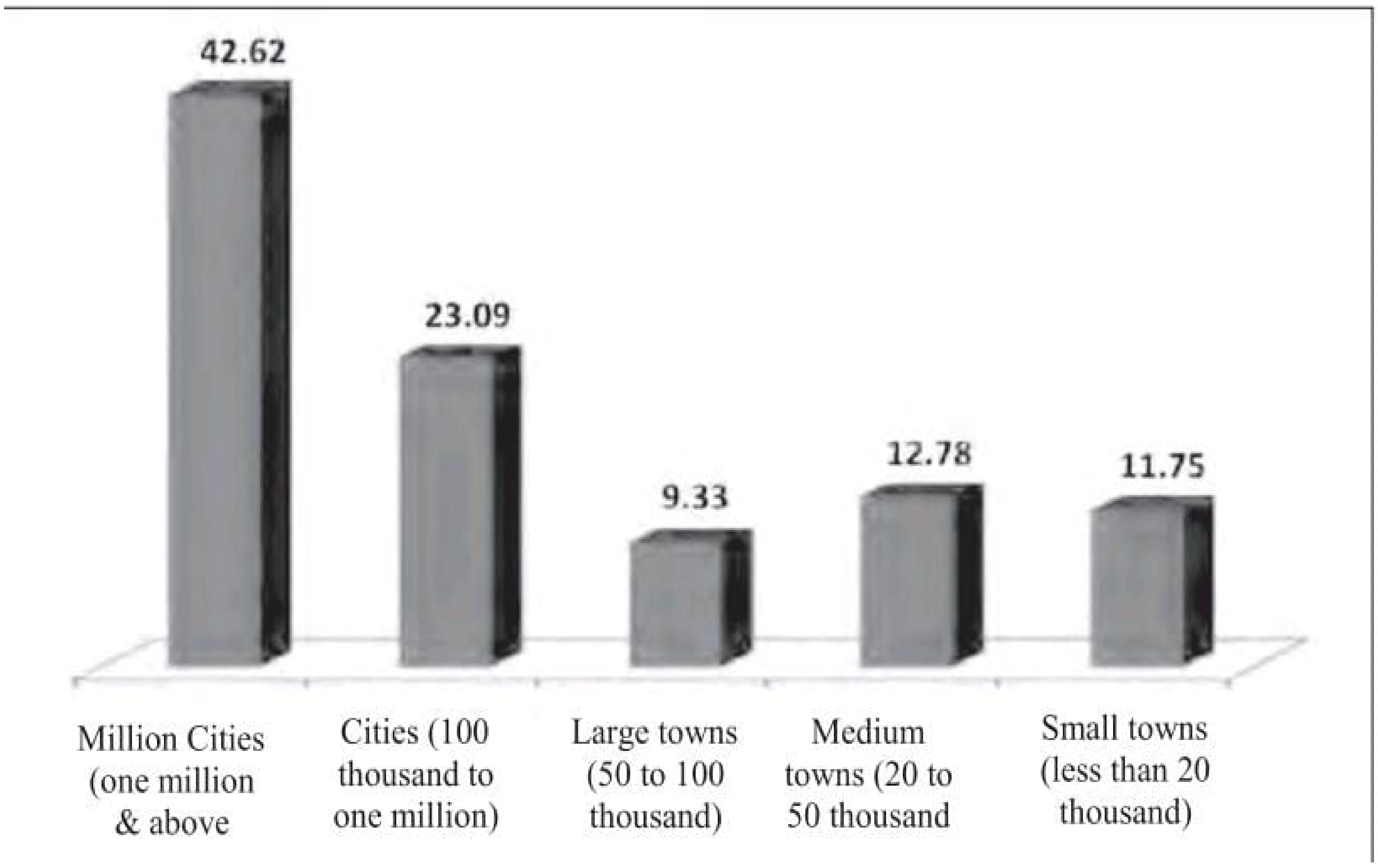
Various studies show that urbanisation has been closely related to economic development, and is the single most important factor in the organization of production and access to services. Cities are considered to be engines of economic growth and temples of modern civilization. Thus to know how our cities are growing assumes enormous significance for understanding the problems of economy and society.

Table-3 urban population by size class of cities and towns in India 1901-2011

Census year	Million Cities (one million & above)	Cities (100 thousand to one million)	Large Towns (50 to 100 thousand)	Medium Towns (20 to 50 thousand)	Small Towns (less than 20 thousand)
1901	5.86	20.11	11.29	15.64	47.10
1911	10.89	16.74	10.51	16.40	45.46
1921	11.30	18.40	10.39	15.92	43.99
1931	10.34	20.86	11.65	16.80	40.35
1941	12.19	26.04	11.42	16.35	34.00
1951	19.07	25.57	9.96	15.72	29.69
1961	23.34	28.08	11.23	16.94	20.41
1971	26.02	31.22	10.92	16.01	15.83
1981	26.93	33.49	11.63	14.33	13.62
1991	33.18	32.01	10.95	13.19	10.66
2001	37.80	30.78	9.73	12.29	9.36
2011	42.62	23.09	9.33	12.78	11.75

Source: various cencus reports

- **Percentage of urban population by size class of cities and towns India,2011**



Access to Basic Amenities by Size Class of Cities/Towns

According to 2011 Census, 55 per cent households in rural areas and 92 per cent of households in urban areas have access to electricity. So far the toilet facility is concerned, it was abysmally low in rural (30 per cent) compared to urban areas (81 per cent). Whereas about one-fifth of households do not have access to toilet facility in urban areas that means about 75 million urban populations have no access to toilet facility as per 2011 Census. Another aspect of sanitation closely associated with toilet facility is the wastewater outlet through the provision of drainage. The proportion of households either with open or closed drainage was 81 per cent in urban areas. Compared with toilet and drainage facility, access to drinking water provided either through tap or hand pumps was reported to be 74 per cent in rural areas compared to and 82 percent households in urban areas as per 2011 Census. Use of clean fuel is very important from health point of view. In rural areas, about one-tenth of households were found using LPG/PNG compared to three-fifths in the urban areas. This shows that a very high proportion (twofifths) of households was still using polluting fuels which are not only hazardous for. health but also contributes to greenhouse gases and global warming.

Urbanisation in India: Trend, Pattem and Policy Issues

India's urban population is distributed across 8000 odd towns and cities with different sizes, economic base and ability to generate resources from tax and nontax sources. Class I cities (100 thousand and more) have higher employment in organized sector compared to small urban centres. In many small urban centres, a sizeable proportion of workforce is also dependent on agriculture. Thus, size as a measure of urban centres not only reflects population concentration but also their economic strength as well. It is expected that the provision of basic services is directly related to the size of urban centres. Table 7 presents basic amenities by size class of urban centres. It confirms that except toilet facility all other amenities like electricity, drainage, LPG/PNG, etc., increases with increasing size class of cities and towns.

Model and results

data <- data.frame(

State_India = c("Andhra Pradesh", "Arunachal Pradesh", "Assam", "Bihar", "Chhattisgarh", "Goa", "Gujarat", "Haryana", "Himachal Pradesh", "Jammu & Kashmir", "Jharkhand", "Karnataka", "Kerala", "Madhya Pradesh", "Maharashtra", "Manipur", "Meghalaya", "Mizoram", "Nagaland", "Orissa", "Punjab", "Rajasthan", "Sikkim", "Tamil Nadu", "Tripura", "Uttar Pradesh" ,"Uttarakhand", "West Bengal", "Andaman & Nicobar Islands", "Chandigarh","Dadra & Nagar", "Daman & Diu", "Delhi", "Lakshadweep", "Pondicherry", "India"),

Urban_Population = c(28.35, 0.31, 4.38, 11.72, 5.93, 0.90, 25.71, 8.82, 0.68, 3.41, 7.92, 23.57, 15.93, 20.05, 50.82, 0.82, 0.59, 0.56, 0.57, 6.99, 10.38, 17.08, 0.15, 34.94, 0.96, 44.47, 3.09, 29.13, 0.13, 1.02, 0.15, 0.18, 16.33, 0.05, 0.85, 377.10),

Percent_Urban = c(33.4, 22.6, 14.0, 11.30, 23.2, 62.1, 42.5, 34.7, 10.0, 27.2, 24.0, 38.5, 47.7, 27.6, 45.2, 30.2, 20.0, 51.5, 28.9, 16.6, 37.4, 22.8, 24.9, 48.4, 26.1, 22.2, 30.5, 31.8, 35.6, 97.2, 46.6, 97.5, 78.0, 68.3, 31.1)

)

Fit linear regression model

model <- lm(Urban_Population ~ Urban_Population, Percent_Urban, data=data)

```
# Summary of the regression model
summary(model)

# Scatterplot of Urban Population vs. % Urban
plot(data$Percent_Urban, data$Urban_Population,
    xlab="% Urban", ylab="Urban Population (in million)",
    main="Urban Population vs. % Urban")
abline(model, col="red")
```

Conclusion

The declining trend in the urban population growth rate observed during 1980s and 1990s was reversed at the national level, and the level of urbanisation increased faster during 2001-2011. The urban population grew from 286 million in 2001 to 377 million in 2011- an increment of 91 million which is larger than the rural population increment of 90.5 million for the first time since independence. A substantial increase in urban population is contributed by net rural-urban classification and rural to urban migration. A huge number of new towns emerged during the last decade contributing significantly to the speeding up of urbanisation.

On the other hand, although the contribution of natural increase in urban growth has declined in terms of proportions, its share in absolute numbers (about 40 million) continues to be huge due to large base of the urban population. This has implications not only for providing the increased urban infrastructure and civic amenities, but also of the reproductive and child health services in urban areas.

Urban areas face acute shortage of civic amenities. In order to deal with the rapid increase in urban population and faster urbanisation, India has to push through several urban reforms and policy changes that have been initiated in the early 1990s. In India, urban development is a state subject; however Central Government used to provide guidelines and also promise increased funds through centrally initiated urban development programmes like Jawaharlal Nehru National Urban Renewal Mission (JNNURM) currently replaced by Smart Cities Mission and AMRUT (Altal Mission for Rejuvenation and Urban Transformation).

Urbanization has emerged as a critical factor influencing economic growth and development globally. This paper has explored the multifaceted impact of urbanization on economic growth and development, drawing upon a comprehensive review of existing literature and empirical evidence.

Urbanization significantly contributes to economic growth by concentrating resources, infrastructure, and markets, thus fostering innovation, productivity, and efficiency. Moreover, cities attract a skilled workforce, promoting human capital development, which is vital for sustained economic advancement. However, it is essential to acknowledge that the benefits of urbanization are not uniformly distributed and can exacerbate inequality if not managed properly.

References

Bhattacharyya, S. (2013). Urbanization and Urban Governance in India: Challenges and Opportunities. Economic & Political Weekly, 48(45), 49-57.

Chakraborty, D., & Mukherjee, J. (2018). Urbanization, Openness, and the Real Exchange Rate. World Development, 103, 297-310.

Kundu, A., & Gupta, S. (2016). Urbanization and Urban Governance: How Cities Should Be Governed. Economic & Political Weekly, 51(50), 45-51.

Mohanty, P. K., & Mahalik, M. K. (2017). Urbanization and Economic Growth in India: Cointegration and Causality Analysis. Journal of Economic Studies, 44(1), 82-97.

Muralidharan, K., & Chaudhury, N. (2016). School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010. Education Economics, 24(5), 495-509.

Rao, M. G., & Singh, N. (2007). Economic Growth, Poverty, and Inequality in Indian States in the Pre-Reform and Reform Periods. Asian Development Review, 24(1), 137-163.

Sundararaman, T., & Vashisht, P. (2012). Public Private Partnerships in Health Care: A Case Study of India. World Health Organization (WHO) - South-East Asia Journal of Public Health, 1(2), 156-162.

Ghatak, S., & Siddiki, J. U. (2011). The Effects of Political Institutions on Poverty and Income Inequality. Public Choice, 148(3-4), 593-612.

Sen, A. (2012). Development as Freedom. Oxford University Press, India.

Bhan, G. (2009). This is no longer the City I Once Knew: Evictions, the Urban Poor, and the Right to the City in Millennial Delhi. Environment and Urbanization, 21(1), 127-142.

Aggarwal, R., & Parthasarathy, B. (2013). India's Economic Growth and Urbanization: Evidence from Census 2011. Economic & Political Weekly, 48(45), 65-72.

Dasgupta, B. (2010). Urbanization and Social Inequality in India. Social Scientist, 38(3/4), 65-82.

Majumdar, S., & Ghatak, A. (2012). Urbanization, Poverty and Development in India. International Journal of Development Research, 2(4), 413-422.

Kundu, A. (2008). Urbanization and Migration: A New Paradigm. Sage Publications, India.

Desai, S., & Vanneman, R. (2015). India Human Development Survey-II (IHDS-II). University of Maryland & National Council of Applied Economic Research (NCAER), India.

Ramachandran, R., & Swaminathan, M. (2007). Urbanization, Poverty and Development in India. Economic & Political Weekly, 42(10), 878-885.

Sridhar, K. S. (2010). Urbanization in India: Emerging Dimensions. Indian Journal of Labour Economics, 53(4), 581-595.

Roy, A. (2011). Slumdog Cities: Rethinking Subaltern Urbanism. International Journal of Urban and Regional Research, 35(2), 223-238.

Das, M. (2017). Urbanization and Economic Growth: Evidence from India. Journal of Urban Economics and Management, 5(2), 19-28.

Chatterjee, L. (2018). Urbanization and Development: Challenges and Opportunities in India. International Journal of Advanced Research, 6(5), 1254-1262.

Econometrics Research Project

Title: The impact of Microfinance credit on rice paddy productivity for small farms in major rice growing states in India.

By:

Shikhar Verma
12708

Vaibhav Seth
12752

Econometrics Project

I. Working Title

The Impact of Microfinance credit on rice paddy productivity for small farms in major rice growing states in India.

II. Research Question

What is the impact of extending microfinance credit on rice paddy productivity for small firms in major rice growing states in India?

III. Introduction/Rationale

It is vital to conduct more research on how to enhance the productivity of rice paddy production, and one promising avenue seems to be micro-credit. It is necessary to increase the productivity of rice, as rice is easily the most 'dominant' food crop of India: it covers more than 35% of the total cropped area, provides food for more than 65% of the populace and provides employment to more than half of the country's workforce.

Increasing yield rates for rice production could be beneficial in more ways than one. It could end up substantially decreasing the amount of land used for rice cultivation - if India's yield rate (currently 2.4 tonnes per hectare) was increased to Chinese levels (4.7 tonnes per hectare), we could end up halving the amount of land used for rice production, freeing it up for other purposes, while simultaneously increasing production to 205.52 million tonnes, almost a doubling from India's rice production levels at 106.19 million tonnes of rice a year. (Livemint, 2023).

Another issue that increasing rice paddy productivity would combat would be the high rates of food security and malnutrition for the steadily growing Indian populace, by providing an easily accessible source of folic acid, potassium, & magnesium. In 2022, India was ranked 68th amongst 113 countries on the Global Food Security Index, and in 2021, India was reportedly 'off-course' in meeting seven of the thirteen global neutron targets (as reported by the Global Nutrition Report), having incurred zero progress on anemia, childhood wasting and other issues (Lumoindong et al., 2021).

However, rice production has been relatively stagnant in recent years - this has been attributed to farmers using lower amounts of fertiliser per unit and lack of affordability of/access to water resources (Rahman, 2020). Extending micro-credit would help accelerate agricultural development by allowing the purchase of production-enhancing inputs such as fertilisers and herbicides and would help farmers to afford the irrigation their crops need (i.e. through tube well installation, better irrigation facilities). Amalgamating these two factors will likely improve soil moisture and fertility that will lead to a (short term) higher yield and greater water productivity (Agboklou and Özkan, 2023). Additionally, increased access to credit would lead to greater ability to purchase pesticides to combat highly prevalent pest infections such as bacterial leaf blight, blast, sheath blight, brown spot and tungro (Shivappa et al., 2022) - in fact, crops worth 5000 crore INR a year are lost to pest infections, which is attributed to low pesticide consumption in the country. (Crops Worth Rs 50,000 Crore Are Lost a Year Due to Pest, Disease: Study, 2014). These low rates are associated with the scant purchasing power of farmers, which is what micro-financing may help to alleviate. Pesticide consumption for plant protection in India has been estimated at around 600 grams per hectare, while countries such as Taiwan, China, and Japan have been noted to consume substantially more, at 17 kg, 13 kg and 12 kg respectively (*Pesticide Usage by Cotton Farmers in India*, 2018). India is the world's largest exporter of rice, and yet it ranks 10th in total pesticide consumption (*Agricultural Exports by Country 2024*, n.d.). Farmers who access credit can purchase higher level inputs such as quality seed and use more labor as well as more fertilizer, as shown in a West African study. (Ouattara et al., 2020)

Micro-finance is the panacea that must be targeted; rural farmers have recognised the vast potential of increased credit, but have been at the mercy of usurious moneylenders that charge exorbitantly high interest rates, trapping them into a vicious cycle of debt and dependency. Current government policies are (unknowingly) in support of this phenomenon, as affluent farmers receive subsidised credit while small farmers must pay off these extortionate rates to moneylenders. This is due to the government's interest intervention subsidy scheme which allocates farmers with relatively larger land parcels credit at subsidised interest rates of 7% and prompt repayers at 4%, which leads to a greater chasm between the wealthy farmers who own significant amounts of land and the local farmers with land parcels below 0.1 hectares.

Out of the 124 million small and marginal farmers in India (those who have below 2 hectares of land), only 36 million borrow from formal sources. These small and marginal farmers constitute 86.2% of the total farmers in the country, and out of the 168 billion USD in agricultural credit, over half was offered to medium and large farmers with pre-existing access to financial capital (Naik, 2020). Traditional banks are unwilling to offer credit to small and marginal farmers due to lack of access, lack of collateral offered by loan-seekers, and limited information, making them a risky bet for banks whose focus is nearly always on the bottom-line and are reluctant to incur the risks of default. (Jimi et al., 2019). Additionally, the average ticket size of the loans smaller farmers require is around 37,500 INR, which is below the loan size traditional sources of credit tend to offer (Staff & Bfsi, 2023). Traditional lending institutions have not set up branches in rural areas and net banking is not a viable panacea due to the lack of telecommunication infrastructure. In order to break this trend and offer credit to those who need it the most, we must adopt a micro-finance model to get the financing directly in the hands of the small farms who may then use it to fill their dearth of resources. (Bonso et al., 2021)

It is imperative that research on the effect of micro-credit on agricultural productivity be conducted. As India's population steadily grows, its resource base remains the same - with more mouths to feed, we must figure out how to make efficient use of what we already have, and with much of our rural economy stuck in the poverty cycle, unable to seek credit from formal institutions, our food growth is stunted. While the population growth rate remains more-or-less constant at 0.8%, agricultural growth has fallen to 3% (Dhoot, 2023). The government must take matters into its own hands and consider extending micro-credit to our farmers - this could substantially help the rural economy, by increasing access to water facilities, fertilisers, and pesticides, increasing yields while keeping costs down, which would in turn increase production and prices while leading to a reduction in the necessary land for cultivation, in turn improving the food insecurity and malnourishment situation. As shown above, it is a fully viable solution that could lead to immense benefits, and is worth looking into.

IV. Literature Review

A Tamil Nadu study (Sudha, R et. al, 2020) showed that microfinance could satisfy farmers' credit needs, which led to an increase in purchasing power and well-timed investment on agricultural investment. While this study did differentiate between tribal and non-tribal farmers, it did not create a distinction between short and longer term loans, which is important when implementing a complete policy.

Another study (Obagbemi, 2022) demonstrated that cooperative society was a veritable tool for capital creation and agricultural development, as it was the main source of credit accessed by rice farmers with 85%, and further stating that rice farmers faced high levels of pest attacks and required financial assistance to deal with it. This proves one-dimensional, but certainly a dimension of the study we are conducting as we are the primary source of credit to be farmer cooperatives.

A research study (Jimi et al., 2019) also showed that household resilience to shocks could be ameliorated via micro-credit, therefore minimising the risk-averseness of farmers and increasing their production activities, as well as letting them invest in productive technology.

Additionally, a study (Mahoukede et al., 2015) indicated the non-homogeneity of impacts of credit on rice farmers - female rice farmers gained more than male rice farmers. A challenge was also introduced - all rice farmers who managed to get credit did not use it in rice farming - this provides a new avenue for research in further studies, as we may monitor the exact further expenditures to see what engenders the most benefits (i.e, drawing the line between a government ‘hand-out’ programme versus a micro-credit programme, as we operate under the assumption that greater micro-credit leads to greater yields and incomes rather than just a transfer payment, which could be used to increase household consumption indirectly, vis-a-vis directly, so we must make sure the allocated credit is being used effectively).

An Indonesian study (Y Lumoindong et al, 2021) showed that the productivity of rice farms before obtaining credit was 3.33 ton/hectare and after credit was obtained, it increased by 0.86 ton/hectare to 4.19 ton/hectare, due to an increase in production due to greater productivity. The increase in productivity was partially due to the usage of pesticides in a timely and synchronised fashion to suppress and control pests and diseases daily, which ended up doubling the farmers’ incomes.

Another study, (Jimi et. Al, 2016), illustrated that on an average, small-scale rice farms with access to subsidized credit are approximately 12.25% more productive than farms with no credit access. Within the treatment group, the average effect of access to credit on technical efficiency of rice production is higher in female-headed households and in households with poultry and livestock activity. However, the average effect of credit access on efficiency is around 7.8% lower for the treatment households with very small farm size. Additionally, a study of small-hold farmers in Sri Lanka found, via a sample t-test, that on average, rice paddy production is greater for farmers receiving credit than farmers who did not receive credit by 798 kgs.

V. Regression Analysis a) Regression equation

Paddy yield = Amount of micro finance credit received + Cultivated area + Seed Rate + Seed Cost + Land Preparation Cost per month + Fertiliser Cost + Pesticide Cost + Labour cost per season

Residuals:

Min	1Q	Median	3Q	Max
-43872	-17010	-2228	11583	126369

Coefficients:

	Estimate	Std. Error	T-value	Pr(>|t|)
(Intercept)	4.991E+04	1.084E+04	4.606	0.000407 ***
Amount of Micro-credit Received (less than 60,000 INR, the 1st quartile)	-1.340E+04	6.573E+03	2.039	0.04420 *
Cultivated area (in acres)	-2.713E+03	1.070E+03	-2.536	0.01283 *
Seed rate (kg)	-1.216E+02	2.347E+01	-5.182	1.23e-06 ***
Seed cost (INR)	-6.502E-03	9.430E-02	-0.069	0.94518
Land Preparation Cost per month (INR)	6.701E-01	2.056E-01	3.259	0.00155 **
Fertiliser Cost (INR)	2.644E-01	9.797E-01	0.270	0.78786
Pesticide Cost (INR)	1.451E-01	2.693E-01	0.539	0.59138
Labour cost per per season (INR)	1.999E-01	2.785E-01	0.718	0.47458

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 24110 on 95 degrees of freedom

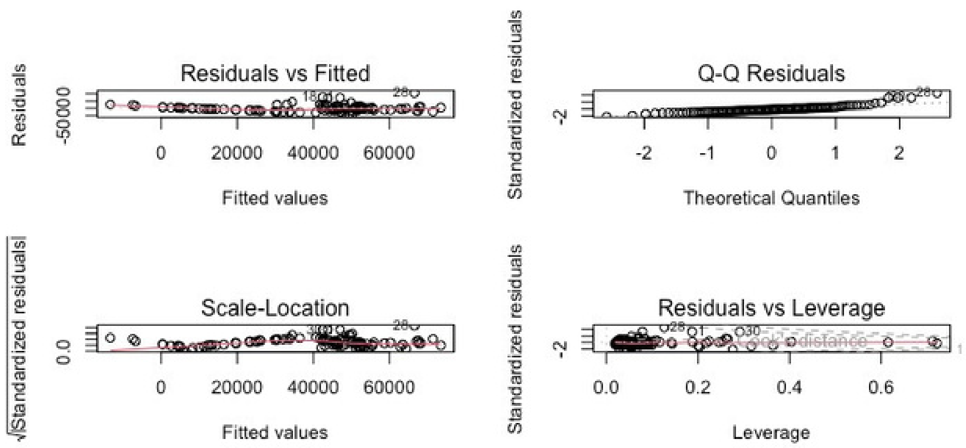
Multiple R-Squared	Adjusted R-squared	F-statistic	p-value
0.4295	0.3815	8.941 on 8 and 95 DF	4.585E-09

b) Interpretations The negative coefficient (-2.713e+03) for cultivated area suggests that for every unit increase in the area of
land cultivated for paddy farming, there is an associated decrease in the predicted yield of approximately 2,713 mounds per acre. This negative relationship may be a product of diminishing returns to scale in agriculture - as the cultivated area increases, factors like soil degradation, nutrient depletion, and pest pressure lead to reduced yields. The factor seed rate also has a negative coefficient (-1.216e+02), which suggests that for every one INR
increase in the seed rate, the predicted yield falls by approximately 121.6 mounds per acre, indicating that that higher seeding rates lead to plant overcrowding, competition for resources like sunlight, water, and nutrients, which lead to a reduction in mounds per acre for a farm.
Land preparation cost per month, however, exhibits a positive coefficient (6.701e-01), suggesting that for every one INR increase in land preparation costs, per month, the predicted yield rises approximately 0.6701 mounds per acre. This positive relationship demonstrates that greater investment in land preparation activities like leveling, plowing, and soil conditioning may cause higher productivity and hence lead to greater yields. Adequate land preparation can help to improve the structure of the soil, ensure better water infiltration, and enhance availability of nutrients, which are crucial for promoting healthy crop growth and yield maximisation.

The analysis also shows that farmers who receive less than 60,000 INR in micro-credit have faced an associated decrease in paddy yield. This could be due to greater vulnerability to external shocks, such as pest attacks or diseases, or extreme weather conditions, as well as the inability to invest in vital inputs such as quality seeds, pesticides, and fertilisers. It could also be due to lack of investment in infrastructure, such as efficient irrigation systems or storage facilities, which would increase the efficiency of the land. Insufficient funding also works to hinder farmers from investing in modern technology and practices, keeping the productivity of the land low.

Overall, our regression model model explains approximately 42.95% of the variance in Paddy yield, as indicated by the Multiple R-squared value, and the F-statistic of 8.941 with a very low p-value indicates that our model is statistically significant.

b) Test for homoscedasticity:



The residuals show no bias, therefore we can say our model fits the assumption of homoscedasticity.

c) Test for multicollinearity

Variable	VIF Value
Amount of Micro-credit Received (INR)	1.177267
Monthly Income (INR)	1.466881
Cultivated area (acres)	14.510476
Seed rate (kg)	2.597989
Seed cost (INR)	1.633295
Land Preparation Cost per month (INR)	5.038834
Fertiliser Cost (INR)	1.953756
Pesticide Cost (INR)	1.512722
Labour cost per per season (INR)	10.684904

Interpretations:

A.Amount of micro-credit received: VIF = 1.18. This is relatively low, suggesting low multicollinearity.

B.Monthly Income: VIF = 1.47. This VIF is is also relatively low, and depicts low multicollinearity.

C.Cultivated area: VIF = 14.51. This is a rather high VIF value, potentially suggesting multicollinearity between Cultivated area and other independent variables in the model.

D.Seed Rate: VIF = 2.60. This is moderately high, and suggests some degree of multicollinearity.

E.Seed Cost: VIF = 1.63. This is relatively low and suggests low multicollinearity.

F.Land Preparation Cost per month: VIF = 5.04. This high value indicates potential multicollinearity with other independent variables.

G.Fertiliser Cost: VIF = 1.95. The VIF suggests low multicollinearity.

H.Pesticide Cost: VIF = 1.51. The VIF also suggests low multicollinearity.

I.Labour cost per season: VIF = 10.68. This is a rather high VIF value, indicating some multicollinearity.

To summarise, cultivated area, land preparation cost per month, and labour cost per season have high VIF values, suggesting potential multicollinearity with other independent variables in the model.

VI. Limitations

One potential limitation could be selection bias - the farmers that already belong to the cooperative in Odessa we have taken data from already have greater access to resources, and may give a skewed picture of paddy productivity that we cannot generalise to the Indian populace. The sample group must be extended to farmers not receiving micro-credit, and farmers receiving more varied amounts of micro-credit. Additionally, the findings of the study may be specific to the selected cooperative in Odessa, and the results may not be easily generalized to other regions or countries with different agricultural paradigms and practices. However, if the breadth of the research were to expanded to more obscure regions, finding and consolidating data would prove to be much more difficult, even though it is likely that they are farmers who need the financial resources the most although lack of funding and limited time may restrict the scope and depth of the research. A more extensive study could yield more nuanced findings.

Furthermore, it is extremely hard to correct for confounding variables, which also means that causality cannot be inferred from correlation. Exogenous variables, such as market fluctuations, weather fluctuations, variable costs, and government policies can also exercise substantial impacts on paddy productivity, distorting the data and making it difficult to draw policy-applicable conclusions. Macroeconomic factors such as inflation, exchange rates, and national economic policies can influence the value of microfinance credit and its impact on farmers' financial situations.

VII. Conclusion

For practical applications, our findings can inform the design of agricultural and financial inclusion policies in India and similar agricultural economies. Policymakers can use the research to make informed decisions on resource allocation and program design. Additionally, non-governmental organizations (NGOs) and development agencies can apply the research to design and implement rural development programs that focus on providing financial services to rural communities. This could also lead to greater ‘banking of the unbanked' - including more people in the formal financial sector, as banks could utilise the data to better understand the needs and challenges of the small farmers. This would help them introduce better financial lending packages, instruments and credit assessments, breaking the financial trap and reducing reliance on usurious moneylenders.

Agricultural extension services can also use the findings to educate farmers on best practices for optimizing agricultural productivity, including the use of credit for inputs. Hnhanced rice productivity could potentially contribute to poverty alleviation among small and marginal farmers by increasing their income and improving their overall living conditions. The findings from this research can be a valuable resource for those advocating for more inclusive and supportive agricultural policies.

The regression analysis concludes by highlighting important variables affecting paddy output. The significance of investing in soil conditioning is highlighted by the positive coefficient for land preparation expenses, whereas the negative coefficients for farmed area and seed rate indicate diminishing returns to scale. Furthermore, yield is linked to microcredit amounts less than 60,000 INR, most likely as a result of infrastructural and input investment constraints. Despite accounting for 42.95% of yield variance, the model's reliability is validated by the statistically significant F-statistic. These results highlight the need for focused interventions to enhance infrastructure, financial access, and agricultural practices in order to increase yields sustainably.

VII. References

Najmuddin, O., Rasul, G., Hussain, A., Molden, D., Wahid, S., and Debnath, B. (2018b). Low water productivity for rice in Bihar, India—A critical analysis. *Water*, *10*(8), 1082. https://doi.org/10.3390/w10081082

Growth Performance and Profitability of Rice Production in India: An Assertive Analysis. (n.d.). *Munich Personal RePEc Archive*. https://mpra.ub.uni-muenchen.de/110635/1/MPRA_paper_110635.pdf Bridging the "Missing Middle" between Microfinance and Small and Medium-Sized Enterprise Finance in South

Asia. (2016). *Asian Development Bank Institute Working Paper Series*. https:// www.adb.org/sites/default/files/publication/189390/adbi-wp587.pdf

Lumoindong, Y., Bakri, R., and Syafirah, R. (2021). Analysis credit accessibility of rice farmers to the informal financial institutions in the efforts of increasing productivity and income. *IOP Conference Series*, *681*(1), 012119. https://doi.org/10.1088/1755-1315/681/1/012119

Ouattara, N., Xiong, X., Traoré, L., Turvey, C. G., Sun, R., Ali, A., and Ballo, Z. (2020). Does Credit Influence Fertilizer Intensification in Rice Farming? Empirical Evidence from Côte D'Ivoire. *Agronomy*, *10*(8), 1063. https://doi.org/10.3390/agronomy10081063

Agboklou, K. E., and Özkan, B. (2023). Analysis of rice farmers access to agricultural credit and profitability in South Togo. *Journal of Business, Economics and Finance*. https://doi.org/10.17261/ pressacademia.2023.1722

Rahman, M. S. (2020). The role of agricultural credit in rice production and food security. *ResearchGate*. https://www.researchgate.net/publication/

342343645_The_role_of_agricultural_credit_in_rice_production_and_food_security#:~:text=Ela s ticity%20of%20rice%20production%20with,rice%20production%20increased%20by%208.5%25. Naik, A. R. (2020, June 2). 50% Of Small And Marginal Farmers In India Are Financially Underserved. *Inc42 Media*. https://inc42.com/buzz/50-of-small-and-marginal-farmers-in-india-are-financially underserved/ Statesman News Service and Statesman News Service. (2017, August 1). *Moneylenders still rule India's*

rural economy. The Statesman. https://www.thestatesman.com/opinion/moneylenders-still-rule-india-s rural-economy-1501622495.html

Crops worth Rs 50,000 crore are lost a year due to pest, disease: Study. (2014, February 13). The Economic Times. https://economictimes.indiatimes.com/news/economy/agriculture/crops-worth-rs-50000- crore-are-lost-a-year-due-to-pest-disease-study/articleshow/30345409.cms?from=mdr

Bonso, A. B., Amare, A., Gurmessa, B., Bekele, Y., and Chalchisa, T. (2021). Does microcredit use helps farmers win battle against food insecurity: evidence from Jimma zone of Southwest Ethiopia. *Agriculture and Food Security*, *10*(1). https://doi.org/10.1186/s40066-021-00323-8

The Mirrority. (n.d.). *Food Security Index | India | 2013 - 2022 | Data, Charts and analysis*. https:// www.themirrority.com/data/food-security index#:~:text=India%27s%20performance%20on%20the%20Food,its%20its%20position%20in % 202021.

Aung, Swe Zin (2015). IMPACT OF MICROFINANCE ON RICE PRODUCTION IN MYANMAR. *KDI School of Public Policy and Management*.

Jimi, N. A., Nikolov, P. V., Malek, M. A., and Kumbhakar, S. (2019). The effects of access to credit on productivity: separating technological changes from changes in technical ef ficiency. *Journal of Productivity Analysis*, *52*(1/3), 37–55. https://www.jstor.org/stable/48701569

Chandio, A. A., Jiang, Y., Wei, F., and Guangshun, X. (2018). Effects of agricultural credit on wheat productivity of small farms in Sindh, Pakistan. *Agricultural Finance Review*, *78*(5), 592–610. https://doi.org/10.1108/afr-02-2017-0010

Jimi, N. a. and. K. S. and. N. P. V. and. M. M. A. (n.d.). Does Credit Access Improve Firm Output? Evidence from a Field Experiment in Bangladesh. *ideas.repec.org*. https://ideas.repec.org/p/ags/ aaea16/236230.html

Obagbemi, D.S. and Bamidele J. (2022). View of Effects of Micro-Credit Scheme on Rice Production among Smallholder Farmers in Kwali Area Council, Abuja. https://www.ejbmr.org/index.php/ ejbmr/ article/view/1666/949

Sudha, R and Ashok K.R. (2020). *Impact of Microfinance on Farm Income of Small and Marginal Farmers in Western Tamil Nadu*. (2020). Publication | FinDev Gateway. https://www.findevgateway.org/ paper/2020/12/impact-microfinance-farm-income-small-and-marginal-farmers-western-tamil nadu

Mahoukede, K. F., and Aliou, D. (2015). Impact of Use of Credit in rice farming on rice productivity and income in Benin. *International Conference of Agricultural Economists*.

Staff, E., & Bfsi, E. (2023, February 16). Rural credit markets, loan disbursements grow 64%: Report. *ETBFSI.com*. https://bfsi.economictimes.indiatimes.com/news/financial-services/rural-credit-markets-loan-disbursements-grow-64-report/97981857

Diseases - IRRI Rice Knowledge Bank. (n.d.). http://www.knowledgebank.irri.org/step-by-step-production/growth/pests-and-diseases/diseases

Pesticide usage by cotton farmers in India. (2018, May 17). Economic and Political Weekly. https:// www.epw.in/journal/2018/19/special-articles/pesticide-usage-cotton-farmers-india.html

Shivappa, R., Navadagi, D. B., Baite, M. S., Yadav, M. K., Rathinam, P. S., Umapathy, K., Pati, P., & Rath, P. C. (2022). Emerging minor diseases of rice in India: Losses and management strategies. In

IntechOpen eBooks. https://doi.org/10.5772/intechopen.99898

Questions asked:

Have you received any agricultural credits? If yes, what is the amount (INR), and what is the duration of your loan?

What is your monthly income (INR)?

What is your cultivated area (acre)?

What is your paddy yield (maunds)?

What is your seed rate (kg)?

What is your seed cost (INR) ?

What is your land preparation cost (INR)?

What is your fertilizers cost (INR) ?

What is your pesticides cost (INR) ?

What is your irrigation cost (INR)?

What is your labor cost (INR)?

Raw data:

Age	Amount	Monthly Income (INR)	Cultivated area (in acres)	Paddy yield (in maunds per acre) - Rabi	Paddy yield (in maunds per acre) - Kharif	Seed Rate (kg)	Seed Cost (INR)	Land Preparation Cost (INR) per month	Fertiliser Cost (INR)	Pesticide Cost (INR)	Irrigation cost (INR per acre) - Kharif	Irrigation cost (INR per acre) - Rabi	Labour cost per season (INR)
38	100000	10000	8	120000	120000	240	9500	48000	1600	11000	101.2	182.2	50000
62	480000	5000	4	60000	60000	30	7000	24000	7000	5000	101.2	182.2	28000
48	80000	10000	4	48000	48000	25	9000	26000	12000	8000	101.2	182.2	35000
71	480000	10000	6	78000	78000	25	11000	38000	10000	8000	101.2	182.2	41000
74	60000	800	3	32000	32000	25	4800	18000	8000	6000	101.2	182.2	26000
49	480000	10000	6	78000	78000	300	7500	88000	10000	8000	101.2	182.2	10000

55	100000	10000	5	60000	60000	25	9000	30000	9000	7000	101.2	182.2	27000
57	70000	8000	5	56000	56000	250	6250	30000	10000	8000	101.2	182.2	30000
46	100000	9000	4	50000	50000	25	6000	40000	7000	6000	101.2	182.2	22000
52	80000	10000	6	56000	56000	25	9000	45000	8000	6000	101.2	182.2	27000
44	50000	9000	6	60000	60000	30	11000	21000	8000	7000	101.2	182.2	39000
48	60000	7000	2	20000	20000	30	3600	10000	10000	8000	101.2	182.2	12000
54	80000	8000	8	60000	60000	25	12000	25000	7000	5000	101.2	182.2	60000
42	80000	7000	5	56000	56000	25	7500	17000	6000	6000	101.2	182.2	45000
42	10000	8000	4	49000	49000	25	6000	42000	6000	5000	101.2	182.2	24000
58	80000	7000	3	38000	38000	30	5400	15000	8000	6000	101.2	182.2	20000
57	50000	9000	6	60000	60000	30	4000	21000	8000	7000	101.2	182.2	39000
59	100000	10000	8	120000	120000	240	11500	48000	16000	11000	101.2	182.2	50000
41	100000	8000	5	48000	48000	30	9000	18000	8000	6000	101.2	182.2	55000
48	40000	6000	3	29000	29000	30	5400	12000	9000	7000	101.2	182.2	32000
60	60000	8000	3	3200	3200	25	4000	18000	8000	6000	101.2	182.2	26000
47	80000	1000	6	78000	78000	300	7500	38000	10000	8100	101.2	182.2	38000
69	80000	10000	4	48000	48000	25	9000	26000	12000	8000	101.2	182.2	35000
38	80000	10000	6	78000	78000	25	11000	38000	10000	8000	101.2	182.2	41000
50	80000	5000	4	60000	60000	240	58000	24000	8000	5000	101.2	182.2	30000
43	68000	3000	3	45000	45000	25	4500	18000	5000	2000	101.2	182.2	22000
39	72000	7000	6	90000	90000	250	7500	48000	9000	7000	101.2	182.2	42000
51	85000	6000	9	172000	172000	25	14000	54000	16000	12000	101.2	182.2	36000
56	468000	10000	4	48000	48000	25	6000	32000	7000	5000	101.2	182.2	32000
44	80000	8000	8	112000	112000	300	144000	48000	12000	9000	101.2	182.2	60000

56	70000	8000	8	55000	55000	30	144000	24000	10000	8000	101.2	182.2	50000
62	90000	8000	7	4900	49000	35	147000	29000	10000	8000	101.2	182.2	40000
72	470000	8000	8	55000	55000	30	14400	24000	8000	10000	101.2	182.2	50000
56	90000	8000	7	49000	49000	35	14700	29000	10000	8000	101.2	182.2	40000
51	45000	8000	3	39000	39000	40	7200	11000	10000	9000	101.2	182.2	25000
41	100000	10000	8	65000	65000	25	12000	30000	9000	70000	101.2	182.2	50000
57	100000	10000	8	65000	65000	25	12000	30000	9000	7000	101.2	182.2	50000
44	60000	80000	6	52000	52000	40	144000	22000	9000	80000	101.2	182.2	50000
37	45000	6000	4	40000	40000	30	7200	12000	12000	10000	101.2	182.2	30000
53	45000	8000	3	39000	39000	40	7200	11000	10000	9000	101.2	182.2	25000
57	100000	10000	8	65000	65000	25	12000	30000	9000	7000	101.2	182.2	50000
44	60000	80000	6	52000	52000	40	14400	22000	9000	8000	101.2	182.2	50000
37	45000	6000	4	40000	40000	30	7200	12000	12000	10000	101.2	182.2	30000
63	60000	8000	6	52000	52000	40	14400	22000	9000	8000	101.2	182.2	50000
65	45000	6000	4	40000	40000	30	7200	12000	12000	12000	101.2	182.2	30000
51	30000	6000	2	27000	27000	50	6000	9000	12000	9000	101.2	182.2	12000
40	72000	8000	6	42000	42000	25	9000	24000	9000	6000	101.2	182.2	39000
43	30000	6000	2	27500	27500	50	6000	9000	12000	9000	101.2	182.2	12000
52	72000	8000	6	42000	42000	25	9000	24000	9000	6000	101.2	182.2	39000
44	80000	9000	6	48000	48000	30	10800	21000	9000	8000	101.2	182.2	39000
60	60000	7000	8	65000	65000	25	12000	25000	10000	8000	101.2	182.2	39000
70	80800	9000	6	48000	48000	30	10800	21000	9000	8000	101.2	182.2	39000
64	60000	7000	8	65000	65000	25	12000	25000	10000	8000	101.2	182.2	39000

72	62000	8000	8	68000	68000	25	12000	25000	12000	10000	101.2	182.2	50000
66	40000	6000	2	25000	25000	40	48000	9000	12000	8000	101.2	182.2	30000
76	62000	8000	8	68000	68000	25	12000	25000	12000	10000	101.2	182.2	50000
57	40000	6000	2	25000	25000	40	4800	9000	12000	8000	101.2	182.2	30000
40	78000	9000	7	60000	60000	30	12600	24000	10000	8000	101.2	182.2	39000
56	59000	8000	6	51000	51000	32	11520	22000	9000	7000	101.2	182.2	34000
60	78000	9000	7	60000	60000	30	12600	24000	10000	8000	101.2	182.2	39000
40	59000	8000	6	51000	51000	32	11520	22000	9000	7000	101.2	182.2	34000
44	50000	10000	6	52000	52000	30	10800	23000	9000	7000	101.2	182.2	39000
52	48000	7000	7	65000	65000	25	10500	29000	9000	7000	101.2	182.2	42000
63	29000	7000	3	30000	30000	30	54000	15000	10000	8000	101.2	182.2	24000
56	40000	6000	4	32000	32000	25	60000	16000	10000	8000	101.2	182.2	30000
58	55000	8000	8	75000	75000	30	14500	32000	9000	7000	101.2	182.2	38000

54	65000	7000	5	35000	35000	40	12000	22000	10000	8000	101.2	182.2	34000
51	100000	10,000	15	8000	8000	450	21,000	75,000	15,000	12,000	101.2	182.2	60,000
52	100000	8,000	14	6000	6000	450	39,000	52,000	14,000	12,000	101.2	182.2	70,000
50	100000	10,000	13	7000	5000	300	24000	50,000	12,000	8,000	101.2	182.2	70,000
80	100000	9000	12	8000	5500	300	30,000	55,000	11,000	9,000	101.2	182.2	60,000
59	100000	8000	10	7000	4500	300	18,000	24,000	8000	6000	101.2	182.2	60,000
45	100000	8000	7	7000	4000	300	13,000	20,000	10,000	8,000	101.2	182.2	40,000
50	100000	10,000	22	10000	7000	400	49,000	45,000	10,000	7,000	101.2	182.2	90,000
38	100000	9,000	20	9000	5000	350	36,500	40,000	12,000	8,000	101.2	182.2	85,000
55	100000	10,000	30	10,000	8000	300	54,000	1,00,000	15,000	11,000	101.2	182.2	90,000

30	100000	10,000	37	15,000	8,000	400	90,000	70,000	20000	14000	101.2	182.2	1,50,000
32	100000	8000	37	13000	7000	400	90,000	70,000	20,000	14,000	101.2	182.2	1,30,000
51	100000	9000	5	8000	5000	500	15,000	40,000	21,000	18,000	101.2	182.2	65,000
72	100000	8000	11	6000	5000	450	30,000	48,000	17,000	15,000	101.2	182.2	46,000
60	100000	9000	29	9000	8000	400	70,000	85,000	14,000	12,000	101.2	182.2	1,10,000
56	100000	10,000	17	8000	6000	400	41,000	62,000	10,000	9,000	101.2	182.2	63,000
50	100000	9000	27	8000	6000	400	65,000	80,000	13,000	10,000	101.2	182.2	1,20,000
50	100000	8000	18	8000	7000	300	33,000	68,000	14,000	12,000	101.2	182.2	75000
46	100000	9000	13	7000	4900	350	28,000	65,000	12,000	10,000	101.2	182.2	68,000
60	100000	8000	26	8000	6000	300	47,000	70,000	12,000	11,000	101.2	182.2	80,000
55	100000	9,000	40	10000	8000	400	96,000		1,50,000 21,000	21,000	101.2	182.2	1,50,000
29	100000	10,000	19	10000	7000	400	46,000	68,000	11,000	9,000	101.2	182.2	85,000
50	100000	8000	16	8000	5000	350	34,000	65,000	12,000	9,000	101.2	182.2	63,000
57	100000	10,000	25	10000	7000	400	60,000	80,000	14,000	10,000	101.2	182.2	90,000
54	100000	8000	10	7000	5000	300	18,000	42,000	10,000	12,000	101.2	182.2	65,000
56	100000	10,000	23	8000	7000	400	55,200	78,000	13,000	11,000	101.2	182.2	71,000
90	100000	8000	14	7000	5000	450	38,000	60,000	10,000	9,000	101.2	182.2	66,000
30	100000	10,000	21	8000	5000	350	44,500	78,000	12,000	11,000	101.2	182.2	85,000
52	100000	9000	27	9000	7000	300	49,000	80,000	18,000	16,000	101.2	182.2	90,000
50	100000	8000	10	8000	5000	450	28,000	40,000	9,000	8500	101.2	182.2	52,000
52	100000	9000	19	9000	7500	350	40,000	60,000	10,000	9000	101.2	182.2	76,000
59	100000	9000	18	8000	6000	450	50,000	50,000	18,000	18,000	101.2	182.2	80,000
60	100000	10,000	17	9000	7000	400	41,000	48,000	10,000	9,000	101.2	182.2	65,000
63	100000	9000	12	7000	4000	300	22,000	51,000	11,000	10,000	101.2	182.2	59,000

50	100000	10,000	16	8000	5000	300	29,000	53,000	10,000	8,000	101.2	182.2	72,000
79	100000	9000	22	9000	6000	400	52,800	76,000	16,000	12,000	101.2	182.2	79,000

52	100000	8000	14	8000	6000	300	25,500	56,000	13,000	10,000	101.2	182.2	80,000
40	100000	10,000	30	10000	8000	400	75,000	1,20,000	18,000	16,000	101.2	182.2	1,30,000

ECONOMIC GROWTH IN INDIA, IS IT A JOBLESS GROWTH?

AN EMPIRICAL EXAMINATION USING OKUN'S LAW

1.RAJU KUMAR (12747)

2. MANISH KUMAR (12726)

3. YOGESH KUMAR (12749)

Introduction

One of the cardinal objectives of macroeconomic policy is to achieve economic growth. Whyt is, why are all government policies designed in such a way that they will be consistent with the growth objective of the government? But the question is why is the government so concerned about economic growth? The answer to this question is multi-dimensional, but the most important one is to get political popularity. This is how the trend is: when there is an output expansion, the level of employment will rise, and the reduction in unemployment makes the government popular and this leads to winning elections. Therefore, the government will always seek output expansion even at the cost of inflation to get re-election by the public. The leading question to this episode is that does output expansion always lead to increased employment? There are three answers to the above question: when output expands rapidly, unemployment will reduce. Secondly, when output grows slowly, unemployment will rise, and finally, when output growth is equal to the potential output, unemployment will be constant.

At a theoretical level, Okun's law was the first effort made to describe the relationship that exists between output and unemployment. The whole idea started with the Keynes notion that at any time the economy is operating below full employment equilibrium, so that whenever there is an increase in aggregate demand above the present level of output, firms will hire more workers to increase the supply of output that will match the existing demand. In the process of doing so, aggregate employment will increase. This will continue until output reaches its optimal level. Therefore, to Keynes, the government can stimulate the economy through either an increase in spending or reduction in tax to boost aggregate demand which will imply output expansion and employment. At the empirical level, a lot of efforts exist trying to verify empirically the existence of Okun's law to different economies. See, for example, the works of Lancaster and Tulip (2015), Haggis (2011), Wen and Chen (2012), and Al-delaine (2016), Blinov (2014), Ball et al. (2012), and Lal (2010) among others. These researchers were conducted with the aim of determining whether the deviation of output from its potential level leads to unemployment. In other words, does an increase in output always lead to economic growth? The findings of these researchers vary across countries as some studies provide evidence that an increase in economic growth is followed by an increase in employment, whereas to others, an increase in output must be reasonable enough before it can lead to an increase in employment. The general conclusion in the literature is that economic growth does not in all the time imply a reduction in unemployment.

The main objective of this work is to analyze the relationship between economic growth and unemployment in India using Okun's law. Although the law has existed for a long period of time, the literature on its application to examine the relationship between employment and economic growth is very scanty; to the limited search of the literature, we only found the work of An et al. (2016) which fits the Okun's law for the low- and lower-middle-income countries like India inclusive. Therefore, the study intends to investigate whether economic growth in India always leads to a reduction in the level of unemployment' this will include the use of both linear and nonlinear econometric models so that the robustness of the estimates with several these approaches can be ascertained.

KEYWORDS: *unemployment, Okun's law, output*[i]

---

i rajukumareco22@pgdav.du.ac.in
Rajukumar1, Manishkumar2, Yogeshkumar3
PGDAV(M) college, Ring Road

Abstract

This study empirically examines the relationship between unemployment and output in India by fitting the Okun's law. To achieve this objective, an annual time series data for unemployment and output were collected from the World Bank and Reserve Bank of India. The empirical estimation starts with testing the unit root evidence using Perron (1997) and Elliot et al. (Econometric 64(4):813–836, 1996) DF-GLS, and the impact models were estimated using linear and nonlinear

econometric models. The estimates show that quantitatively, there is no significant difference between the two modelling approaches. The evidence indicates that the relationship between unemployment and output for the Indian economy is consistent with Okun's law. We find most of the coefficients to be negative, less than unity and statistically significant and this conforms to the theoretical expectation. The study also finds that based on the estimated evidence the 11.75% nominal GDP growth rate as targeted by the Indian government will only result in 0.52% decrease in unemployment which is insignificant for the Indian population. The study found that, to get a 1% decline in unemployment, 25 % nominal GDP growth rate is required which is twice the targeted value. Therefore, the study concludes that although Indian output is growing, the growth is jobless because it is not up to the threshold level that ensures a decline in unemployment.

Keywords: Unemployment, Okun's law, Output

Theoretical model

The existence of a meaningful relevance between economic growth rate of GDP and unemployment rate has been extensively studied in the economic literature (Misbah Akram, 2014). In 1962, Okun, in a paper, described two simple empirical relationships between unemployment and real production. In the first model, the seasonal changes in the unemployment rate Y (expressed as a percentage) were related to the seasonal percentage of changes in GDP X. This relationship was estimated by Okun using 55 observations, from the second season of 1947 to the fourth season of 1960 for America, as below:

$$Y=0.3-0.3X \qquad R^2=0.79 \qquad (1)$$

According to Okun's estimate, if the real GNP (X) doesn't change, the rate of unemployment (Y) will rise 0.3 cent (3%) from season to another season. Also, the unemployment rate will fall 0.3 percent for a one percent increment in the GNP. In alternative words, one percent (1%) higher unemployment rate means 3.3 percent lower GNP. The second Okun model, called the Gap Model, links the unemployment rate to the gap between actual and potential GDP and is expressed as follows:

$$U=a + b(gap) \qquad (2)$$

Where in this relation potential GDP "The amount of GDP at 4% unemployment rate (U = 4)" is considered. The b parameter for this equation is estimated to range between 0.28 and about 0.38 (Okun, 1962).
This equation was estimated in 1961 by the Board of Economic Advisers of the US Joint Economic Committee using the seasonal data for the period 1953-1960 as below:

U=3.72 +0.36 gap R^2=0.93 (3)

According to this equation, 1 percent increase in the unemployment would reduce the gap between real and potential GDP by 0.28 percent.Also, the estimated unemployment rate for a zero gap is 3.72 per cent, which is not far from the ideal 4 per cent level for full employment (Okun, 1962). According to Okun's interpretation, the parameter (a) in the above equation can be expound as the unemployment rate correlated with full employment. Also, since the actual rate of production is expected to be lower than its potential value, the sign of the coefficient (b)is positive.

The problem with this model is the use of full employment conditions and the amount of potential output that neither of these two data sets are directly visible and available in the macro economy. However, Okun was able to generate a series of data on potential outputs, assuming a four percent unemployment rate for full employment. But by changing the assumption of unemployment rates of full employment, different amounts of potential output are obtained. Okun also argues that simplifying these equations can be potentially problematic (Edward S. Knotek, 2007). The existence of these problems forced economists to make some changes to Okun's core relationships. Of course, these relationships and models are still called Okun's law, despite the considerable differences in their initial form of equations (Roma & Valde, 2017). One of these models is the "dynamic model", built on the assumption that in Okun's law some variables eliminated from the right side of the model; And more recently, many economists have used it in their studies (Knote, II.E.S., 2007) (Boulton, T., 2010) and Mossa, I. (2008). In the current dynamic model form, current and past years real production growth and past years unemployment rates as the right variables of the equality and the current changes in the unemployment rate are on the left equality, (Knut, 2007) Moses (2008) in his study has estimated this relationship.

$$U_t^c = \alpha + \sum_{i=1}^{m} \beta_i U_{t-i}^c + \sum_{i=0}^{n} \gamma_i Y_{t-i}^c + \upsilon_t$$

Although this model has some similarities to the differential model, the problem is that it is not as easily interpreted as the original differential model of Okun law. Another problem that is commonplace in the law of Okun (in all cases) is that labor unemployment is considered a representative of all idle resources. However, unemployment is only one of the determinants of the amount of labor, and other factors such as Population, the fraction of the population that is part of the labor force, and the number of hours worked by workers not included. These problems and deficiencies in Okun's law led to the creation of a "Production-Function Version". This model is derived by combining a theoretical production function with Okun's law-based slit model. In the theoretical production function used, products are derived from a combination of labor, capital and technology (Knote, II.E.S., 2007). According to (Freeman, 2001) study of total production (written in the form of a natural logarithm) it is stated as follows:

$$Y_t = \tau_t + \alpha K_t + \beta N_t$$

$Y_t = {}_t + \alpha K_t + B\ N_t$ These include Yt: product, Kt: volume of capital used, Net: employment, and t: technical progress. In this respect, according to (Paldam, M., 1987) and (Parktown, M. F. G., 1993) unemployment is defined as U=L-N, where L represents the total labor force. The " Okun's law gap model" is thus defined (* sign indicates the equilibrium values of the variables).

$$Y_t - Y_t^* = \alpha(K_t - K_t^*) + \beta(L_t - L_t^*) - \beta(U_t - U_t^*)$$

It is necessary to explain that the technical progress at all-time points is assumed to be equilibrium (*). The Gap Model 6, by adding a random (*t*) component to account for productive shocks, can be specified and estimated for country (i) during the (t) period (Freeman, D. G. 2001) as below:

$$Y_{it}^{'} = \beta_{ui} U_{it}^{'} + \beta_{li} L_{it}^{'} + \beta_{ki} K_{it}^{'} + \varepsilon_{it}$$

Where the " ' " signifies the gaps of the variables or their deviation from the equilibrium values. This model, unlike earlier models, was more empirical; it had a theoretical structure that was considered an advantage, but the criticism of this model was difficult to measure in institutions such as capital reserves (Knote, II.E.S., 2007).

Although the existence of a stable empirical relationship such as the Okun relationship may be important for political modeling, few macroeconomic theories have modeled the association of GDP and unemployment. According to the conservative Keynesian view, it is very easy to interpret and explain Okun's law. Due to changes in aggregate demand, companies are changing their production plans, which results in changes in labor demand and thus affecting the unemployment rate. But this kind of conception and logic only becomes problematic when prices and wages are implicitly assumed to be fixed. Neo-Keynesian economics attempts to overcome these problems by assuming the nominal and true inflexibility of these variables, for example if we consider, as in (Blanchard, O. and Kiyota, N.1987), a monopolistic competition model by introducing a cost list (or the same). Nominal Flexibility in the Commodity Market and Flexibility in the Labor Market It can easily be shown that changes in aggregate demand will affect production and employment and thus unemployment (SOGNER , L. and Stiassny, A., 2000).

In 1994, (Aghion, P. and Howitt, P.) and Hewitt examined the long-term effects of growth on unemployment. Their analysis revealed two rivalry effects of GDP growth on unemployment: The first effect was called investment, whereby increase in growth, through job creation, increases investment return and thus reduces the equilibrium unemployment rate. The second is called Creative Destruction, whereby increases in growth, due to increased demand for labor, increase the duration of work, thereby increasing the equilibrium level of unemployment. Lee, J. (2000) estimates that the eccentricity of Japan is 6.12 percent, indicating that this anomaly represents a significant institutional inflexibility in the Japanese labor market, especially given the level of job security available. Stephen, J. N. (1997) attributes the high unemployment in many European countries to the high rigidity in the labor market in these countries due to the high penetration power of unions, poor educational standards and the benefits of long-term unemployment. On the other hand, there has been an asymmetric relationship of unemployment rate and economic growth, according to studies conducted in recent years by some researchers, such as Cuaresma, J.C. (2003), Chuan Huang, H. and Chang, Y. K. (2005), Jardin, M. and Stephan, G. (2011). It shows the asymmetry of Okun's law. Courtney, H.G. (1991) is the first to propose that the coefficient of Okun may be different during times of boom and stagnation. Taking the "aggregate production function" approach, he has attributed the asymmetry of Okun's law to the succession of factors over periods, Multi-Factor Productivity Fluctuations and Changes of the Distribution in Sectoral Growth Rates (Harris, R. & Silverstone, 2001). Also, the results of some studies such as Fouquet, J. (2005) and Isthmian, M. (2010) indicate the instability of Okun's law over time.

Blanchard. (1999) has assumed that the stability of the Okun coefficient decreases over the time. change in GDP, we will always see a stronger impact on unemployment. He described the reasons for this as strong international competitiveness, less legal protection for workers, and a general shift in firm spending toward a reduction in the labor force (Sanger antialien, 2000). Isthmian, M. (2010) considers the Okun's coefficient as a reduced form of several structural parameters. According to his analysis, the Okun's coefficient tends to change over time. He argues that this coefficient of instability is primarily due to temporary changes in the parameters of the structural relationships between supply-side and labor demand variables, which are also likely to be due to changes in institutional legal characteristics and other related features. The labor market is a commodity. In fact, Isthmian, M. (2010) shows that Okun's law is inherently inclined to change over time, especially in response to structural changes in the legal-institutional characteristics and other characteristics of the labor and commodity markets. However, despite all the above considerations, Mossa, I. (2008) concluded that unemployment and production were irrelevant in four Arab countries. The reasons for this lack of communication include:

1. Unemployment is not cyclical. In some countries unemployment is structural or frictional, not periodic. The results in Structural unemployment from the mismatch of changes in the economy with changes of education. This means that unemployment is not because the economy is in recession, but because there is no need to learn the skills necessary to do the job. Frictional unemployment also stems from the mismatch of job vacancies with unemployed people. That is, people may have the skills to do specific tasks, but their lack of awareness of the positions that correspond to their skills has made them unemployed. According to some definitions of these two types of unemployment, increased production in such countries cannot reduce such unemployment.

2. Inflexible in the labor market. The flexibility of the labor market (for example, employer's freedom to hire or fire employee and the absence in labor laws like minimum wage) makes unemployment more affected by production. Therefore, in countries where the government plays a dominant role in the labor markets, the labor markets will be inflexible, which will lead to a lack of relation between production and unemployment, and

3. In economies dominated by the government or only one sector (such as, the oil sector), growth in this sector (which generates growth in this sector (which generates growth in this sector (which generates growth in this sector (which generates growth of economic) cannot decrease unemployment if most of the workforce is not concentrated.

Data and definition of variables

To estimate the relationship between unemployment and output for the Indian economy, annual time series data from 1991 to 2022 were collected from World Bank and Reserve bank of India . Two measures of output and unemployment were used, that is, unemployment (*U*) as a percentage of the total labor force which was sourced from the World Bank and youth unemployment which was sourced from the
Reserve Bank of India. The two measures of output are Gross national income.
(GNI) and GDP growth rate for output. The study chooses the sample size due to the unavailability of data for a long period. The data of unemployment that exist for Indian economy in World Bank and Reserve Bank of India.

Table: 1

Results of Stationarity analysis (Augmented Dickey Fuller Unit Root Tests) of GDP Growth rate

H0 Hypothesis: GDP growth has a unit root (Non-Stationary of variable)

		t-Statistic	Prob.*
ADF test statistic		-5.302579	0.0002
Test critical values:	1% level	-3.699871	
	5% level	-2.976263	
	10% level	-2.627420	

Source: Authors Calculations

According to Table 1, the (ADF) test results show that all variables are Stationary at 99%, 90%, and 95% confidence level. It means that our H_0 hypothesis (GDP growth has a unit root) is rejected, and our alternative hypothesis is confirmed (When prop value (0.0002) is less than 0.05 (in case 95% of confidence level) the Null Hypothesis is rejected and vice versa.), meaning that the growth rate of GDP is a Stationary variable.

Table 2- Hypothesis analysis table

confidence levels	critical values	Hypotheses
99% of confidence level	5.302579>3.699871	❖ H0 (Null) Hypothesis is accepted ❖ H1 (alternative) hypothesis is rejected
95% of confidence level	5.302579>3.699871	❖ H0 Hypothesis is accepted ❖ H1 hypothesis is rejected
90% level of confidence	5.302579>3.699871	❖ Null Hypothesis is accepted ❖ alternative hypothesis is rejected

Source: Authors Calculations

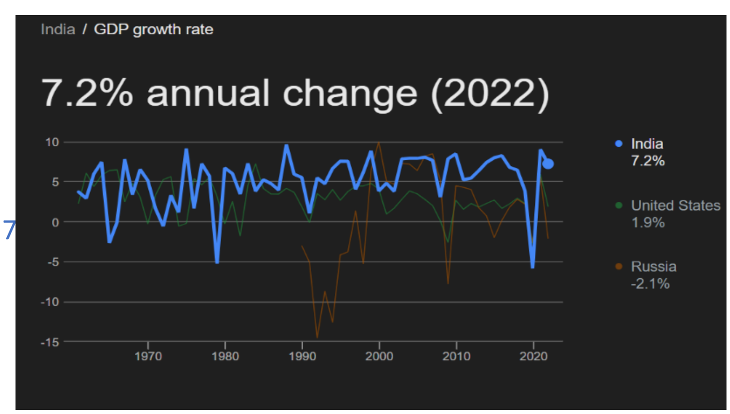


Table: 3
Results of Stationarity analysis (ADF Unit Root Tests) of unemployment rate

H0: Unemployment has a unit root (Non-Stationary of variable)

		t-Statistic	Prob.*
Augmented DF test statistic		-2.956217	0.0526
Test critical values:	1% level	-3.711457	
	5% level	-2.981038	
	10% level	-2.629906	

Source:Researchers Findings

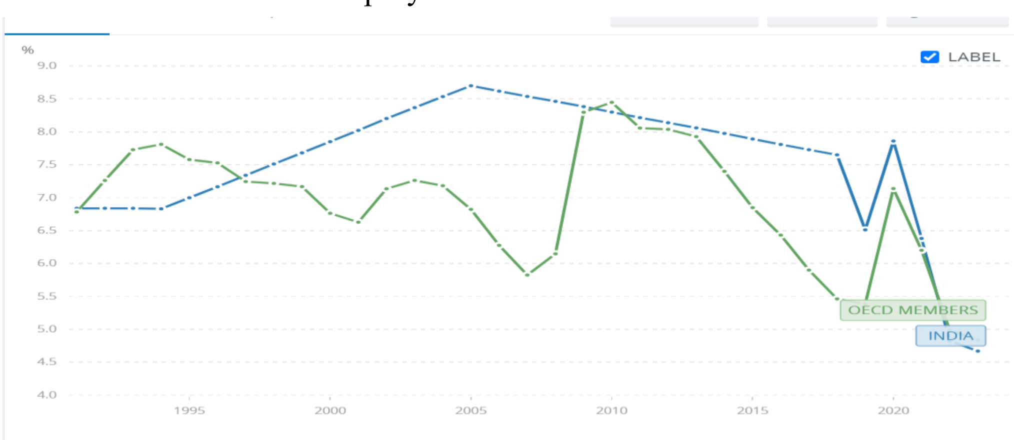
Table 4- Hypothesis analysis table

Confidence levels	Critical values	Hypotheses
99% of confidence level	2.956217<3.711457	❖ H0 Hypothesis is accepted ❖ H1 hypothesis is rejected
95% of confidence level	2.956217<2.981038	❖ Null Hypothesis is accepted ❖ alternative hypothesis is rejected
90% of confidence level	2.956217>2.629906	❖ H0 (Null Hypothesis) is rejected ❖ H1 (alternative hypothesis) is accepted

Source: Authors Calculations

According to Table 1 and 2, the Augmented DF test results show the non-Stationary of all variables at 99% confidence level and 90% confidence level but at 90% confidence level all variables are Stationary. And on the other hand, we chose our confidence level to analyze the relationships between variables at 90 %, so we aren't faced with the non-Stationary of variables. It means that our H0 (GDP has a unit root) at 90% confidence level is rejected, and our alternative hypothesis is confirmed, meaning that unemployment is a Stationary variable at 90% confidence level.

unemployment



The graph-1 relates to GDP growth over the years 1991-2018. As we can see from this graph, it is clear that the GDP growth variable is stable, and our data are distributed around the mean. Which can be deduced to be Stationarity of this variable. The graph-2 shows unemployment during the years 1991-2018. The distribution of the data and as well as the unemployment graph is also clear that unemployment data is stationary too. Which can be deduced as being this Stationarity.

Analysis of the OLS model results:

After performing the Augmented Dickey-Fuller Unit Root Tests we came to the conclusion that at 90% confidence level our variables are Stationary and we do not experience Unit Root problems and now in this section, using the EViews program and the Ordinary least-squares method, we obtained the regression equation, the results of which are shown in table 5 below.

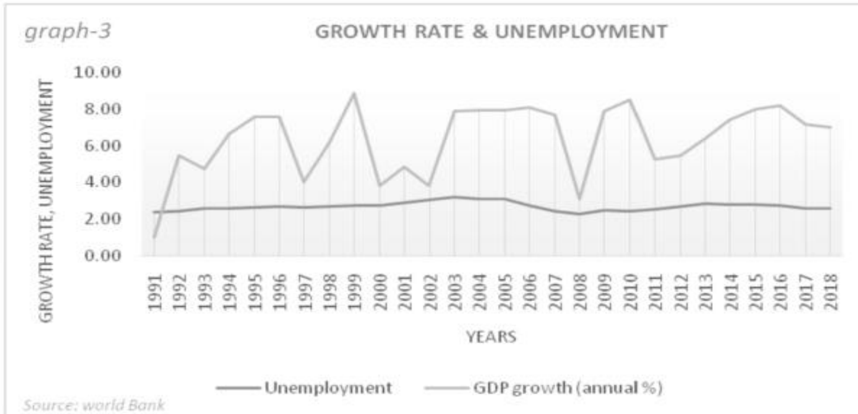
***Table 5- Results of* the OLS model**

Dependent Variable: Unemployment
Method: Least Squares
Sample: 1991-2018
Included observations: 28
Null Hypothesis: There is no significant relationship between variables

Variable	Coefficient	Error	t-Statistic	Prob.
C	2.455890	0.148719	16.51359	0.0000
GDPgrowth	0.035030	0.022386	1.564771	0.1297
R^2	0.086068	Mean dependent var		2.678929
Adjusted R^2	0.050917	S.D. dependent var		0.230480

Source: Authors Calculations

According to table 5: The results of the OLS (ordinary least square method) model estimation show that the relationship between GDP growth and unemployment is positive (Okun's law violation). In other words, with the increase in GDP growth rate, unemployment will go up. Conversely, unemployment decreases as GDP growth declines (Okun's law violation). On the other, since (Prob. = 0.1297) is greater than (0.05) H_0 hypothesis is accepted and H_1 is rejected. That is, there is no significant relationship of economic growth and unemployment.
The regression equation of GDP and unemployment is (***Unemployment = 2.40.035*GDPGROWTH***). This indicates a direct relationship between GDP and unemployment. The GDP growth coefficient is 0.035, which indicates that when changes one percent GDP growth it causes a 0.035 percent change in unemployment, in other words whenever the 100 percent change in GDP growth causes a 0.035 percent change in unemployment in a positive direction. The R-squared is 0.0861, which represents the total impact of the GDP growth on unemployment. It means 8.61 Changes in unemployment are due to GDP growth. The intercept is 2.46 which represents our unemployment without GDP growth which means if we exclude GDP growth, unemployment will be 6.14.



Reasons for rejecting Okun's law:

Violation of Okun's law in the Indian Economy In addition to the three that Mossa, I. (2008) has mentioned, there are several other reasons to be mentioned below:

a) Labor force transfer among sectors: The growth of the service and industry sectors in India has led to the shift of the labor force from agriculture to services and industry. That is, the labor force that was formerly in the agricultural sector as disguised unemployment and not productive, now with Transition to services and industry have contributed to economic growth.

b) Increasing labor productivity by increasing education and expertise has increased the share of each unit of labor in national production.

c) Increasing the labor force in India is another reason that the Okun's law is being violated. That is, every year a large volume of new labor force is joining the body of the Indian economy, which is a major challenge for the economy and the economy must create job for them. That is, no matter how high India's economic growth is, it is still affected by the growth of the labor force and cannot change unemployment.

d) There is another possibility that the current rate of unemployment in India is the natural Unemployment rate or lower than the natural unemployment rate. In this case too, one can understand the violation of Okun's law.

e) The inaccuracy of data on unemployment rate and as well as economic growth is another reason for not accurately calculating and still violating Okun's law. (It's a presumption.)

The above reasons may be part of the cause of the violating of the Okun's law. While there are many other reasons that need to be researched and each of the reasons mentioned above requires a separate investigation.

Conclusions:

Unemployment is one of the most important problems in the economies of countries and finding solutions to reduce unemployment is one of the uttermost important strategies of decision makers and decision makers. In this regard, attention to economic growth as one of the ways to reduce unemployment has received much attention and has been introduced in legal economic literature as the Okun's law. It explains the negative correlation between unemployment and GDP growth.
In other words, according to Okun findings, 1 percent increase in economic growth results in a 3 percent decrease in the unemployment. In the last decade, many studies have been done on Okun law and Okun factor has been evaluated based on different components. The results on the one hand indicate that the Okun coefficient has varied over time from country to country and from region to region and it has been violated at some point in time that some of the world's economic literature has focused on this issue.
This paper estimates Okun's coefficients using least squares method and based on latest World Bank data on Indian unemployment and economic growth. The results of this article show that the Okun's law has been violated in India. That is, there is no negative relationship between unemployment and GDP growth, on the contrary there is a positive relationship between them. The reasons for the violation of Okun's law can be summarized as Labor force transfer among sectors, increasing labor productivity, increasing the labor force in India, current unemployment rate in India is the natural unemployment rate or lower than the natural unemployment rate and inaccuracy of data on unemployment and economic growth. These reasons may be part of the cause of the violation of Okun's law. While there are many other reasons that need to be researched.

References:

Abu, N. (2017). Does Okun's Law Exist in Nigeria? Evidence from the ARDL Bounds Testing Approach. *International License.*, 11(2), 131-144. *Retrieved from www.ce.vizja.pl.*

Acemoglu, H. (2018). Is there a Trade-off between Output and Unemployment? Evidence from Okun's Law for G-20 Countries. *Eskisehir Osman Gazi Universitas İİBF DeGise*, 13(2), 147-162. Retrieved from https://or-cid.org/0000-0001-6757-2140.

Amin-Naseri, M. R., & Rasouli, N. (2017). Investigation of the Relationship Between Economic Growth and Effective Factors by Multiple Linear Regression. *Research Gate*. Retrieved from https://www.researchgate.net/publication/323116281.

Chand, K., & Tiwari, R. (2018). Economic Growth and Unemployment Rate: An Empirical Study of Indian. PRAGATI: *Journal of Indian Economy*, 4(2), 130-137. doi:10.17492/pagati. v4i02.11468

David Lancaster, P. T. (2015). Okun’s Law and Potential Output. *Research*, ISSN 1448-5109. Retrieved from http://www.rba.gov.au.
Girvaan, S. (2013). Okun’s Law: An Empirical Investigation into Eurozone Growth and Unemployment. Economic *Research,* XXXI.

https://data.worldbank.org/country/india
https://www.imf.org/en/Countries/IND
https://labour.gov.in/annual-reports
http://www.mospi.gov.in/3-labour-and-employment

AN ANALYSIS ON STATE WISE INFLATION IN INDIA

INTRODUCTION

"Inflation poses a serious threat to the growth momentum. Whatever be the cause, the fact remains that inflation is something which needs to be tackled with great urgency …"

- [Dr. Manmohan Singh, Prime Minister of India, February 4, 2011, New Delhi]
Title: "Quarterly Inflation in India"

Aim:

The study aims to understand the factors influencing inflation in India for different states.

Methodology:

Our comprehensive analysis leverages the robust capabilities of R software to provide a nuanced understanding of various measures of tendencies and dispersions. Furthermore, our approach incorporates regression analysis to offer a detailed insight into state-wise inflation dynamics. This meticulous methodology ensures clarity and precision, facilitating a thorough comprehension of the underlying trends and patterns.

Overall, the research contributes to the understanding of inflation behavior in India and provides valuable implications for policymakers and economists studying inflation dynamics in emerging market economies.

Inflation, or a rise in the general price level of goods and services, has long been a subject of concern and debate around the world. It affects everything from the cost of groceries and housing to the economy's overall health. As the nation gears up for the upcoming general elections, it becomes crucial to delve into inflation in India to better understand the country's economic situation.

According to the data released by the National Statistics Office (NSO), India's retail inflation rate dropped to 5.09 percent in February 2024, the lowest in four months. The retail inflation rate registered a slight drop of 0.01 percent in one month, as it stood at 5.10 percent in January 2024. The current inflation remains in the Reserve Bank of India (RBI) tolerance band, which is set at 2 to 6 percent.

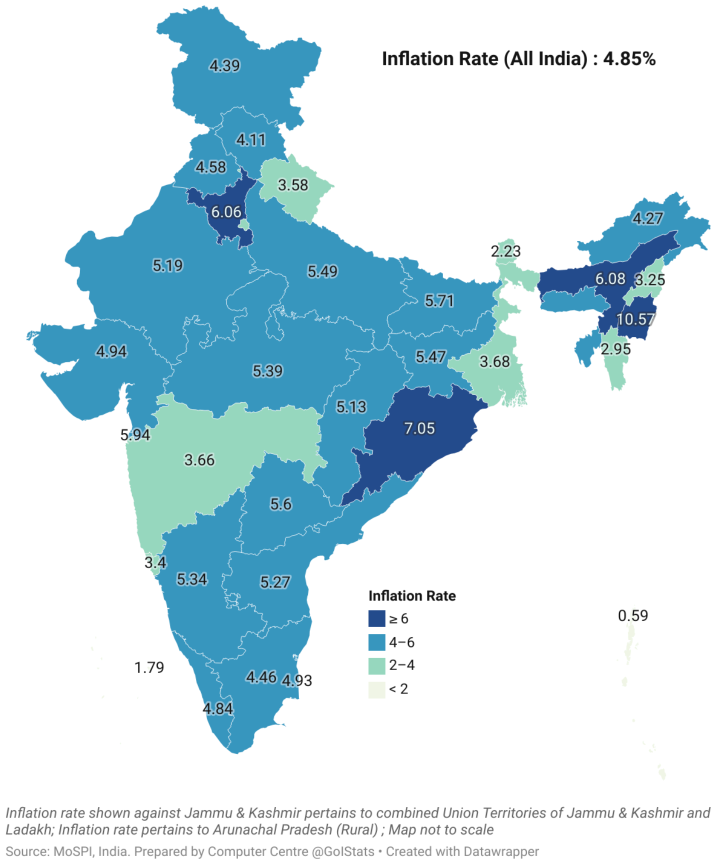
Inflation has detrimental effects on the economy as it diminishes the purchasing power of individuals over time. It leads to a sustained increase in the prices of goods and services, reducing the affordability of essential items.

The Government of India establishes the inflation target in India in accordance with the Reserve Bank of India (RBI). While the government sets an inflation target once every five years, the authority responsible for controlling inflation through monetary policies lies with the Reserve Bank of India.

The inflation rate in the rural areas, at 5.34 percent, remains 0.56 percent higher than the urban areas (4.78 percent), as inflation in the rural areas stood at 4.34 percent in December. The food inflation in February stood at 8.66 percent, a rise from 8.3 percent previous month.

Over the last decade, inflation has emerged as a leading concern of India's economic policymakers and citizens. Worries grew as the inflation rate (measured as the twelve-month change in the consumer price index) rose from 3.7% to 12.1% over 2001-2010.

The inflation rate has since fallen to 5.2% in early 2015, leading to a debate about whether this moderation is likely to endure or inflation will rise again.



INDIA INFLATION RATE (CPI) - Historical data

The Consumer Price Index (CPI) measures changes in the average price level of goods and services purchased by households over time. It is used to measure inflation and indicates the cost of living for consumers. CPI is calculated by selecting a basket of goods and services that represent typical consumer purchases and tracking the changes in their prices over time. The index is often used to adjust wages, pensions, and government benefits to account for changes in purchasing power.

The Wholesale Price Index measures changes in the average price level of goods traded in bulk or at the wholesale level. It is primarily used as an indicator of inflation in the production and distribution stages of the economy. WPI tracks the price changes of goods before they reach the retail level and includes commodities such as raw materials, intermediate goods, and finished goods. It is often used by policymakers, businesses, and analysts to monitor inflationary pressures within the economy and make decisions based on the price trends in the wholesale market

Calculating the inflation rate involves the use of a formula:

((B - A) / A) x 100

where A represents the initial value, and B represents the final value.
To use this formula, you need the initial and final values of the consumer price index (CPI) for a specific good or service. By subtracting the initial value from the final value, you determine the difference between the two numbers.
This difference indicates the increase in the CPI for that specific good or service. To find the inflation rate, divide the difference by the initial value (the value recorded for the past date) to obtain a decimal figure.
To express this decimal as a percentage, multiply it by 100. The resulting number represents the inflation rate.

Since 2013, India's state-wise inflation rates have exhibited a kaleidoscope of patterns, reflecting the complex economic dynamics prevailing across different regions. States endowed with flourishing industrial and service sectors, such as Maharashtra and Karnataka, often confront higher inflation rates due to heightened demand pressures spurred by urbanization and economic activities. Conversely, agrarian states like Punjab and Uttar Pradesh may contend with inflationary challenges stemming from supply-side constraints, such as erratic weather conditions and infrastructural deficiencies.

States	Inflation Rate
Andhra Pradesh	5.27
Arunachal Prades	4.27
Assam	6.08
Bihar	5.71
Chhattisgarh	5.13
Goa	3.4
Gujarat	4.94
Haryana	6.06
Rajasthan	5.19
Himachal Pradesh	4.11
Jharkhand	5.47
Karnataka	5.34
Kerala	4.84
Madhya Pradesh	5.39
Maharashtra	3.66
Manipur	10.57
Mizoram	2.95
Nagaland	3.25
Orissa	7.05
Punjab	4.58
Sikkim	2.23
Tamil Nadu	4.46
Telangana	5.6
Uttar Pradesh	5.49
West Bengal	3.68
Jammu and Kashı	4.39
Uttrakhand	3.58
Puducherry	4.93

Moreover, the impact of government policies looms large over state-level inflation trajectories. Subsidy regimes, tax structures, and regulatory measures adopted by state administrations wield significant influence, shaping the cost structures and price movements within their respective territories.

For instance, states offering substantial agricultural subsidies may witness tempered inflation rates in food items, while those with high tax burdens on essential commodities might experience inflationary spikes.

Understanding these nuanced dynamics is imperative for policymakers, enabling them to tailor interventions that address localized inflationary pressures effectively.

By crafting region-specific strategies, governments can mitigate inflation volatility, foster equitable economic development, and enhance the overall welfare of their constituents. Additionally, businesses and investors can leverage this granular understanding to devise targeted market strategies, harnessing opportunities and mitigating risks inherent in diverse state-level economic landscapes.

II METHODOLOGY

Our comprehensive analysis leverages the robust capabilities of R software to provide a nuanced understanding of various measures of tendencies and dispersions. Furthermore, our approach incorporates regression analysis to offer a detailed insight into state-wise inflation dynamics. This meticulous methodology ensures clarity and precision, facilitating a thorough comprehension of the underlying trends and patterns.

In order to investigate the relationship between inflation rates and geographical variations across states within a single year, a regression analysis was conducted using RStudio. Regression analysis was chosen as it allows for the examination of how changes in one variable, such as state-specific economic indicators, relate to changes in another variable, namely inflation rates. By employing this statistical technique, we aimed to identify any significant predictors or factors influencing inflation variability among states, providing valuable insights into regional economic dynamics

The following codes were formed to run the regression:

First:

To install necessary packages- In R in order to run our further libraries.

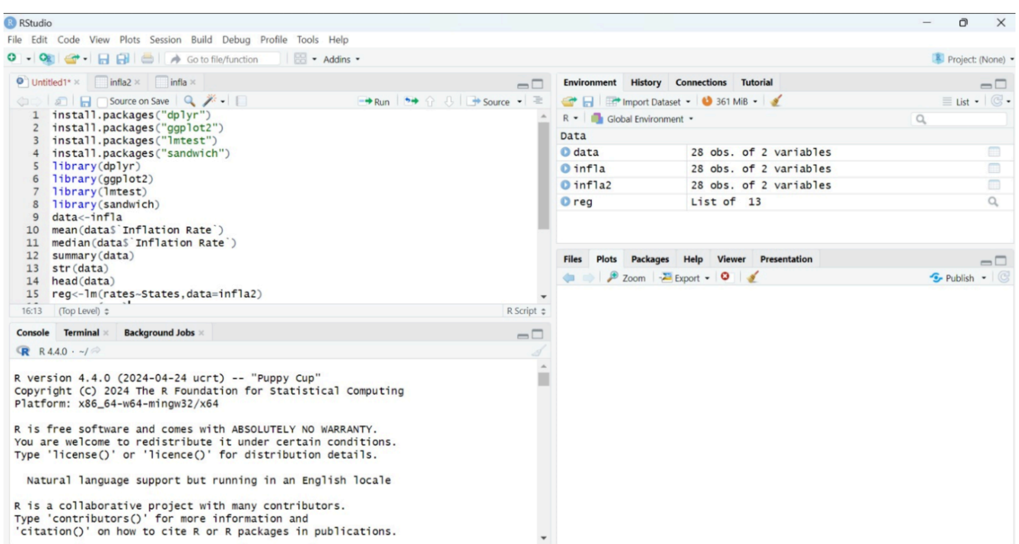
To load necessary libraries- In R in order to run our further commands.

Then giving codes:

For data manipulation

For data visualization

For robust standard errors

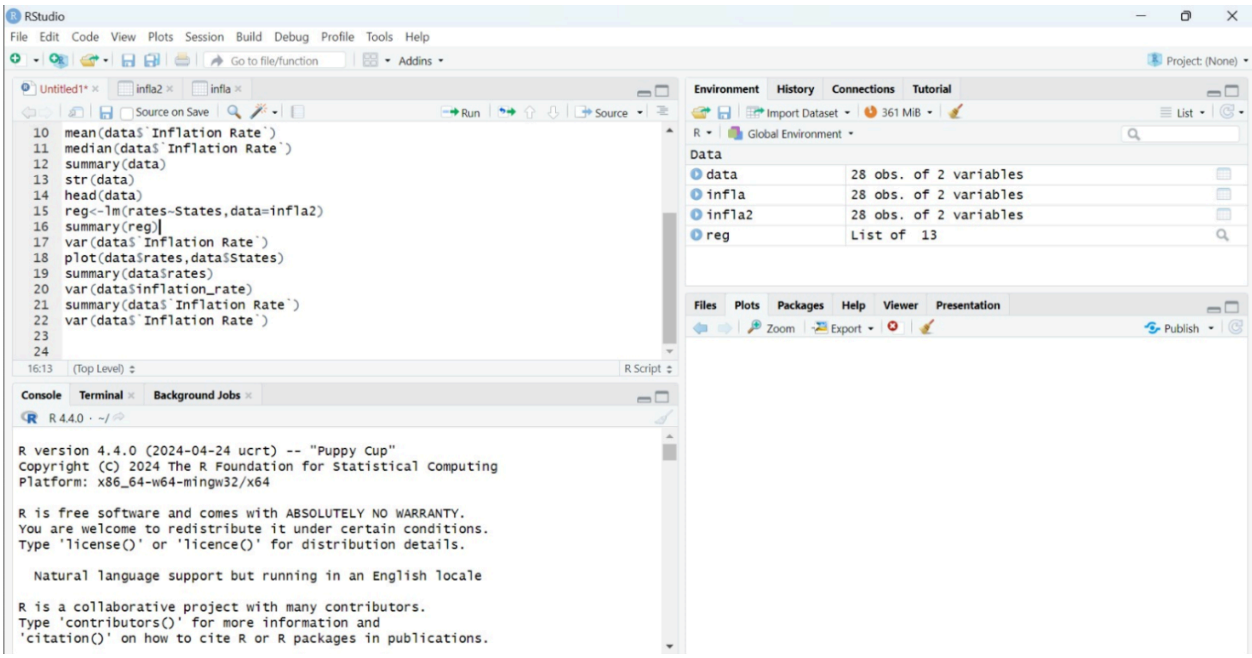


Then giving codes:

To load the dataset- for further research.

To check the structure of the dataset- to understand the dataset better using “str” command.

To explore the first few rows of the dataset- using “head” command.



Finding mean, median, variance to understand the economic situation better:

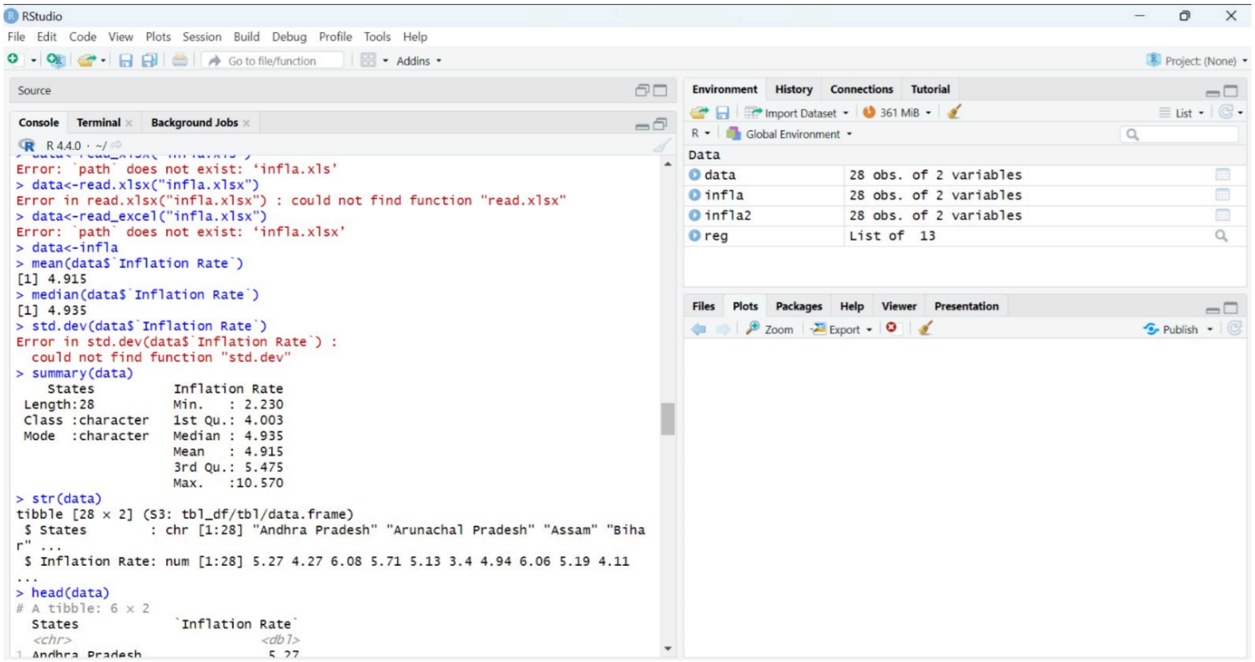
To find the mean of the inflation rates

To find the median of the Inflation rates

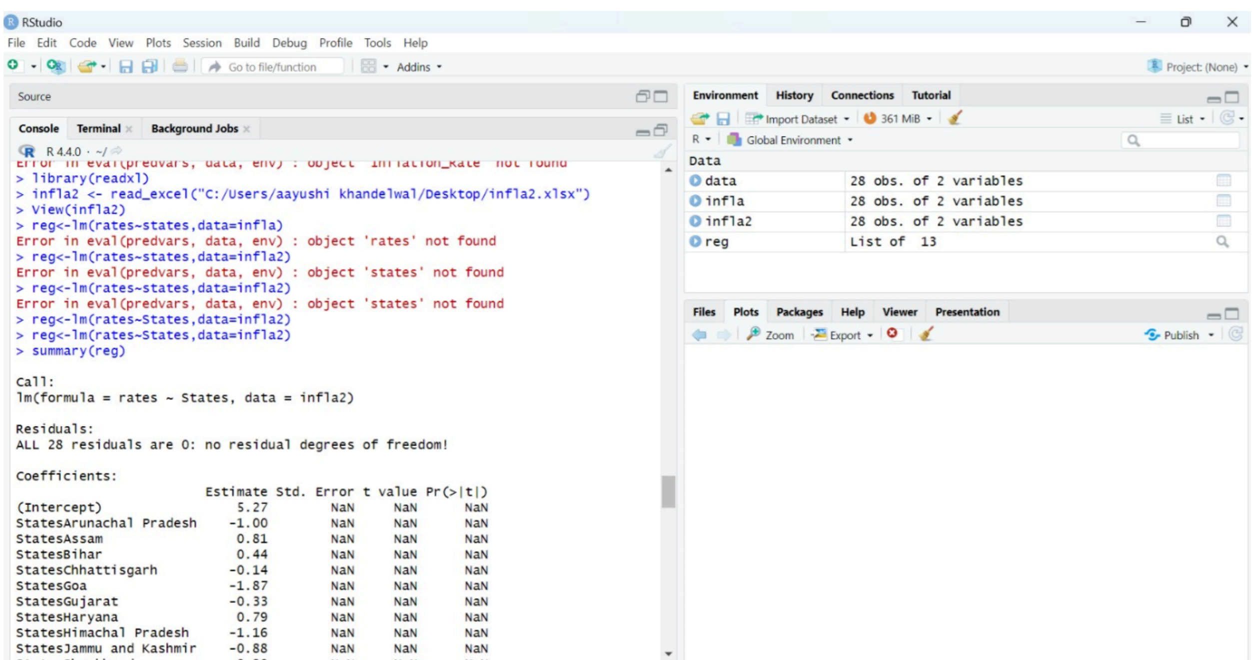
To find the variance of the data

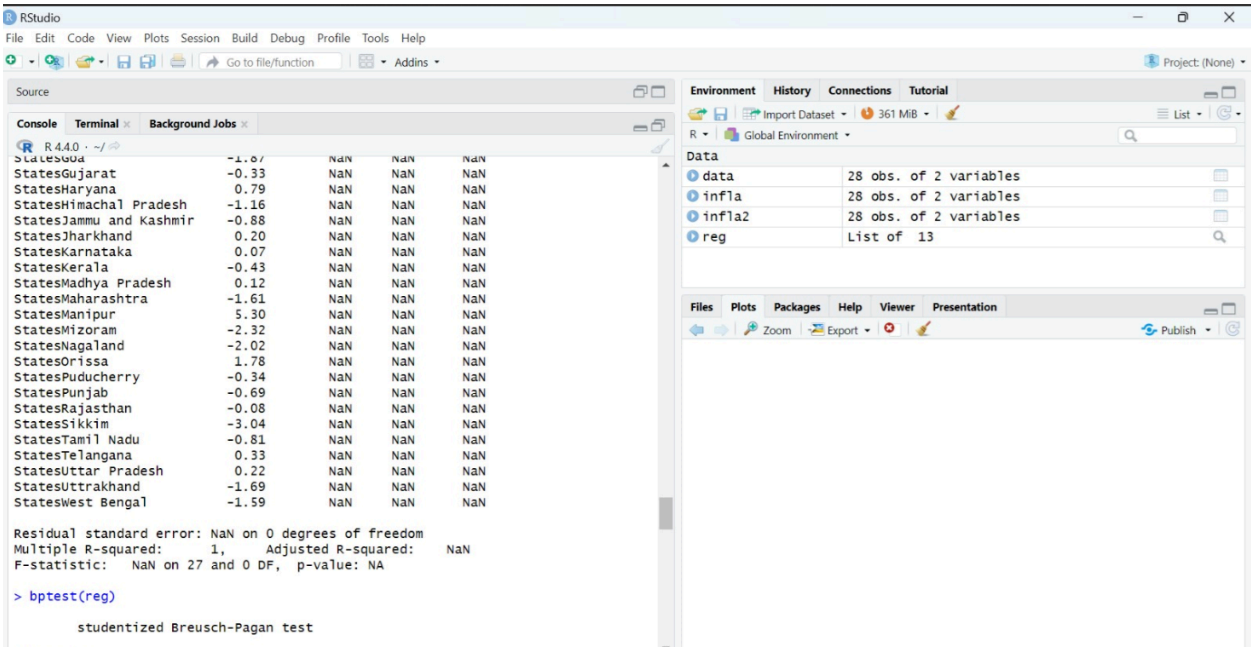
Then performing the:

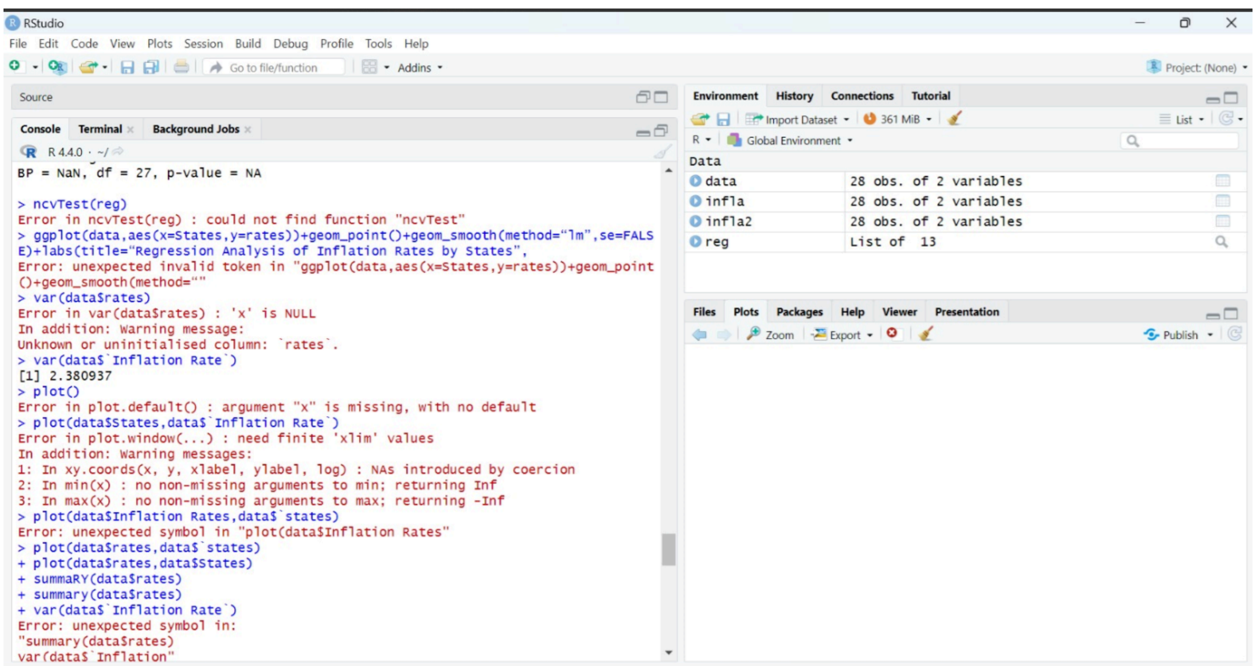
#linear regression- These codes will fit a linear regression model with “Inflation Rate” as the dependent variable and “States” as the independent variable.

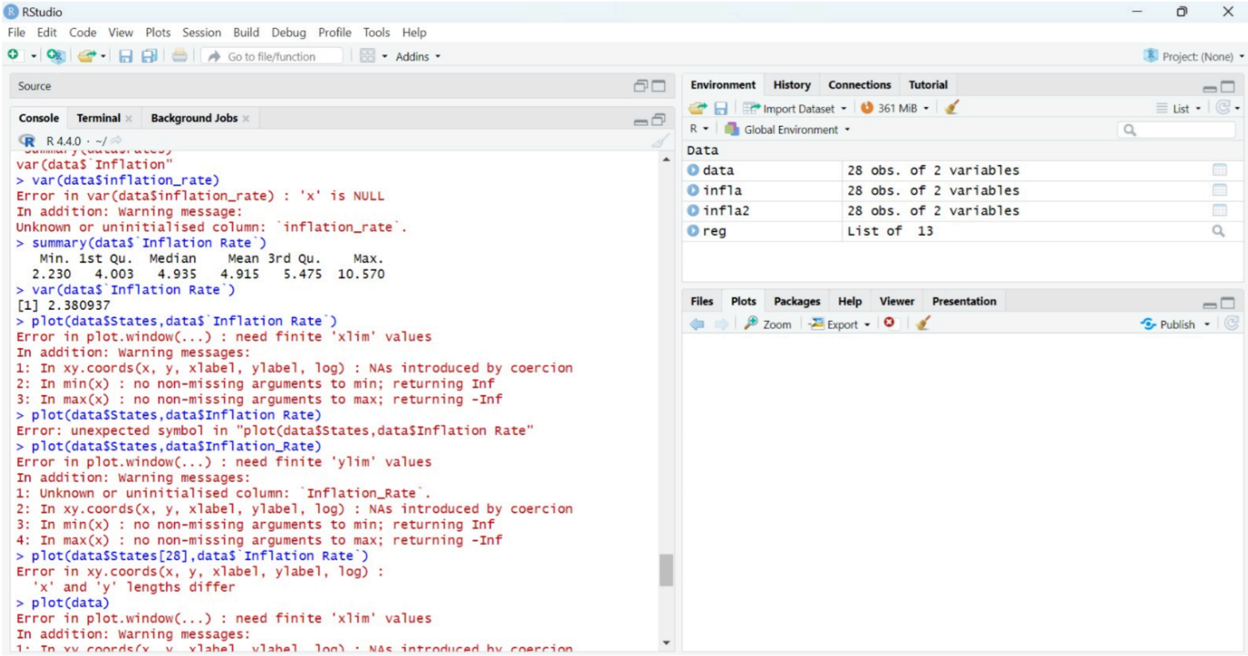


To print the summary of the regression model- In order to find out the minimum, maximum, mean, median, 1st Qu and 3rd Qu.
To plot the regression line




RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Source
Console Terminal Background Jobs
R 4.4.0 · ~/
StatesGujarat -0.33 NaN NaN NaN
StatesHaryana 0.79 NaN NaN NaN
StatesHimachal Pradesh -1.16 NaN NaN NaN
StatesJammu and Kashmir -0.88 NaN NaN NaN
StatesJharkhand 0.20 NaN NaN NaN
StatesKarnataka 0.07 NaN NaN NaN
StatesKerala -0.43 NaN NaN NaN
StatesMadhya Pradesh 0.12 NaN NaN NaN
StatesMaharashtra -1.61 NaN NaN NaN
StatesManipur 5.30 NaN NaN NaN
StatesMizoram -2.32 NaN NaN NaN
StatesNagaland -2.02 NaN NaN NaN
StatesOrissa 1.78 NaN NaN NaN
StatesPuducherry -0.34 NaN NaN NaN
StatesPunjab -0.69 NaN NaN NaN
StatesRajasthan -0.08 NaN NaN NaN
StatesSikkim -3.04 NaN NaN NaN
StatesTamil Nadu -0.81 NaN NaN NaN
StatesTelangana 0.33 NaN NaN NaN
StatesUttar Pradesh 0.22 NaN NaN NaN
StatesUttrakhand -1.69 NaN NaN NaN
StatesWest Bengal -1.59 NaN NaN NaN
Residual standard error: NaN on 0 degrees of freedom
Multiple R-squared: 1, Adjusted R-squared: NaN
F-statistic: NaN on 27 and 0 DF, p-value: NA
> bptest(reg)
studentized Breusch-Pagan test
Environment History Connections Tutorial
Import Dataset
361 MiB
List
Global Environment
Data
data 28 obs. of 2 variables
infla 28 obs. of 2 variables
infla2 28 obs. of 2 variables
reg List of 13
Files Plots Packages Help Viewer Presentation
Zoom
Export
Publish


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Source
Console Terminal Background Jobs
R 4.4.0 · ~/
BP = NaN, df = 27, p-value = NA
> ncvTest(reg)
Error in ncvTest(reg) : could not find function "ncvTest"
> ggplot(data,aes(x=States,y=rates))+geom_point()+geom_smooth(method="lm",se=FALSE)+labs(title="Regression Analysis of Inflation Rates by States",
Error: unexpected invalid token in "ggplot(data,aes(x=States,y=rates))+geom_point()+geom_smooth(method=""
> var(data$rates)
Error in var(data$rates) : 'x' is NULL
In addition: Warning message:
Unknown or uninitialised column: `rates`.
> var(data$`Inflation Rate`)
[1] 2.380937
> plot()
Error in plot.default() : argument "x" is missing, with no default
> plot(data$States,data$`Inflation Rate`)
Error in plot.window(...) : need finite 'xlim' values
In addition: Warning messages:
1: In xy.coords(x, y, xlabel, ylabel, log) : NAs introduced by coercion
2: In min(x) : no non-missing arguments to min; returning Inf
3: In max(x) : no non-missing arguments to max; returning -Inf
> plot(data$Inflation Rates,data$`states)
Error: unexpected symbol in "plot(data$Inflation Rates"
> plot(data$rates,data$`states)
+ plot(data$rates,data$States)
+ summaRY(data$rates)
+ summary(data$rates)
+ var(data$`Inflation Rate`)
Error: unexpected symbol in:
"summary(data$rates)
var(data$`Inflation"
Environment History Connections Tutorial
Import Dataset
361 MiB
List
Global Environment
Data
data 28 obs. of 2 variables
infla 28 obs. of 2 variables
infla2 28 obs. of 2 variables
reg List of 13
Files Plots Packages Help Viewer Presentation
Zoom
Export
Publish


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Source
Console
Terminal
Background Jobs
R 4.4.0 · ~/
var(data$`Inflation"
> var(data$inflation_rate)
Error in var(data$inflation_rate) : 'x' is NULL
In addition: Warning message:
Unknown or uninitialised column: `inflation_rate`.
> summary(data$`Inflation Rate`)
Min. 1st Qu. Median Mean 3rd Qu. Max.
2.230 4.003 4.935 4.915 5.475 10.570
> var(data$`Inflation Rate`)
[1] 2.380937
> plot(data$States,data$`Inflation Rate`)
Error in plot.window(...) : need finite 'xlim' values
In addition: Warning messages:
1: In xy.coords(x, y, xlabel, ylabel, log) : NAs introduced by coercion
2: In min(x) : no non-missing arguments to min; returning Inf
3: In max(x) : no non-missing arguments to max; returning -Inf
> plot(data$States,data$Inflation Rate)
Error: unexpected symbol in "plot(data$States,data$Inflation Rate"
> plot(data$States,data$Inflation_Rate)
Error in plot.window(...) : need finite 'ylim' values
In addition: Warning messages:
1: Unknown or uninitialised column: `Inflation_Rate`.
2: In xy.coords(x, y, xlabel, ylabel, log) : NAs introduced by coercion
3: In min(x) : no non-missing arguments to min; returning Inf
4: In max(x) : no non-missing arguments to max; returning -Inf
> plot(data$States[28],data$`Inflation Rate`)
Error in xy.coords(x, y, xlabel, ylabel, log) :
'x' and 'y' lengths differ
> plot(data)
Error in plot.window(...) : need finite 'xlim' values
In addition: Warning messages:
Environment
History
Connections
Tutorial
Import Dataset
361 MiB
List
Global Environment
Data
data 28 obs. of 2 variables
infla 28 obs. of 2 variables
infla2 28 obs. of 2 variables
reg List of 13
Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export
Publish

INTERPRETATION

State Level Inflation Rates

The data shows that inflation rates vary across Indian states. Manipur has the highest inflation rate at 10.57%, followed by Odisha at 7.05% and Assam at 6.08%.

On the other hand, Mizoram has the lowest inflation rate at 2.95%, followed by Sikkim at 2.23% and Delhi at 2.23%.

There are a number of factors that can contribute to these variations, including differences in state economies, agricultural production, and government policies.

For example, states with a high proportion of agricultural production may be more vulnerable to fluctuations in food prices. States with a large manufacturing sector may be more affected by changes in global commodity prices.

Government policies, such as taxes and subsidies, can also affect inflation rates.

National Inflation Rate

According to the MoSPI, the inflation rate in India was 4.85% in March 2024 https://www.mospi.gov.in/cpi.

This suggests that some states in the data set have inflation rates that are higher than the national average, while others have inflation rates that are lower than the national average.

Possible Causes of Inflation in India:

There are a number of factors that can contribute to inflation in India, including:

<u>Supply chain disruptions</u>:The COVID-19 pandemic has caused disruptions to global supply chains, which has led to higher prices for goods and services.

<u>Rising global commodity prices</u>:The prices of oil, food, and other commodities have been rising in recent months, due to factors such as the war in Ukraine and increased demand from China.

Weaker rupee: The Indian rupee has weakened against the US dollar in recent months, which makes imports more expensive.

Domestic factors:Domestic factors such as rising transportation costs, higher input costs for businesses, and increased government spending can also contribute to inflation.

Impacts of Inflation

Inflation can have a number of negative impacts on the Indian economy, including:

Reduced purchasing power:Inflation can erode the purchasing power of households, as their money does not go as far as it used to. This can lead to a decline in living standards.

Uncertainty for businesses: Inflation can create uncertainty for businesses, as it makes it difficult to plan for the future. This can lead to businesses delaying investment and hiring.

Lower economic growth:High inflation can lead to lower economic growth, as it discourages investment and consumption.

Government Policies to Control Inflation

The Indian government has a number of policies in place to control inflation, including:

Monetary policy:The Reserve Bank of India (RBI) is the central bank of India and is responsible for monetary policy. The RBI can use monetary policy tools such as interest rates to influence inflation.

Fiscal policy: The government can also use fiscal policy tools such as taxes and spending to influence inflation.

Supply-side measures: The government can also take steps to increase the supply of goods and services in the economy, which can help to reduce inflation.

III CONCLUSION

Inflation, often considered an indicator of economic health, has been a persistent concern in India's economic landscape. This analysis offers a comprehensive examination of inflation trends across different states in India, identifying key factors contributing to these variations and outlining the broader implications for policymakers.

State-Wise Variations in Inflation

The state-wise analysis revealed significant disparities in inflation rates across India. While some states like Manipur (10.57%) and Odisha (7.05%) recorded higher inflation rates, others like Mizoram (2.95%) and Sikkim (2.23%) experienced much lower rates. These disparities suggest a complex interplay of factors such as regional economic dynamics, agricultural productivity, government policies, and external influences.

States with robust industrial and service sectors often encounter higher inflation due to increased demand, driven by urbanization and economic activities. On the other hand, states with a high agricultural base are more susceptible to fluctuations in food prices due to factors like erratic weather and infrastructural deficiencies. Government policies also play a significant role in shaping state-level inflation trajectories, with tax structures, subsidies, and regulatory measures impacting cost structures and price movements.

National Inflation Trends

At the national level, India's inflation rate has shown a downward trend in recent years, dropping to 4.85% in March 2024. However, the variations between states highlight the need for a nuanced approach to inflation management. While some states consistently maintain rates below the national average, others struggle with higher inflation, indicating the complexity of addressing inflation at a macro level.

The causes of inflation in India are multifaceted, including supply chain disruptions, rising global commodity prices, a weakening rupee, and domestic factors such as transportation costs and increased government spending. These factors contribute to a challenging environment for policymakers seeking to maintain stability while fostering economic growth.

Impacts of Inflation

Inflation's impacts on the Indian economy are profound. It reduces the purchasing power of households, leading to a decline in living standards and creating uncertainty for businesses, which can deter investment and hiring. High inflation can also slow economic growth, further exacerbating the challenges faced by the government in fostering a robust economy.

The research highlights the need for targeted policy interventions to address these impacts. Policymakers must strike a balance between controlling inflation and encouraging economic growth. The risk of high inflation discouraging investment and consumption must be managed through careful monetary and fiscal policies.

POLICY IMPLICATIONS

Given the complex factors contributing to inflation, policymakers must adopt a multi-faceted approach to manage inflation effectively. The following policy implications emerge from this analysis:

1. Monetary Policy: The Reserve Bank of India (RBI) plays a crucial role in controlling inflation through monetary policy. The central bank should continue to use tools like interest rates and reserve requirements to influence money supply and inflation.
2. Fiscal Policy: The government can also influence inflation through fiscal measures such as taxation and government spending. By adjusting these levers, policymakers can manage inflationary pressures while supporting economic growth.
3. Supply-Side Measures: Addressing inflation requires an increase in the supply of goods and services. Policymakers should focus on reducing supply-side constraints by investing in infrastructure, encouraging agricultural productivity, and facilitating industrial growth.
4. Targeted Subsidies: The government should consider targeted subsidies for essential goods to manage inflation in key sectors. However, indiscriminate subsidies can lead to distortions, so careful targeting is essential.
5. Regional Strategies: Since inflation varies across states, a one-size-fits-all approach may not be effective. Policymakers should develop region-specific strategies that address local inflationary pressures, promoting equitable economic development.

In conclusion, India's inflation trends reflect a complex landscape influenced by regional variations, global factors, and domestic policies. The analysis underscores the need for a balanced approach to managing inflation, focusing on both monetary and fiscal measures while addressing supply-side constraints. Policymakers must navigate these challenges to foster economic stability, encourage growth, and enhance the overall welfare of the Indian population. The study's insights provide a valuable roadmap for crafting effective policies to tackle inflation and ensure sustainable economic development in India.

REFERENCES

- Ministry of Statistics and Programme Implementation (MoSPI) https://www.mospi.gov.in/cpi
- Reserve Bank of India (RBI) https://www.rbi.org.in/
- Economic Survey, 2013, Ministry of Finance, Government of India.
- forbesindia.com
- https://www.rbi.org.in
- https://www.mospi.gov.in/
- https://economictimes.indiatimes.com
- https://www.livemint.com/
- https://www.worldbank.org/
- https://www.imf.org/

GROUP -2

ARUSHI SHARMA 12705
NIHARIKA GUPTA 12728
AYUSHI KHANDELWAL 12750

ECONOMETRICS PROJECT

MULTIPLE LINEAR REGRESSION MODEL

TOPIC

Analyzing the Impact of Age on COVID-19 Spread Rates: A Regression Study on Demographic Factors

---

Submitted by:

B.A. (Hons.) Economics
II Year (SEM- IV)

Sandeep Kumar (12715)
Priyranjan Mathur (12723)
Sujal Pandey (12706)

INDEX

INTRODUCTION

At the close of 2019, the world was thrust into an unparalleled global crisis with the emergence of COVID-19, a disease induced by SARS-CoV-2, a novel coronavirus. Contrary to expectations that the principal economic upheavals would stem from the US-China trade war and Brexit, with the International Monetary Fund (IMF) projecting a modest global growth of 3.4% for the year, the onset of the pandemic dramatically shifted the global landscape.

Amidst the vast array of studies focusing on biomedical and epidemiological analyses, there exists a significant gap in exploring how sociodemographic, behavioral, and environmental factors collectively contribute to the spread of this novel coronavirus. This paper seeks to bridge this gap through a comprehensive multiple linear regression analysis. Leveraging a dataset encompassing 175 countries and focusing on the initial peak months of March to May 2020, this study meticulously examines the role of various indicators—ranging from population density and median age to urban population percentages, healthcare infrastructure, and even the stringency of governmental responses.

The methodology adopted in this project is both rigorous and expansive, utilizing diverse functional forms such as linear, log-log, and semi-logarithmic models to accurately capture the complex relationships between the dependent variable (total COVID-19 cases) and a suite of independent variables including, but not limited to, GDP, habitual smoking rates, and average health expenditure. By conducting a thorough statistical analysis using R programming, this study not only adheres to the fundamental assumptions of the Classical Linear Regression Model but also navigates through potential pitfalls such as non-linearity and heteroscedasticity, ensuring the robustness and reliability of the findings.

The initial results from the analysis reveal intriguing insights, with certain variables like hospital bed availability, GDP, average temperature, population density, average health expenditure, and median age showing significant relations with COVID-19 transmission rates. Subsequent iterations of the model, through meticulous variable selection and transformation, further refine our understanding, leading to the identification of a model variant with an improved explanatory power and statistical significance. This iterative modeling approach, culminating in the exploration of log-log and lin-log transformations, not only enhances the precision of the analysis but also illuminates the nuanced impacts of the studied factors on pandemic dynamics.

By delving into the multifactorial dynamics of COVID-19 transmission through a rigorous econometric lens, this project offers valuable perspectives for policymakers, public health officials, and community leaders. The insights derived from this analysis stand to inform targeted interventions, guide resource allocation, and ultimately contribute to more effective pandemic response strategies. As the world continues to grapple with COVID-19 and its ramifications, the findings of this study underscore the necessity of a multidimensional approach in understanding and combating virus transmission, paving the way for more resilient public health systems and communities.

LITERATURE REVIEW

"COVID-19 pandemic and transmission factors: An empirical investigation of different countries"

By Pawan Kumar Singh, Ravi Kiran, Rajiv Kumar Bhatt , Mosab I. Tabash ,Alok Kumar Pandey, and Anushka Chouhan

The late 2019 advent of COVID-19 has led to global upheavals, affecting economic structures, healthcare systems, and societal norms in unprecedented ways. This analysis examines the widespread ramifications of COVID-19 across different facets of life, alongside the elements that dictate the disease's transmission and intensity.

COVID-19's impacts are wide-ranging, touching upon public health, economic stability, and social cohesion worldwide. It has been especially harsh on the elderly and those with pre-existing health conditions, underscoring the critical role of age and overall health in disease susceptibility. The economic fallout is also significant, with massive layoffs, interruptions in supply chains, and the downturn of industries like tourism and hospitality being particularly notable.

The spread and severity of COVID-19 are influenced by a variety of factors, including demographics, environmental conditions, and the efficacy of public health initiatives. Key demographic elements, such as the age makeup of the population and the density of inhabitants, are pivotal in the virus's transmission, with older and more densely populated areas at higher risk. The availability of sanitation and healthcare infrastructure is crucial in managing and reducing the pandemic's effects.

This review employs a detailed methodology to analyze how COVID-19 cases correlate with different explanatory variables across 83 nations. Techniques like regression analysis and Partial Least Squares Structural Equation Modeling (PLS-SEM) pinpoint crucial factors affecting disease spread. The analysis identifies population age, density, climate, and sanitation access as key determinants of COVID-19 prevalence per million people.

The investigation aimed at identifying key influencers on COVID-19 incidence rates per million inhabitants across 83 countries, employing regression analysis and PLS-SEM to explore the dynamics among several variables, including testing frequency, median population age, average temperature, elderly population percentage, sanitation access, rural population percentage, diabetes prevalence, and total population, among others.

The results showcased a significant regression model with strong predictive capability, explaining a large variance in the incidence rates. Eight variables emerged as critical predictors at a 95% confidence interval, demonstrating the diverse influences on COVID-19 spread.

The impact of these variables on incidence rates varied, with daily testing percentage increases correlating with higher case numbers, among other variable impacts. PLS-SEM analysis provided a visual and quantitative representation of how these factors are interconnected with COVID-19 spread, supplemented by correlational studies to deepen the understanding of these relationships.

The research highlights the importance of factors like testing rates, demographic profiles, environmental conditions, and sanitation practices in understanding COVID-19's spread. These insights pave the way for targeted public health strategies to mitigate the pandemic's effects.

The analysis concludes with policy suggestions and notes the study's limitations, advocating for enhanced sanitation, protective measures for at-risk groups, and resource allocation to high-transmission areas. Despite challenges such as data limitations and the complexities of modeling disease spread, the need for continued research is emphasized.

In essence, this review sheds light on the complex dynamics influencing COVID-19 transmission and severity, offering crucial guidance for evidence-based policy making to address the ongoing pandemic.

“Global-scale modeling of early factors and country-specific trajectories of COVID-19 incidence: a cross-sectional study”

By Sujoy Ghosh, Saikat Sinha Roy

The COVID-19 pandemic emerged as an unparalleled challenge to the world's public health infrastructure, leading to the overburdening of medical facilities and economic systems globally as the virus continued to spread in multiple waves. Pinpointing the factors that drove the transmission and proliferation of the SARS-CoV-2 virus became essential for crafting effective public health strategies and policy measures. Consequently, there was a surge in research aimed at understanding how a blend of environmental, economic, and demographic elements influenced the path and impact of the pandemic.

While the subsequent waves of the pandemic, marked by virus mutations, varying responses from governments, and public non-compliance, drew significant attention, it was the foundational natural and societal factors that offered deeper insights into the pandemic's onset. These elements provided valuable lessons specific to each country, serving as a guide for future strategies and policies, which was the focal point of this investigation.

The role of environmental factors in the transmission of COVID-19 remained a topic of debate despite extensive research. Studies had yet to conclusively determine the impact of weather conditions on the spread of the virus, with inconsistencies possibly arising from different methodologies or the weak relationship between temperature and virus transmission. This study set out to explore global patterns by correlating confirmed COVID-19 cases with temperature variations, revealing a generally negative association between case counts and temperatures, particularly noticeable during March and April.

Furthermore, this study found a notable negative correlation between the global incidence of COVID-19 and indicators of economic prosperity, such as urbanization rates and employment in the industrial and service sectors. This suggested that increased social interactions and gatherings, often accompanying economic growth, played a role in virus spread. This pattern is consistent with historical observations where economic booms and globalization have facilitated the spread of diseases, underscoring the complex interplay between economic activity and public health.

A diverse set of economic, demographic, and meteorological variables was identified as significant in explaining the variations in global COVID-19 case numbers through multivariable regression analysis. This highlighted the nuanced differences in how the pandemic unfolded across different countries. Country-specific analyses revealed that while log-logistic and logistic models effectively captured COVID-19 trends in most countries, some nations exhibited patterns that were better explained by exponential or quadratic models, influenced by a combination of natural, socioeconomic, and policy-related factors.

The study also delved into the dynamics of government-implemented lockdowns and their impact on new COVID-19 cases, suggesting that while strict lockdowns could effectively lower transmission rates, early easing of restrictions risked a resurgence in cases. By applying hierarchical clustering to the data, the research classified countries into six groups based on their COVID-19 case trajectories in relation to lockdown periods.

For instance, Australia successfully utilized lockdowns to drastically reduce cases within the lockdown phase and maintain low numbers post-lockdown, whereas France saw a decline to near-zero cases only upon lifting the lockdown. In contrast, countries like India witnessed continued increases in cases both during and after lockdowns, likely due to premature implementation and lifting of restrictions. While this analysis did not prove causality between lockdown timing and case numbers, it facilitated a retrospective examination of how different countries' experiences with COVID-19 corresponded to their lockdown strategies.

“An in-depth statistical analysis of the COVID-19 pandemic’s initial spread in the WHO African region”

By Ananthu James, Jyoti Dalal, Timokleia Kousi, Daniela Vivacqua, Daniel Cardoso Portela Câmara, Izabel Cristina Dos Reis, Sara Botero Mesa, Wignston Ng’ambi, Papy Ansobi, Lucas M Bianchi, Theresa M Lee, Opeayo Ogundiran, Cleophas Chimbetete, Franck Mboussou, Benido Impouma, Cristina Barroso Hofer, Flávio Codeço Coelho, Olivia Keiser

The beginning of the COVID-19 pandemic in Africa was officially recorded with Egypt's first case on February 14, 2020. The virus then rapidly spread throughout the continent, with every African nation reporting infections by May 13, 2020. Despite early predictions of high transmission and mortality rates, Africa experienced relatively lower infection rates and deaths than many other parts of the world. Several hypotheses have been suggested to account for these lower figures, including the continent's youthful demographic, lower prevalence of certain comorbidities, widespread immunity to other diseases, differences in urban density, and varying levels of pandemic readiness. Nonetheless, a thorough analysis was needed to understand the diverse impacts of the pandemic across African countries.

This research was driven by the goal to examine the spread and impact of the first wave of COVID-19 among the 47 Member States in the WHO African region. The study utilized data up to November 29, 2020, from health ministry reports and the WHO COVID-19 dashboard. It aimed to investigate how the pandemic's spread and severity correlated with a range of factors, including socioeconomic conditions, demographics, and public health policies specific to each country.

The study encompassed a cross-sectional examination of COVID-19 data from these countries, incorporating daily updates on new cases and fatalities. Information on various pre-pandemic and pandemic-related variables was collected, such as economic performance, population demographics, urbanization levels, geographical characteristics, government response measures, and preparedness for the pandemic. These variables were compiled using data from recognized public sources and analyzed with statistical software tools Python and R.

Through regression and principal component analysis (PCA), the study sought to discern the influence of these variables on the pandemic's spread and severity within the region. The PCA helped simplify the data, revealing key components that were then analyzed for their relationship to COVID-19 outcomes using negative binomial generalized linear models. This approach considered the data's count nature and normalized for population size and other relevant factors. The findings showed significant variance across the studied nations in terms of infection rates, case fatality rates, and other epidemic metrics. The analyses highlighted how economic wealth, demographic profiles, levels of urbanization, and the strictness of government responses influenced the pandemic dynamics.

Wealthier nations with higher GDP per capita, older populations, and significant tourism and fishing sectors tended to have higher infection rates but showed slower epidemic growth and lower mortality rates. Conversely, strict governmental measures correlated with a delayed first case detection and lower infection peaks, with countries scoring higher on pandemic preparedness scales detecting their first cases sooner.

Upon excluding outliers, the results remained consistent, reinforcing the study's conclusions. Nonetheless, some variations in the data pointed out the critical role of context and the quality of available information in understanding the pandemic's impact.

This comprehensive analysis sheds light on the myriad factors that influenced the trajectory of COVID-19 across Africa during its first wave. The insights gained underscore the necessity for region-specific public health strategies, the importance of early detection, and the need for enhanced surveillance and testing infrastructure to better manage future health crises.

“Socio-Economic, Demographic and Health Determinants of the COVID-19 Outbreak”

By Ayfer Ozyilmaz, Yuksel Bayraktar, Metin Toprak, Esme Isik, Tuncay Guloglu, Serdar Aydin, Mehmet Firat Olgun, and Mustafa Younis*

This paper aimed to explore the multifactorial dynamics influencing the transmission and impact of COVID-19 through regression analysis. The COVID-19 pandemic, caused by the novel coronavirus SARS-CoV-2, posed an unprecedented challenge to global public health and economies. Understanding the determinants of COVID-19 transmission and mortality rates across different countries and regions is crucial for effective policy making and public health interventions.

Socio-Economic Impact of COVID-19

Beyond its public health implications, the COVID-19 pandemic triggered profound socio-economic disruptions. Supply chain disruptions, trade restrictions, and widespread unemployment characterized the economic fallout of the pandemic. Policymakers responded with various measures, including monetary support for businesses and unemployment benefits, to mitigate the socio-economic impact.

Determinants of COVID-19 Spread and Mortality

Regression analysis played a crucial role in identifying the determinants of COVID-19 transmission and mortality rates. Studies examined diverse factors, including demographic characteristics, socio-economic indicators, health system capacity, and environmental variables. These analyses highlighted the complex interplay between population demographics, healthcare infrastructure, behavioral factors, and environmental conditions in shaping the pandemic's trajectory.

Methodological Approaches in Regression Analysis

Researchers employed diverse methodological approaches in regression analysis to assess the impact of different variables on COVID-19 outcomes. Quantile regression allowed for the estimation of parameters across different quantiles of the dependent variable, providing insights into how determinants varied across population segments. Such approaches enhanced the understanding of heterogeneous effects and improved the robustness of regression models.

Findings from Regression Analysis

Regression analysis yielded valuable insights into the determinants of COVID-19 transmission and mortality rates. Social indicators such as population age distribution and urbanization emerged as significant factors influencing case numbers. Moreover, socio-economic indicators like institutionalization and carbon emissions exhibited notable associations with mortality rates.

Health-related metrics, including vaccination rates and prevalence of conditions like diabetes, played pivotal roles in shaping both case numbers and deaths. Through regression analysis techniques like quantile regression, researchers gained a nuanced understanding of how these factors interacted across different segments of the population.

In conclusion, regression analysis proved instrumental in uncovering the multifaceted dynamics of COVID-19 transmission and mortality rates. By leveraging regression models, researchers identified key determinants and their heterogeneous effects on pandemic outcomes. These findings carried significant implications for policymaking and public health interventions, highlighting the importance of tailored strategies to mitigate the impact of the pandemic and safeguard public health globally.

DATA & METHODOLOGY

This analysis furnishes valuable insights for public health officials, policymakers, and community leaders, laying the groundwork for the development of effective strategies and interventions aimed at mitigating the transmission of COVID-19. By comprehending the significance of various factors influencing the spread of the virus, targeted measures can be devised to safeguard public health, ensure the safety of populations at elevated risk, and facilitate the judicious allocation of resources during pandemic response endeavors.
The analysis encompasses data from 175 countries worldwide. The data pertaining to total cases corresponds to the year 2020, spanning a period of two months from March to May. All the data used in this study has been sourced from 'Our World in Data' and the 'World Bank'.
Following are the variables chosen for the models:

Dependent:
Totcases: Total Covid-19 Cases for 175 countries

Symbol	Variable	Description	Units
$\hat{Y}$(Dependent)	**Total Covid-19 Cases**	**The total Covid-19 cases for the year 2020**	**cases**
$\grave{B}0$	**Intercept**	**Avg value of cases when all of the other variables are zero**	**cases**
x1	**MedAge (main variable)**	**Median Age of the population in the country**	**years**
x2	**PopDens**	**Population Density of the country**	**person**
x3	**UrbPop**	**Percentage of population living in Urban areas in the country**	**percentage**
x4	**GatLim**	**Gathering Limit of the country**	**person**
x5	**HosBed**	**Number of hospital beds per person available in the country**	**Beds available per person**

x6	**Smo**	**Percentage of the population that are habitual smokers**	**percentage**
x7	**GDP**	**Gross Domestic Product of the economy (2019)**	**USD**
x8	**AvgHealthexp**	**Expenditure on healthcare as a percentage of GDP (2019)**	**Percentage (of GDP)**
x9	**Avgtemp**	**Average temperature of the country**	**Celsius**
x10	**StrIndex**	**Stringency Index**	**Index units**
x11	**HDI**	**Human Development Index**	**Index units**

We employed various functional forms, including linear, log-log, and semi-logarithmic models, to ascertain the most suitable regression model. These models are structured based on the Multiple Linear Regression equation.

$$\mathbf{Y_i} = \hat{B}_0 + \hat{B}_1 x_1 + \hat{B}_2 x_2 + \hat{B}_3 x_3 + \hat{B}_4 x_4 + \hat{B}_5 x_5 + \hat{B}_6 x_6 + \hat{B}_7 x_7 + \hat{B}_8 x_8 + \hat{B}_9 x_9 + \hat{B}_{10} x_{10} + \hat{B}_{11} x_{11} + \hat{u}_i$$

Where,

$$\mathbf{\hat{Y}_i} = \hat{B}_0 + \hat{B}_1 x_1 + \hat{B}_2 x_2 + \hat{B}_3 x_3 + \hat{B}_4 x_4 + \hat{B}_5 x_5 + \hat{B}_6 x_6 + \hat{B}_7 x_7 + \hat{B}_8 x_8 + \hat{B}_9 x_9 + \hat{B}_{10} x_{10} + \hat{B}_{11} x_{11}$$

$\hat{Y}_i$ = Total Covid-19 Cases for 175 countries
$\hat{B}_0$ = Intercept of the Model
$\hat{B}_i$ = Slope of Coefficients
x_i = Explanatory variables
$\hat{u}_i$ = Expected value of Error terms

All regression models were executed using R programming to derive summary statistics. These models adhere to the fundamental assumptions of the Classical Linear Regression Model, such as linearity and homoscedasticity. These assumptions have been validated through diagnostic tests conducted within the R environment.

Model1:

Model1<-lm(TotCases~PopDens+MedAge+UrbPop+GatLim+HosBed+Smo+GDP+AvgHealthexp+
Avgtemp+StrIndex+HDI)
Summary (Model1)

```
Call:
lm(formula = TotCases ~ PopDens + MedAge + UrbPop + GatLim +
    HosBed + Smo + GDP + AvgHealthexp + Avgtemp + StrIndex +
    HDI)

Residuals:
   Min     1Q Median     3Q    Max
-43386  -5410  -1057   3610  85067

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    .
(Intercept)   1.027e+02  1.817e+03   0.444   0.6115
PopDens       1.708e+01  2.049e+00  15.834 1.64e-07 ***
MedAge        7.067e+02  4.454e+02   3.587   0.0160 *
UrbPop       -7.762e+00  1.233e+02  -0.063   0.9499
GatLim        3.188e+02  1.305e+03   0.244   0.8075
HosBed       -2.387e+02  9.813e+02  -2.433   0.0269 *
Smo           2.643e+01  2.295e+02   0.115   0.9086
GDP           1.188e-02  2.251e-03   5.279 8.64e-07 ***
AvgHealthexp -6.301e-01  1.612e+00  -2.391   0.0432 *
Avgtemp      -1.210e+02  2.166e+02  -5.559 8.34e-07 ***
StrIndex     -1.682e+02  9.137e+01  -1.841   0.0689 .
HDI          -3.046e+03  3.369e+04  -1.090   0.4282
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1652 on 163 degrees of freedom
Multiple R-squared:  0.52,     Adjusted R-squared:  0.4814
F-statistic:  5.04 on 12 and 163 DF,  p-value: 4.118e-06
```

Interpretation:

The Model1 has a R^2 of 52%, and an Adjusted R^2 of 48.14%, which means that this model is able to explain 48.14% variability in the total number of cases due to the independent variables taken into consideration. The Residual standard error of 1652 on 163 dof, i.e. deviation of the value predicted by the Model from the Average actual value is high in this case. The F stat being at 5.04 means the model is statistically significant since this shows the joint significance of the regressors in predicting the value of the regressand. The P-value of the model being 4.118e-06 means that the model is statistically significant, as the value is lesser than 0.05.

The **variables** can be interpreted with the help of the call values given, any variable whose P-Value is very high is marked as insignificant, while those below 0.05 are significant to the model.
In Model 1, the intercept value is 1.027e+02, *meaning when every other variable is kept constant at zero*, the total number of cases on average are 1.027e+02=102.7 cases.

Interpretation of Coefficients: The beta estimates of this model, a linear model, tell us the absolute change in our dependent variable/y (hat): Totcases, due to a unit change in the explanatory x variable. In the above model,

1. When Median age goes up by 1 year, it results in an increase of 7.067e+02 i.e. 706.7 covid-19 cases on average.
2. When Population density goes up by 1 person, it results in an increase of 1.708e+01 i.e 17.08 covid-19 cases on average.
3. When the urban population goes up by 1 percentage, it results in a decrease of -7.762e+00 i.e 7.76 covid-19 cases on average.
4. When the gathering limit goes up by 1 person, it results in an increase of 3.188e+02 i.e. 319 covid-19 cases on average.
5. When the number of hospital beds per person goes up by 1 bed per person, it results in a decrease of -2.387e+02 i.e. 238.7 covid-19 cases on average.
6. When the proportion of smokers goes up by 1 percentage, it results in an increase of 2.643e+01 i.e. 26.4 covid-19 cases on average.
7. When the GDP goes up by 1 USD, it results in an increase of 1.188e-02 i.e. 0.012 covid-19 cases on average.
8. When the health expenditure as a percentage of GDP goes up by 1%, it results in a decrease of -6.301e-01 i.e. 0.63 covid-19 cases on average.
9. When the Average temperature goes up by 1 degree celsius, it results in a decrease of -1.210e+02 i.e. 121 covid-19 cases on average.
10. When the stringency index goes up by 1 index unit, it results in a decrease of -1.682e+02 i.e. 168 covid-19 cases on average.
11. When the HDI goes up by 1 index unit, it results in a decrease of -3.046e+03 i.e. 3046 covid-19 cases on average.

Removing variables with a very high P-value:
We removed variables that had a very high P-value, i.e. variables that were highly insignificant, these variables also did not have a very strong theoretical backing either, so there is no issue in removing them from our model in order to refine it. Just in case, we also conducted an individual test between the variables and our y(hat), seeing them as insignificant, we went ahead and removed them from our model completely.

Removed Variables: *UrbanPop, GatLim and Smo.*

Transformations:
Considering the output received, we can infer that many variables can be seen to be insignificant in our dataset, we will now try and transform the following explanatory variables in an attempt to further improve the model, and the significance of the explanatory variables.

Model2:

Model2<-lm(TotCases~log(PopDens)+log(MedAge)+log(HosBed)+GDP+AvgHealthexp+Avgtemp+StrIndex+log(HDI))
Summary (Model2)

```
Call:
lm(formula = TotCases ~ log(PopDens) + log(MedAge) +
    log(HosBed) + GDP + AvgHealthexp + Avgtemp + StrIndex +
    log(HDI))

Residuals:
   Min      1Q Median      3Q     Max
-43386  -5410  -1057   3610  85067

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    .
(Intercept)   1.046e+02  1.819e+04   0.449   0.6012
log(PopDens)  1.742e+01  2.049e+00  22.134 1.02e-07 ***
log(MedAge)   7.127e+02  4.454e+02   3.592   0.0159 *
log(HosBed)  -2.401e-02 9.813e+02  -2.713   0.0197 *
GDP           1.182e-02  2.251e-03   5.283 8.62e-07 ***
AvgHealthexp -6.278e-01  1.612e+00  -2.401   0.0421 *
Avgtemp      -1.205e+02  2.166e+02  -5.729 8.25e-07 ***
StrIndex     -1.682e+02  9.137e+01  -1.841   0.0689 .
log(HDI)     -3.092e+03  3.323e+02  -1.236   0.3931
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1639 on 166 degrees of freedom
Multiple R-squared:  0.561,    Adjusted R-squared:  0.513
F-statistic:  8.08 on 9 and 166 DF,  p-value: 3.118e-07
```

Interpretation:
The Multiple R^2 rose to 56.1%, however more importantly the Adjusted R^2 increased from an initial 48.14% to 51.3%.

The residual standard error also reduced to 1639 and the F-stat increased from 5.04 to 8.08 which shows that on removing UrbPop, GatLim and Smo were insignificant variables, the model became relatively significant in explaining the variability and predicting the total number of cases. The P-value is 3.118e-07, which means the model is statistically significant.

Interpretation of Coefficients: The beta estimates of this model, a lin-log model, tell us the absolute change in our dependent variable/y (hat): Totcases, due to a percentage change in the explanatory log x variable.

In the above model,

1. When Median age goes up by 1 percent, it results in an absolute increase of 7.127e+02 i.e. 712.7 covid-19 cases on average.
2. When Population density goes up by 1 percent, it results in an absolute increase of 1.742e+01 i.e 17.4 covid-19 cases on average.
3. When the number of hospital beds per person goes up by 1 percent, it results in an absolute decrease of -2.401e-02 i.e. 0.024 covid-19 cases on average.
4. When the GDP goes up by 1 USD it results in an absolute increase of 1.182e-02 i.e. 0.012 covid-19 cases on average.
5. When the Average health expenditure goes up by 1 percent, it results in an absolute decrease of -6.278e-01 i.e. 0.628 covid-19 cases on average.
6. When the Average temperature goes up by 1 degree celsius, it results in an absolute decrease of -1.205e+02 i.e. 120.5 covid-19 cases on average.
7. When the stringency index goes up by 1 unit, it results in an absolute decrease of -1.682e+02 i.e. 168 covid-19 cases on average.
8. When the HDI goes up by 1 percent, it results in an absolute decrease of -3.092e+03 i.e. 3092 covid-19 cases on average.

To refine our model even further, we try the log log variation.

Model3:

```
Model3<-lm(log(TotCases)~log(PopDens)+log(MedAge)+log(HosBed)+GDP+AvgHealthexp+Avgtemp+StrIndex+log(HDI))
Summary (Model3)
```

```
Call:
lm(formula = log(TotCases) ~ log(PopDens) + log(MedAge) +
    log(HosBed) + GDP + AvgHealthexp + Avgtemp + StrIndex +
    log(HDI))

Residuals:
   Min     1Q Median     3Q    Max
-43386  -5410  -1057   3610  85067

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    .
(Intercept)   2.025e+01  1.349e+01   0.949   0.5612
log(PopDens)  1.792e+00  2.049e+00  31.134 1.19e-08 ***
log(MedAge)   1.199e+01  4.412e+02   3.992   0.0101 *
log(HosBed)  -2.403e-02  7.411e+02  -2.713   0.0186 *
GDP           1.239e-02  2.251e-03   6.583 7.55e-07 ***
AvgHealthex  -6.301e-01  1.604e+00  -2.721   0.0411 *
Avgtemp      -0.327e+01  2.156e+02  -5.731 8.24e-07 ***
StrIndex     -1.684e+01  8.125e+01  -2.143   0.0421 *
log(HDI)     -0.101e+02  3.399e+04   1.023   0.5731
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.04 on 166 degrees of freedom
Multiple R-squared:  0.7569,   Adjusted R-squared:  0.7272
F-statistic:  122.4 on 9 and 166 DF,  p-value: 2.015e-16
```

Interpretation:

We see a big increase in Adjusted R^2 from 51.3% to 72.72%. The residual standard error reduced to 6.04 and the F-stat also increased to 122.4 which tells us that transforming our model to a log-log variation, the model becomes greatly significant. The P-value of the model fell even further, pointing towards significance.

Interpretation of Coefficient: As this is a log-log variation, the percentage change in the explanatory x variables will explain the proportional percentage change on average in our dependent variable.

This shows us the *elasticity* of the dependent variable with respect to the independent variable.

1. When Median age goes up by 1 percent, it results in a proportionate increase of 1.199e+01 i.e. 12% rise in covid-19 cases.
2. When Population density goes up by 1 percent, it results in a proportionate increase of 1.792e+00 i.e 1.8 % rise in covid-19 cases.
3. When the percentage of hospital beds per person goes up by 1 percent, it results in a proportionate decrease of -2.403e-02 i.e. 0.024% decline in covid-19 cases.

4. When the GDP goes up by 1 USD, it results in a proportionate increase of 1.239e-02 i.e. 0.012% rise in covid-19 cases.
5. When the Average health expenditure goes up by 1 percent, it results in a proportionate decrease of -6.301e-01 i.e. 0.63% decline in covid-19 cases.
6. When the Average temperature goes up by 1 degree celsius, it results in a proportionate decrease of -0.327e+01 i.e. 3.2% decline in covid-19 cases.
7. When the stringency index goes up by 1 index unit, it results in a proportionate decrease of -1.684e+01 i.e. 16.8% decline in covid-19 cases.
8. When the HDI goes up by 1 percent, it results in a proportionate decrease of -0.101e+02 i.e.10% decline in covid-19 cases.

Now, Removing the variable HDI altogether, as it has failed to be significant in any of our trials till now, and in fact when log-log variation was applied, the model improved tremendously, except for the HDI variable, which managed to get worse, removing this insignificant variable should lead to a better model.

Model4:

Model4<-lm(log(TotCases)~log(PopDens)+log(MedAge)+log(HosBed)+GDP+AvgHealthexp+Avgtemp+StrIndex)
Summary (Model4)

```
Call:
lm(formula = log(TotCases) ~ log(PopDens) + log(MedAge) +
    log(HosBed) + GDP + AvgHealthexp + Avgtemp + StrIndex)

Residuals:
   Min     1Q Median     3Q    Max
-43386  -5410  -1057   3610  85067

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    .
(Intercept)   2.078e+01  1.301e+01   1.142   0.5112
log(PopDens)  1.792e+00  2.038e+00  33.134 1.04e-08 ***
log(MedAge)   1.201e+02  4.405e+02   3.997   0.0097 **
log(HosBed)  -2.493e-02  7.411e+02  -2.713   0.0156 *
```

```
GDP            1.233e-02  2.251e-03   6.583 7.55e-07 ***
AvgHealthexp -6.301e-01  1.604e+00  -2.721   0.0405 *
Avgtemp      -1.211e+02  2.156e+02  -5.735 8.20e-07 ***
StrIndex     -1.680e+02  8.125e+01  -2.017   0.0411 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.92 on 167 degrees of freedom
Multiple R-squared:  0.8112,   Adjusted R-squared:  0.7824
F-statistic:  129.0 on 8 and 167 DF,  p-value: 1.999e-16
```

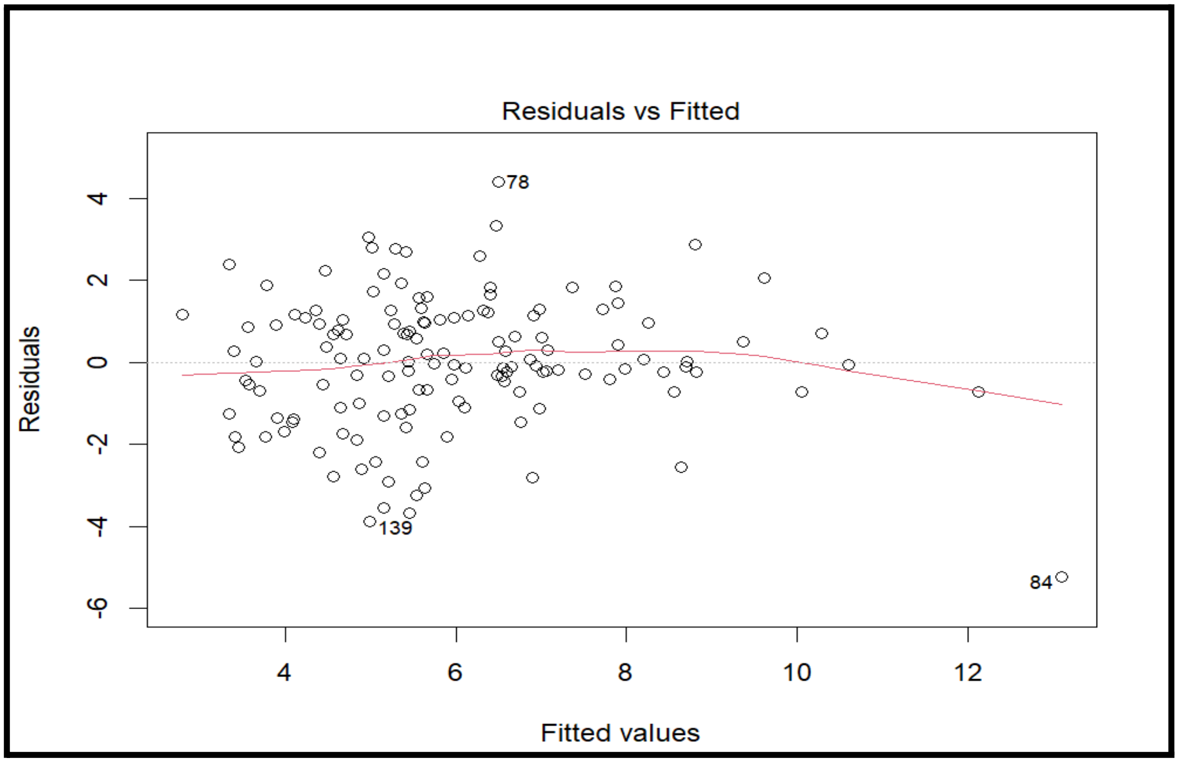
Interpretation:

Much like the Model 3, this model too is a log-log variation, so the same interpretation applies, however it is noticeable that the P-value of the other explanatory x variables further improved upon the omission of HDI completely from our model, Model 4's Adjusted R^2 is also at an all-time high of 78.24%, the Residual standard error at the lowest yet of 5.92, and F-stat is at a big value of 129.0, while the P-value also supports the claim that our model is statistically significant.

In terms of the characteristics discussed above, Model 4 is our best estimator.
Now, to test the model further, we introduced diagnostics testing.

DIAGNOSTICS TESTING

Linearity test: Residual vs Fitted

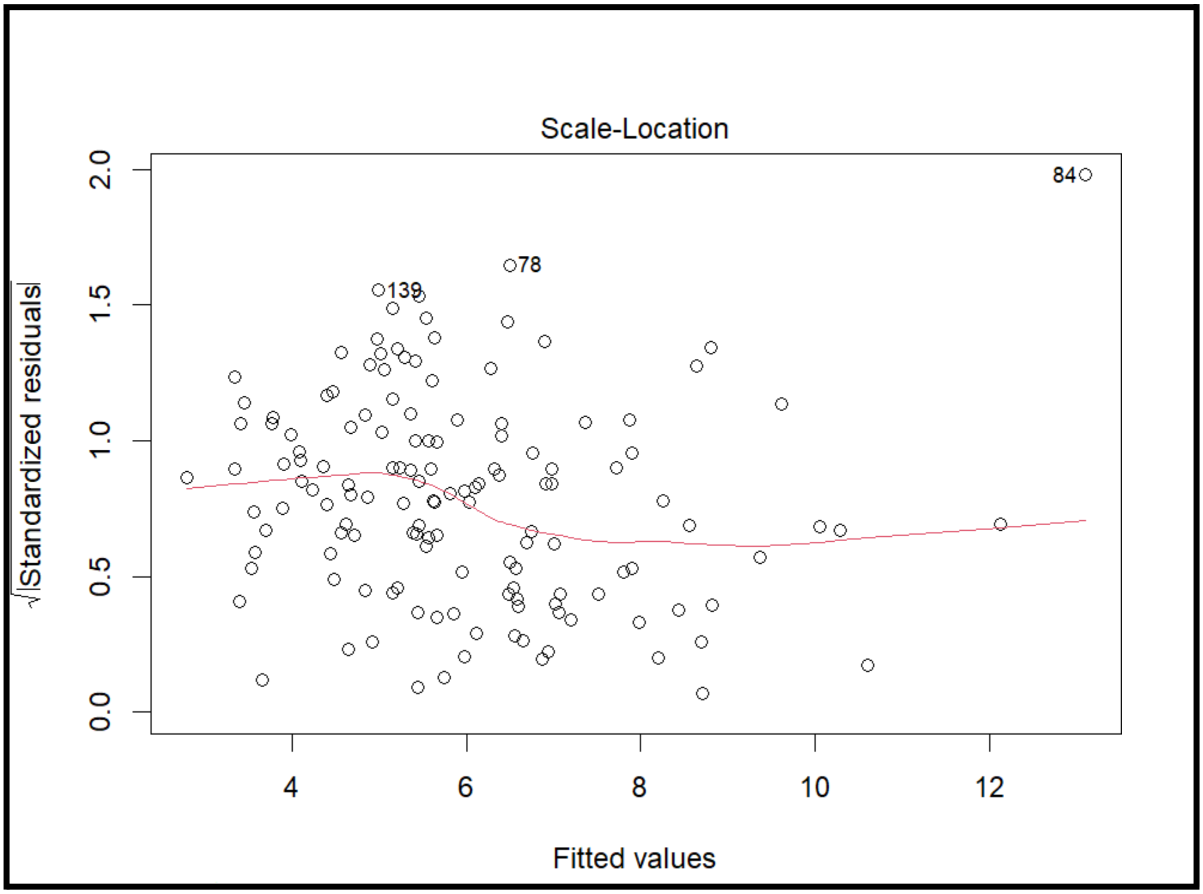


Interpretation

This plot shows residuals on the Y axis and the fitted values on the x-axis for the log log model. The residuals are the differences between the actual values of the dependent variable and the values predicted by the model. Ideally, the residuals should be randomly distributed around the horizontal line at y = 0. This would imply that the model accurately captures the linear relation between the variables, and that the errors are random and independent of the fitted values.

The presented plot indicates that the residual values (y-axis values) are randomly dispersed around the y=0 line, without exhibiting any discernible shape or pattern. This suggests that the assumption of linearity is adequately satisfied.

Homoscedasticity test: Scale-Location Plot

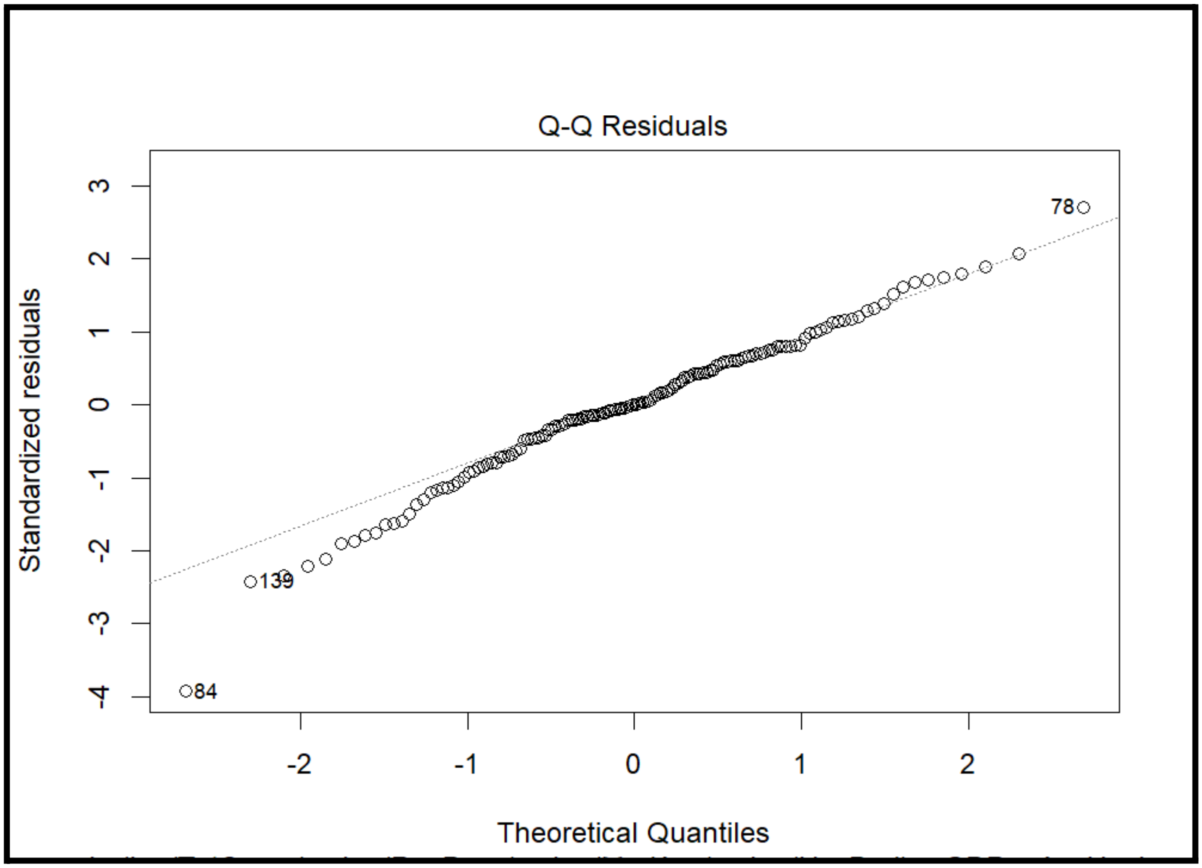


Interpretation

A scale-location plot displays the fitted values of a regression model along the x-axis and the square root of the standardized residuals along the y-axis. It shows if residuals are spread equally along the ranges of predictors. This is how we can check the assumption of equal variance (homoscedasticity). The residuals roughly form a horizontal band around the 0 line suggesting that the variances of the residual terms are equal.

In the presented scale-location plot, the residual values are observed to be randomly dispersed around the constant line, indicative of y=0, without exhibiting any systematic pattern. Although there are some outliers present, the consistent spread of the residuals suggests a reasonable conformity to the homoscedasticity assumption of the linear regression model for the sample dataset.

Normality Test: Q-Q Residuals Plot

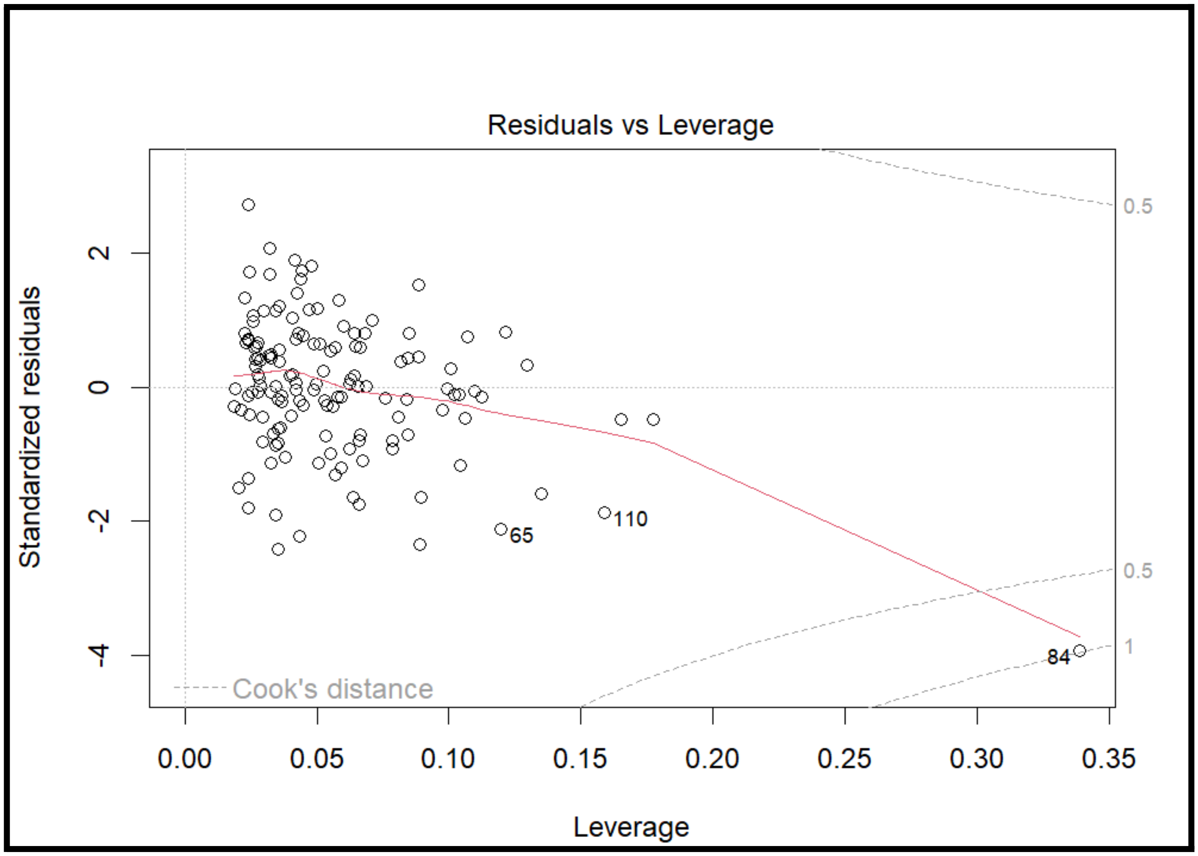


Interpretation

The QQ plot, which stands for Quantile-Quantile plot, is a graphical method used to determine whether a given dataset is likely to have originated from a specific theoretical distribution, such as the normal or exponential distribution. The x axis shows the theoretical quantiles from the standard normal distribution with mean 0 and standard deviation while the Y axis shows the quantile for the sample dataset.

In the presented plot, the sample quantiles closely align with the theoretical quantiles derived from a standard normal distribution. This observation suggests that the sample data exhibits a *reasonably* normal distribution.

Checking for Outliers and Reliability: Residual vs Leverage Plot



Interpretation

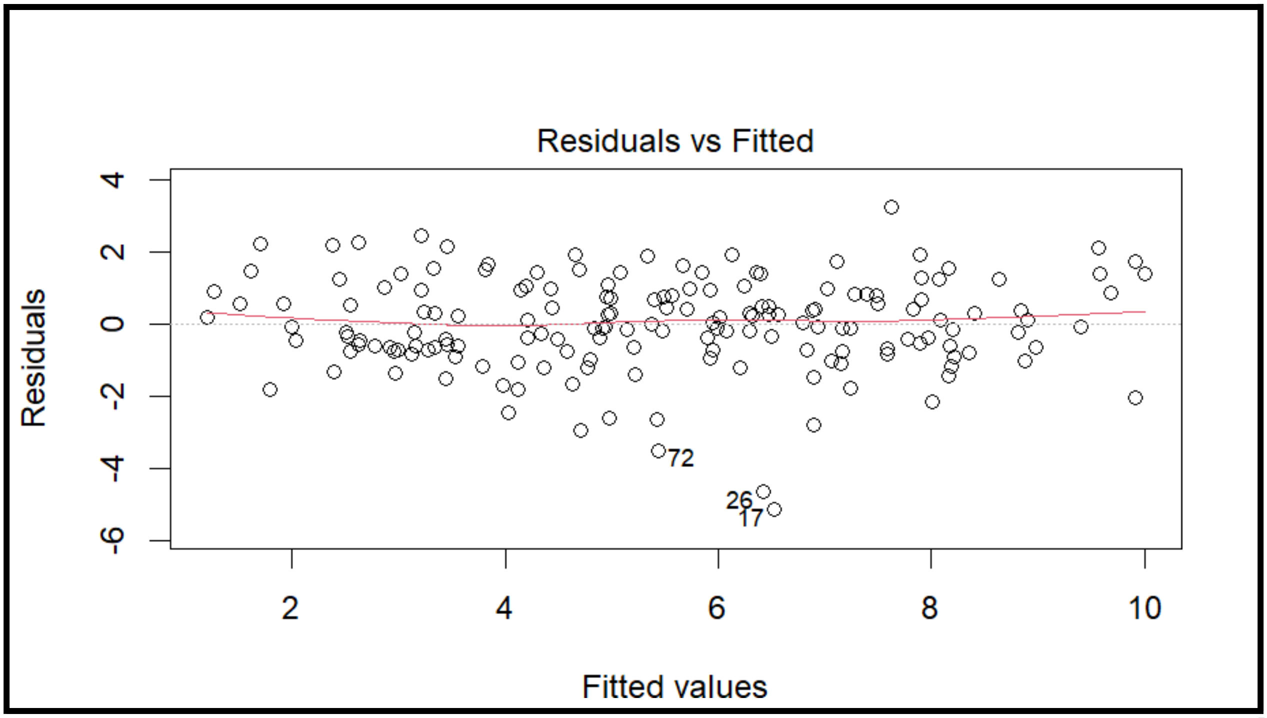
A residuals vs. leverage plot serves as a diagnostic tool designed to identify influential observations within a regression model. The x-axis represents the leverage of each data point, while the y-axis displays the standardized residual for each respective point. Observations with high leverage exert a significant influence on the coefficients of the regression model.

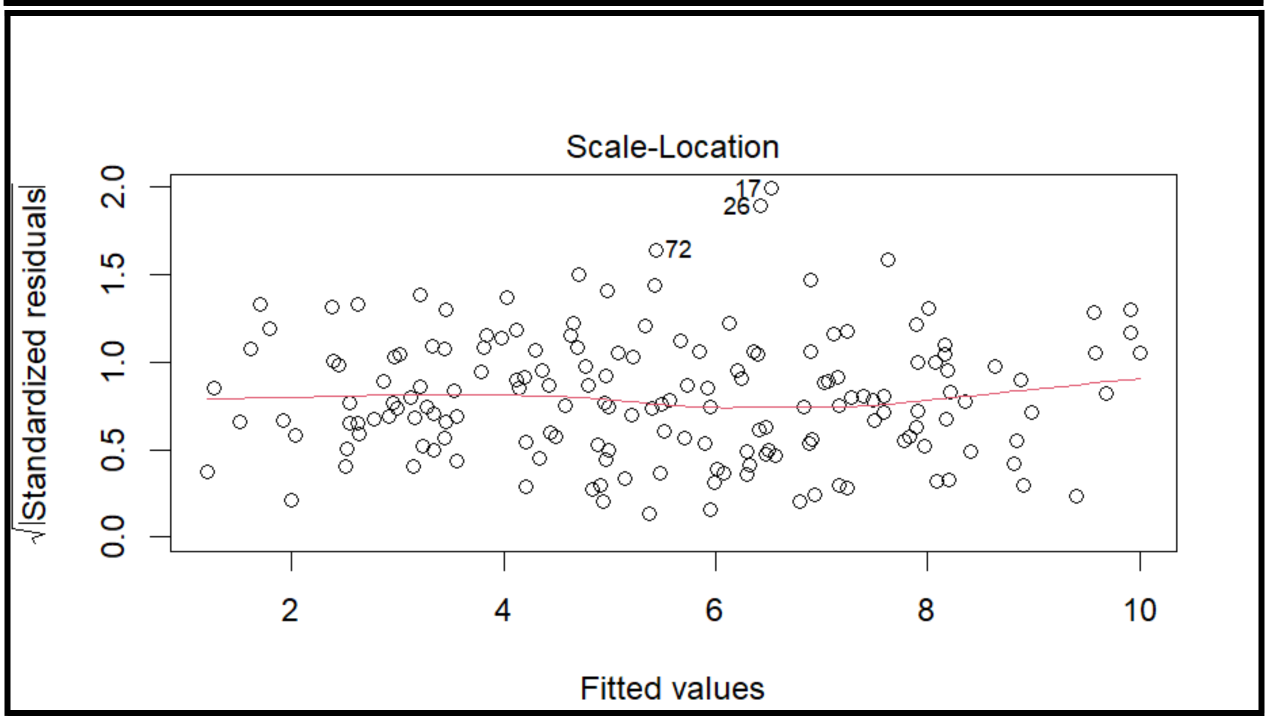
In the plot above, there is just one value, the 84th observation, that lies beyond the Cook's distance, thus being the most influential variable in the sample dataset. This implies that excluding this particular observation from our dataset and re-estimating the regression model would lead to substantial changes in the model's coefficients.

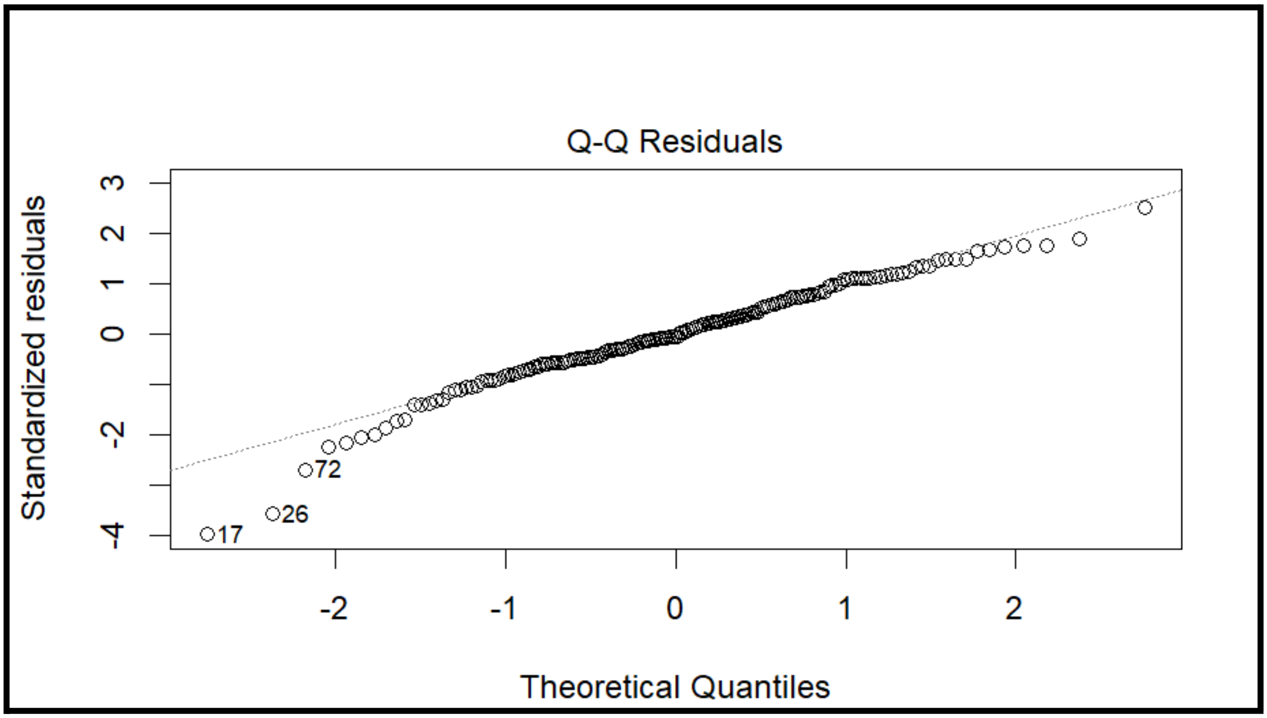
It is now identified that the 84th observation is an influential residual value, which has a significant impact on our model and its plots.

To remedy this observation, we omit the 84th observation from our dataset in order to see what changes it would bring to our model.

New Diagnostics Tests







Upon the removal of the 84th observation from our dataset, notable improvements were observed in the diagnostic measures of our regression model. Specifically, the linearity of the model improved substantially, as evidenced by the red line now closely aligning with the y=0 line, and the residual terms exhibiting a random scatter around this line.

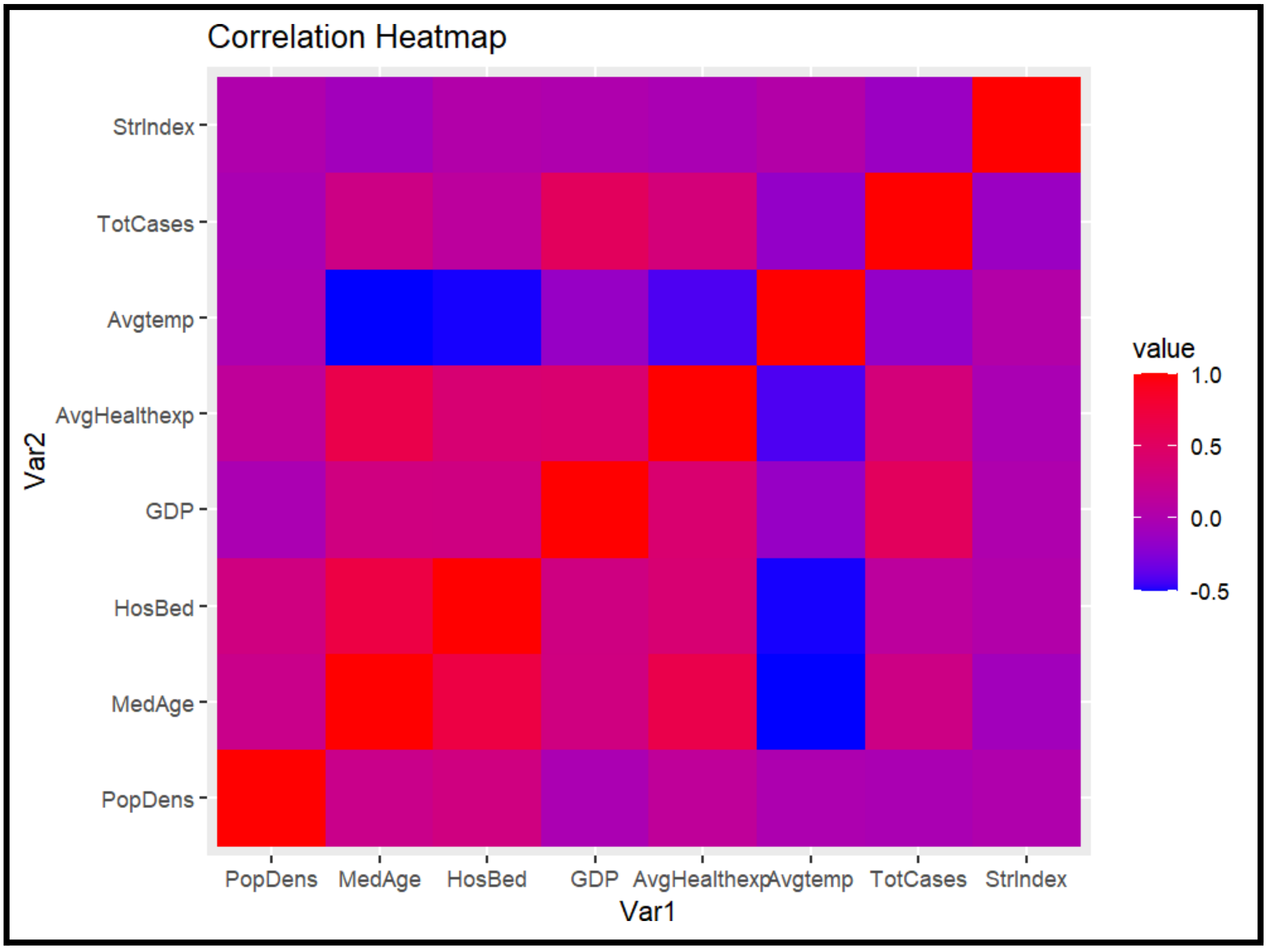
Furthermore, the homoscedasticity of the model also improved, with the residual terms now displaying an equal and consistent scatter, forming a horizontal band along the horizontal line.

In addition, the Q-Q plot revealed enhanced normality of the distribution following the removal of the influential observation, thereby further refining the effectiveness of the log-log transformation applied to the model.

MULTICOLLINEARITY TESTS

TEST 1

Correlation Matrix: A correlation matrix is a square table that summarizes the correlation coefficients between all possible pairs of variables in a dataset.
Correlation matrix visualizations (heatmaps or scatterplot matrix) depict relationships between variables using color intensity.

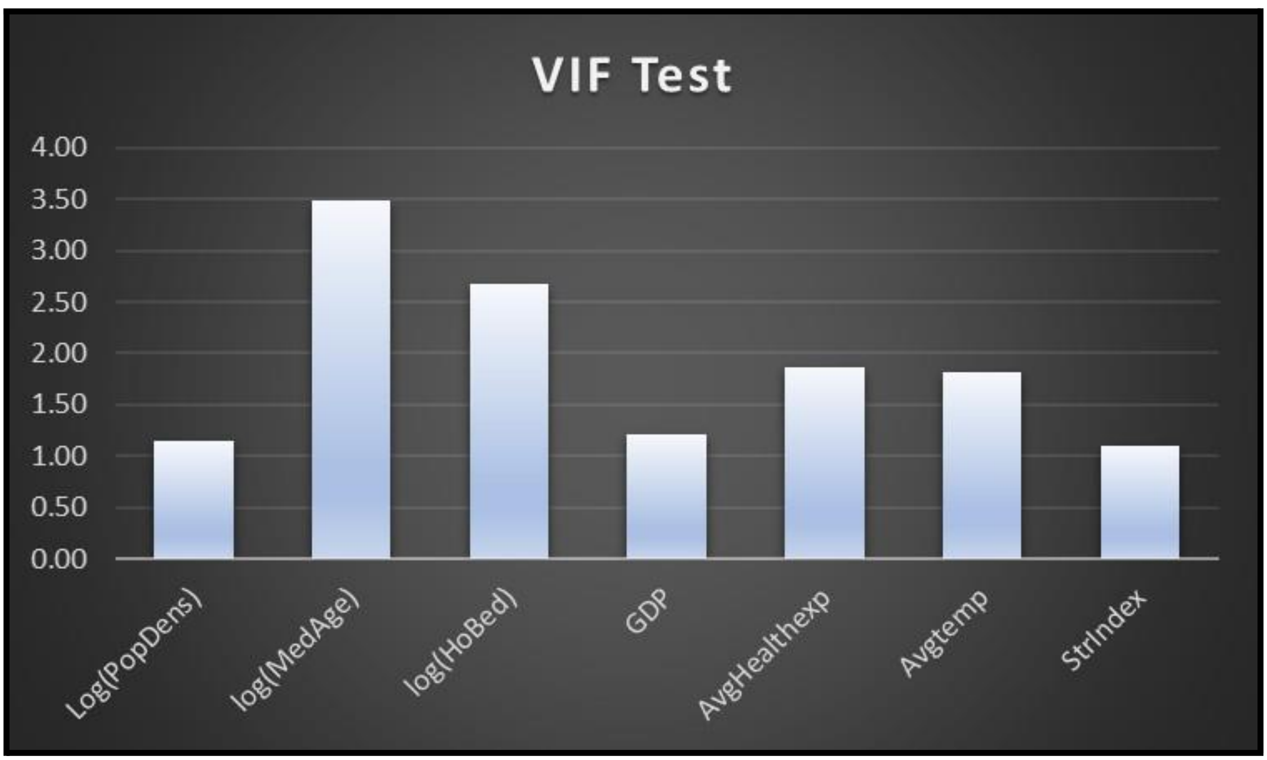


Column1	PopDens	MedAge	HosBed	GDP	AvgHealthexp	Avgtemp	TotCases	StrIndex
PopDens	1.0000	0.238	0.304	-0.0217	0.139	-0.0058	0.028	0.0151
MedAge	0.2381	1.000	0.699	0.3024	0.650	-0.5078	0.275	-0.0848
HosBed	0.3039	0.699	1.000	0.2931	0.393	-0.4994	- 0.114	0.0323
GDP	-0.0217	0.302	0.293	1.0000	0.417	-0.1493	0.538	0.0069
AvgHealth	0.1393	0.650	0.393	0.4173	1.000	-0.4288	-0.348	-0.0286
Avgtemp	-0.0058	-0.508	-0.499	-0.1493	-0.429	1	-0.166	0.0441
TotCases	0.0279	0.275	-0.114	0.5384	-0.348	-0.1658	1.000	-0.1254
StrIndex	0.0151	-0.085	0.032	0.0069	-0.029	0.0441	-0.125	1

In the provided matrix and table, it is evident that no variables exhibit a perfect correlation with each other, aside from their correlation with themselves. Therefore, we can deduce that there is no perfect multicollinearity present among the variables.

TEST 2

VIF TEST




```
log(PopDens)  log(MedAge)  log(HosBed)        GDP   AvgHealthexp
    1.159016     3.494228     2.677644   1.206064       1.864578
     Avgtemp     StrIndex
    1.820341     1.093785
```

VIF testing, an acronym for Variance Inflation Factor testing, is a statistical method utilized to quantify the extent to which multicollinearity, or a linear relationship between independent variables, inflates the variance of regression estimates, potentially compromising their reliability.
A widely accepted rule of thumb in the context of VIF is that a value exceeding 5 indicates potential multicollinearity, which may undermine the reliability of the regression coefficients.
In the given scenario, as all VIF values fall between 1 and 5, it can be inferred that while there is no perfect linear relationship between the variables, some degree of linear relation does exist among them.

TESTING FOR AUTOCORRELATION

H_0: The residuals are not autocorrelated.
Ha: The residuals are autocorrelated.

The Durbin-Watson test:
The DW test always produces a test number range from 0 to 4. Values closer to 0 indicate a greater degree of positive correlation, values closer to 4 indicate a greater degree of negative autocorrelation, while values closer to the middle suggest less autocorrelation.

```
Durbin-Watson test

data:  Model4
DW = 2.1021, p-value = 0.7376
```

Interpretation:

Our DW value comes out to 2.1021, quite close to the value of 2, hinting at a small value for negative autocorrelation.

The P-value is greater than our acceptable 0.05 value, and so we fail to reject the Null Hypothesis. The residuals are not autocorrelated.

ANALYSIS

Through rigorous statistical modeling and diagnostic tests, we have tried to identify the most appropriate regression model for our given sample dataset. Our analysis has revealed that Model4, which is the log-log transformation of the model, is the optimal fit for the data in question.

One of the primary concerns when analyzing sample datasets, as opposed to full population datasets, is the presence of high heteroscedasticity. Heteroscedasticity refers to the unequal variance of the dependent variable across the levels of the independent variable(s). This non-constant variance can severely affect the reliability and validity of the regression model. To mitigate this issue, the log-log transformation of the data has been applied. This transformation serves to reduce the heteroskedasticity of the data, thereby producing a more consistent variance across the levels of the independent variable(s). Additionally, the log-log transformation aids in reducing data skewness, transforming the data into a distribution that approximates normality, which is a fundamental assumption for linear regression.

Another advantage of employing the log-log model is its ability to address the problem of differing scale units between the independent and dependent variables. The log-log transformation allows for the interpretation of the relative change in the dependent variable ('y') in response to a proportional change in the independent variable ('x'), thereby standardizing the scale and facilitating more meaningful and interpretable results.

After running the diagnostics tests, we were able to identify the influential outlier in our data, and so were able to remove that observation altogether, this resulted in a more homoscedastic and linear model.

Furthermore, to ensure the robustness and reliability of our chosen model, we conducted an examination of the assumption of multicollinearity among the independent variables. Multicollinearity occurs when two or more independent variables in a regression model are highly correlated, making it difficult to determine the individual effect of each variable on the dependent variable. To assess multicollinearity, both Variance Inflation Factor (VIF) and correlation matrix tests were performed. Our analysis indicated that there is no evidence of perfect multicollinearity among the variables, affirming the reliability and validity of our log-log regression model.

In summary, based on our comprehensive analysis, we conclude that the log-log model (Model4) is the most suitable and reliable regression model for our sample dataset. This model not only addresses the issue of heteroscedasticity and data skewness but also standardizes the scale of the variables, fulfilling the fundamental assumptions of multiple linear regression.

LIMITATIONS

Presence of Outliers: In our graphs above, we saw that a few observations can be seen as outliers in our data, in the 4th Graph, i.e. the Residuals vs Leverage, we even saw that the 84th observation, is not only an outlier, but an influential one, this may disturb our analysis and models, and the way to avoid that in the easiest and most direct way is to omit the 84th observation.

Skewness in our data: As can be seen, our data is slightly skewed (scale-location plot), this could potentially be attributed to the fact that our data is a sample and not population, this could also lead to our intercepts not being exact. Skewed data can also create points with high leverage, meaning they exert a strong pull on the regression line. This can distort the model fit and lead to misleading results. The way to avoid that would be to take log on both sides of the model, which ultimately helped us improve our model and deal with such problems.

Heteroscedasticity in the data: Our data exhibits heteroscedasticity due to the sampling method, resulting in a non-constant variance of the residual terms and potentially unreliable model interpretations. However, the implementation of a log-log transformation successfully mitigated this issue by creating a more homoscedastic distribution with a consistent variance of the residual terms, enhancing the robustness and reliability of our regression model.

Scope for further research and modeling: As researchers further their understanding, in the future more variables can be attributed to our model, and can potentially improve our model further.

CONCLUSION

This research delves into the intricate world of COVID-19 transmission, employing regression analysis as a powerful tool to shed light on its multifaceted dynamics. Our project meticulously explored the various factors influencing viral spread and utilized a structured approach to identify the optimal model for predicting the total number of cases within the provided dataset.
We embarked on a journey of model comparison, meticulously evaluating different functional forms. This comprehensive analysis encompassed linear models, the workhorses of regression, alongside log-log and semi-log models, each offering unique perspectives on the underlying relationships between variables. We further expanded our investigation by considering various combinations of regressors, independent variables believed to influence COVID-19 transmission. Through this rigorous process, we aimed to pinpoint the most effective and statistically significant model tailored specifically to our data.

By harnessing data from a diverse set of 175 countries across the entirety of 2020, we embarked on a journey of refinement, iteratively honing our model selection process. This meticulous approach culminated in the identification of a log-log model (Model 4) as the most effective fit. This champion model surpassed its competitors, boasting an impressive adjusted R-squared value of 78.24%. This exceptional statistic not only underscores the model's remarkable predictive power but also signifies its ability to explain a substantial portion of the variance in total case numbers observed across the dataset.

The significance of this research extends beyond the realm of model selection. By meticulously unraveling the factors that significantly influence the spread of COVID-19, our findings offer a treasure trove of actionable insights for public health officials and policymakers. Armed with this knowledge, they are empowered to devise targeted interventions that can effectively mitigate the pandemic's impact. These interventions, informed by the robust framework established through our analysis, have the potential to protect vulnerable populations worldwide.

In conclusion, this research not only unveils a powerful predictive model for understanding COVID-19 spread but also delivers valuable insights into the key drivers of this global health crisis. By equipping policymakers and public health officials with a deeper comprehension of these dynamics, we move a step closer to effectively combating the pandemic and safeguarding the well-being of populations across the globe.

REFERENCES

1. Singh, P. K., Kiran, R., Bhatt, R. K., Tabash, M. I., Pandey, A. K., & Chouhan, A. (2021). COVID-19 pandemic and transmission factors: An empirical investigation of different countries. Journal of public affairs, 21(4), e2648. https://doi.org/10.1002/pa.2648
2. Ghosh, S., Roy, S.S. Global-scale modeling of early factors and country-specific trajectories of COVID-19 incidence: a cross-sectional study of the first 6 months of the pandemic. BMC Public Health 22, 1919 (2022). https://doi.org/10.1186/s12889-022-14336-w
3. James A, Dalal J, Kousi T, et alAn in-depth statistical analysis of the COVID-19 pandemic's initial spread in the WHO African regionBMJ Global Health 2022;7:e007295.
4. Ozyilmaz, A., Bayraktar, Y., Toprak, M., Isik, E., Guloglu, T., Aydin, S., Olgun, M. F., & Younis, M. (2022). Socio-Economic, Demographic and Health Determinants of the COVID-19 Outbreak. Healthcare (Basel, Switzerland), 10(4), 748. https://doi.org/10.3390/healthcare10040748
5. DasguptaSusmitaWheelerDavid, R. (n.d.). Modeling and Predicting the spread of COVID-19: Comparative results for the United States, the Philippines, and South Africa. World Bank. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/533861601575025228/modeling-and-predicting-the-spread-of-covid-19-comparative-results-for-the-united-states-the-philippines-and-south-africa

STUDENTS NAME :- MANISH KUMAR

BHAVIK JINDAL

ROLL NO :- 12745

12730

PGDAV COLLEGE ,UNIVERSITY OF DELHI

The Impact of War on a Country's GDP

Introduction

War has long been a significant disruptor of national economies. It can leave lasting scars on the economic landscape, impeding growth and development for years or even decades. Gross Domestic Product (GDP), a key indicator of a country's economic health, often reflects the immediate and long-term impacts of war. This article explores how war affects GDP through various channels including the destruction of infrastructure, loss of human capital, increased military spending, economic instability, and changes in investor confidence. Furthermore, it examines case studies to illustrate these impacts in real-world scenarios.

Destruction of Infrastructure

Infrastructure is the backbone of economic activities in any country. It includes transportation networks (roads, bridges, railways), utilities (water, electricity), and communication systems, all of which are essential for the smooth functioning of an economy. During war, these structures are often targets of direct attacks or collateral damage. The destruction of infrastructure has several adverse effects on GDP:

1. **Production Capabilities**: Factories and industrial sites are often damaged or destroyed, leading to a significant reduction in the production of goods. This, in turn, reduces the overall output of the economy.

2. **Supply Chain Disruptions**: War can sever supply chains, making it difficult for raw materials to reach factories and for finished products to reach markets. This disruption further hampers economic activities.

3. **Increased Costs**: The destruction of infrastructure means that resources must be diverted to rebuild these assets. This reallocation of resources comes at the expense of other productive investments, leading to lower GDP growth.

For example, the Syrian Civil War has caused extensive damage to the country's infrastructure. According to the World Bank, the cumulative loss in GDP from 2011 to 2016 due to the war was about $226 billion, which is roughly four times the country's GDP in 2010.

Loss of Human Capital

Human capital, the economic value of a worker's experience and skills, is another critical component of GDP. War leads to significant loss of human capital through casualties, displacement, and disruption of education and training systems:

1. **Casualties**: The loss of life reduces the workforce, directly impacting productivity. Skilled labor, which is often hard to replace, is particularly affected.

2. **Displacement**: War forces people to flee their homes, creating large numbers of refugees and internally displaced persons. These individuals often lose their jobs and are unable to contribute to the economy.

3. **Education Disruption**: Schools and universities may be destroyed or closed during conflict, interrupting the education of future generations. This disruption leads to a less skilled workforce, reducing the potential for economic growth in the long term.

For instance, in the Democratic Republic of Congo, ongoing conflict has displaced millions and led to significant educational disruptions. The loss of human capital has been a major factor in the country's persistent low GDP growth.

Increased Military Spending

Military expenditure during wartime increases significantly as countries allocate more resources to defense. While this spending can create short-term economic activity in defense-related industries, it has several negative implications for GDP:

1. **Opportunity Cost**: Resources spent on military activities are diverted from other productive uses such as infrastructure development, healthcare, and education. This diversion reduces the overall productive capacity of the economy.

2. **Debt**: War often leads to increased borrowing to finance military operations. The resultant debt burden can constrain future government spending on development projects, further hampering economic growth.

3. **Inflation**: Increased military spending can lead to inflation, as governments may resort to printing more money to finance the war. Inflation erodes purchasing power and can destabilize the economy.

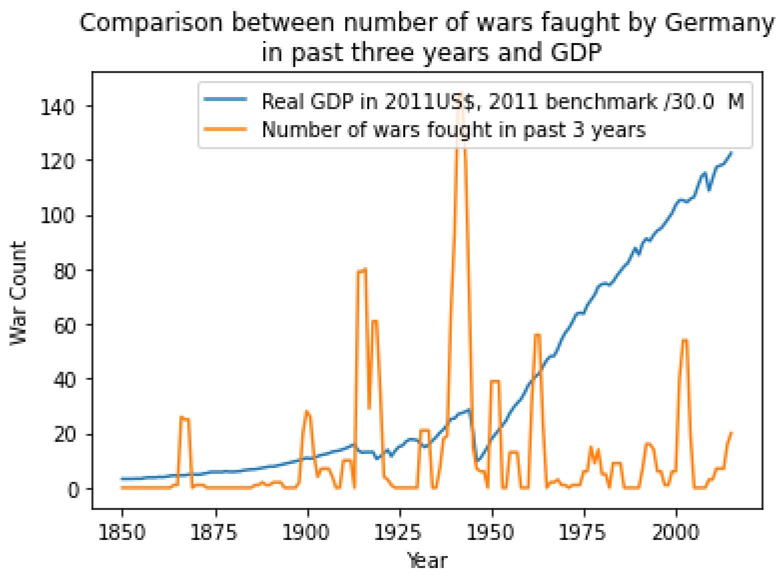
The United States during World War II saw a massive increase in military spending. While this spending boosted the economy in the short term, it led to significant national debt and necessitated economic restructuring in the post-war period.

Let see a graph of germany had fought war over 3 years

	1938	1939	1940	1941	1942	1943	1944	1945
Allied powers								
USA	800	869	943	1094	1235	1399	1499	1474
UK	284	287	316	344	353	361	346	33
France	186	199	82	..	..	..	..	10
Italy	..	..	..	..	..	..	117	92
USSR	359	366	417	359	274	305	362	34
Allied total	1629	1721	1757	1798	1862	2064	2325	2342
Axis powers								
Germany	351	384	387	412	417	426	437	31
France	..	..	82	130	116	110	93	.
Austria	24	27	27	29	27	28	29	12
Italy	141	151	147	144	145	137	..	.
Japan	169	184	192	196	197	194	189	144
Axis total	686	747	835	911	903	895	748	46
Allies–to–Axis								
Overall	2.4	2.3	2.1	2.0	2.1	2.3	3.1	5.
USSR to Germany	1.0	1.0	1.1	0.9	0.7	0.7	0.8	1.

Sources. For 1938, see table 1-1. Other years are interpolated on index numbers as follows: UK, table 2-1 (col. 4); USA, table 3.1 (col. 4); Germany, table 4.1 (col. 1); Italy, table 5-1 (col. 3); Japan, table 6-1 (col. 1); USSR, table 7-7, part (A). Figures for the USSR for 1939 are interpolated on population within 1938 frontiers on the assumption that GDP per head remained unchanged compared with 1938 (for evidence on this score see Harrison (1994), 269; Maddison (1995), 200). For France and Austria see Maddison (1995), appendix B.

Figures in red correct a spreadsheet error in the published version that overstated figures for Soviet GDP.



Economic Instability and Investor Confidence

War creates an environment of uncertainty and instability, which has far-reaching effects on the economy:

1. **Investment Deterrence**: Both domestic and foreign investors are likely to be wary of investing in a country embroiled in conflict. The risks associated with war, including the potential for asset destruction and market volatility, make investment unattractive.

2. **Capital Flight**: Existing investors may withdraw their investments and move capital to safer havens, leading to a reduction in the country's financial resources and further economic decline.

3. **Currency Depreciation**: War can lead to a loss of confidence in the country's currency, resulting in depreciation. This depreciation makes imports more expensive, contributing to inflation and further economic instability.

Iraq, for example, has struggled with attracting foreign investment due to ongoing instability and security concerns. Despite its rich natural resources, the continuous conflict has made it difficult to restore investor confidence and achieve sustained economic growth.

Long-term Economic Consequences

The long-term economic consequences of war can be profound and enduring. Post-war reconstruction often requires substantial financial resources and can lead to a series of economic challenges:

1. **National Debt**: The costs of war and subsequent reconstruction efforts often lead to increased national debt. High debt levels can constrain government spending on essential services and development projects.

2. **Economic Restructuring**: Countries emerging from war often need to undergo significant economic restructuring. This process can be painful and slow, requiring major policy changes and economic reforms.

3. **Opportunity Costs**: The economic opportunities lost due to war are substantial. The potential economic output that could have been achieved in the absence of conflict represents a significant cost.

4. **Social and Economic Inequalities**: War can exacerbate social and economic inequalities, creating long-term challenges for economic stability and growth.

Case Studies

To illustrate the impact of war on GDP, it is helpful to examine specific case studies of countries that have experienced significant conflict.

1. **Germany Post-World War II**: The devastation of World War II left Germany's economy in ruins. However, with the aid of the Marshall Plan, Germany was able to rebuild its infrastructure and economy, leading to the Wirtschaftswunder (economic miracle) in the 1950s and 1960s. The country's GDP grew rapidly as it became one of the world's leading industrial powers.

2. **Afghanistan**: Decades of conflict in Afghanistan have left its economy fragile and dependent on foreign aid. The war has destroyed infrastructure, displaced millions, and disrupted education. Despite significant international assistance, the country struggles with low GDP growth and widespread poverty.

3. **Vietnam**: The Vietnam War caused extensive destruction and economic disruption. However, after the war, Vietnam implemented significant economic reforms (Đổi Mới) in the 1980s, transitioning from a centrally planned economy to a market-oriented one. These reforms have led to rapid GDP growth and significant improvements in living standards.

4. **Bosnia and Herzegovina**: The Bosnian War in the 1990s caused extensive damage to the country's economy and infrastructure. Post-war reconstruction and international aid have helped the country recover, but it still faces challenges related to political instability and economic inequality. The GDP has grown, but the pace of growth remains slower compared to other post-conflict countries.

5. **Sierra Leone**: The civil war in Sierra Leone during the 1990s resulted in severe economic decline and destruction of infrastructure. Since the end of the war, the country has made progress in rebuilding its economy, with GDP growth driven by mining and agriculture. However, the long-term effects of the conflict, such as poverty and unemployment, continue to pose challenges.

Conclusion

The impact of war on a country's GDP is profound and multifaceted. The immediate destruction of infrastructure, loss of human capital, increased military spending, and economic instability all contribute to a significant decline in GDP. The long-term consequences, including debt burdens, economic restructuring, and opportunity costs, further hinder economic recovery and growth. While some countries manage to rebuild and achieve robust economic growth post-conflict, others struggle with persistent challenges. Understanding these impacts is crucial for policymakers and international organizations involved in conflict resolution and post-war reconstruction. By addressing the economic repercussions of war, these entities can help pave the way for sustainable development and stability in post-conflict regions.

Reference

Alfaro, L., Chanda, A., Kalemli-Ozcan, S., Sayek, S. (2004), FDI and economic growth: The role of local financial markets. Journal of International Economics, 64(1), 89-112. Balasubramanyam, V.N., Salisu, M., Sapsford, D. (1996), Foreign direct investment and growth in EP and is countries. The Economic Journal, 106(434), 92-105. Balasubramanyam, V.N., Salisu, M., Sapsford, D. (1999), Foreign direct investment as an engine of growth. The Journal of International Trade and Economic Development, 8(1), 27-40. Basu, P., Chakraborty, C., Reagle, D. (2003), Liberalization, FDI, and growth in developing countries: A panel cointegration approach. Economic Inquiry, 41(3), 510-516. Basu, P., Guariglia, A. (2007), Foreign direct investment, inequality and growth. Journal of Macroeconomics, 29, 824-839. Bengoa, M., Sanchez-Robles, B. (2003), Foreign direct investment, economic freedom and growth: New evidence from Latin America. European Journal of Political Economy, 19, 529-545. Berthélemy, J.C., Démurger, S. (2000), Foreign direct investment and economic growth: Theory and application to China. Review of 330 Development Economics, 4(2), 140-155. Carkovic, M., Levine, R. (2002), Does Foreign Direct Investment Accelerate Economic Growth? Available from: https://www.piie. com/publications/chapters_preview/3810/08iie3810.pdf. Choe, J.I. (2003), Do foreign direct investment and gross domestic investment promote economic growth? Review of Development Economics, 7(1), 44-57. Chowdhury, A., Mavrotas, G. (2003), FDI and Growth: What Causes What? Paper presented at the WIDER Conference on Sharing Global Prosperity, WIDER. Helsinki: World Bank Group. De Mello, L.R. (1999), Foreign direct investment-led growth: Evidence from time series and panel data. Oxford Economic Paper, 51, 133-151. Durham, J.B. (2004), Absorptive capacity and the effects of foreign direct investment and equity foreign portfolio investment on economic growth. European Economic Review, 48, 285-306. Hansen, H., Rand, J. (2006), On the Causal Links Between FDI and Growth in Developing Countries. Copenhagen: University of Copenhagen and Development Economics Research Group. Hermes, N., Lensink, R. (2003), Foreign direct investment, financial development and economic growth. Journal of Development Studies, 40(1), 142-163. Johnson, A. (2006), The Effects of FDI Inflows on Host Country Economic Growth, No. 58, Working Paper Series in Economics and Institutions of Innovation. Tamil Nadu: Royal Institute of Technology, CESIS Centre of Excellence for Science and Innovation Studies. Rahaman, A., Chakraborty, S. (2015), Effects of foreign direct investment on GDP: Empirical evidence from developing country. Advances in Economics and Business, 3(12), 587-592. Rakhmatullayeva, D., Kuliyev, I., Beisenbaiyev, Z., Tabeyev, T. (2020), Assessment of the influence of FDI on the economic growth of the host country: Evidence from Kazakhstan. E3S Web of Conferences, 159, 06007

What is Regression?

Regression analysis is a statistical technique used to model and analyze the relationship between a dependent variable and one or more independent variables. Its purpose is to understand how changes in the independent variables are associated with changes in the dependent variable. Regression analysis aims to find the best-fitting mathematical model that represents this relationship. It's a versatile tool employed in various fields, from predicting sales trends in digital marketing to understanding the impact of advertising campaigns on consumer behavior. There are different types of regression models, such as linear regression, logistic regression, polynomial regression, and more, each suited to different types of data and relationships.

Applications of Regression Analysis in Real-Life Scenarios
Business and marketing:
- Predicting sales based on factors like advertising expenditure, product price, and consumer demographics.
- Analyzing customer behavior and purchase patterns to optimize marketing strategies.
- Forecasting demand for products or services based on historical data and market trends.

Finance and economics:
- Modeling stock prices or commodity prices based on economic indicators and market factors.
- Analyzing the impact of interest rates, inflation, and other variables on economic growth or household income.
- Estimating the effect of policy changes (e.g., tax reforms) on economic variables like employment or consumer spending.

Healthcare and medicine:

- Studying the relationship between risk factors (e.g., age, lifestyle, genetics) and the prevalence of diseases.
- Analyzing the effectiveness of medical treatments or drugs based on patient characteristics and other variables.
- Predicting the length of hospital stays or the likelihood of readmission based on patient data.

Environmental studies:
- Modeling the impact of factors like industrial emissions, deforestation, or climate change on air quality or biodiversity.
- Predicting the spread of wildfires or the occurrence of natural disasters based on weather conditions and environmental factors.

Engineering and manufacturing:
- Optimizing product design or manufacturing processes based on the relationship between input variables and output quality.
- Predicting the lifespan or failure rate of components or systems based on usage patterns and environmental conditions.

Social sciences:

- Analyzing the relationship between socioeconomic factors (e.g., income, education, family structure) and various outcomes like academic performance or crime rates.
- Studying the impact of policy interventions or social programs on variables like unemployment or poverty levels.
- These are just a few examples, and regression analysis has numerous other applications in various fields. By identifying and quantifying the relationships between variables, regression helps in making informed decisions, optimizing processes, forecasting outcomes, and gaining insights into complex real-world phenomena.

Introduction:

The project aims to examine the relationship between revenue deficits and economic development, specifically the percentage of the population living below the national poverty line, for Indian states and union territories in the year 2019-2020. Understanding this relationship is crucial for policymakers and economists in formulating policies to promote economic growth and alleviate poverty.

Data:

State/UT	2019-2020	
	Revenue Surplus (-)/ Deficit (+)	Percentage of population living below the national poverty line
1. Andhra Pradesh	26,440.50	9.2
2. Arunachal Pradesh	-2,669.80	34.67
3. Assam	1,322.20	31.98
4. Bihar	-698.9	33.74
5. Chhattisgarh	9,608.60	21.81
6. Goa	218.8	5.09
7. Gujarat	-1,944.80	16.63
8. Haryana	16,990.10	11.16
9. Himachal Pradesh	-7.5	8.06
10. Jharkhand	-1,961.00	36.96
11. Karnataka	-1,185.40	20.91
12. Kerala	14,495.30	7.05
13. Madhya Pradesh	2,800.90	31.65
14. Maharashtra	17,115.60	17.35
15. Manipur	-445.5	36.89
16. Meghalaya	151.6	11.87
17. Mizoram	-204.3	20.4
18. Nagaland	213.7	18.88
19. Odisha	-2,430.40	32.59
20. Punjab	14,284.90	8.26
21. Rajasthan	36,371.30	14.71
22. Sikkim	1,343.80	8.19
23. Tamil Nadu	35,908.80	11.28
24. Telangana	6,254.10	-
25. Tripura	2,375.30	14.05
26. Uttar Pradesh	-67,560.10	29.43
27. Uttarakhand	2,136.20	11.26
28. West Bengal	19,660.90	19.98
29. Jammu and Kashmir	354.1	10.35
30. NCT Delhi	-7,498.80	9.91
31. Puducherry	54.5	9.69

The updated data shows the revenue surplus/deficit and the percentage of the population living below the national poverty line for various Indian states and union territories in the year 2019-2020.

Dependent variable used in the analysis:

The dependent variable in this analysis is the percentage of the population living below the national poverty line for different Indian states and union territories. This variable serves as a proxy for measuring economic development and well-being within each region. A higher percentage indicates a larger proportion of the population struggling with poverty and lack of access to basic necessities. By examining the relationship between revenue deficits (the independent variable) and this poverty measure, the analysis aims to understand how fiscal policies and budgetary constraints may influence the economic development and living standards of the states' and union territories' populations.

Methodology:

The scatter plot shows the relationship between revenue deficit and poverty percentage. The x-axis represents the revenue deficit, ranging from around -60,000 to 40,000, while the y-axis represents the poverty percentage, ranging from around 5% to 35%.

There is a cluster of data points with high revenue deficits (around 0 to 40,000) and high poverty percentages (around 25% to 35%), indicating a potential association between large revenue deficits and higher poverty rates.
There are also some data points with relatively low revenue deficits (around -40,000 to -20,000) but high poverty percentages (around 20% to 30%), suggesting that factors other than revenue deficit may contribute to poverty in some cases.
There are a few outliers with very high revenue deficits (around 0 to 40,000) but relatively low poverty percentages (around 5% to 15%), which could be exceptions or anomalies in the data.

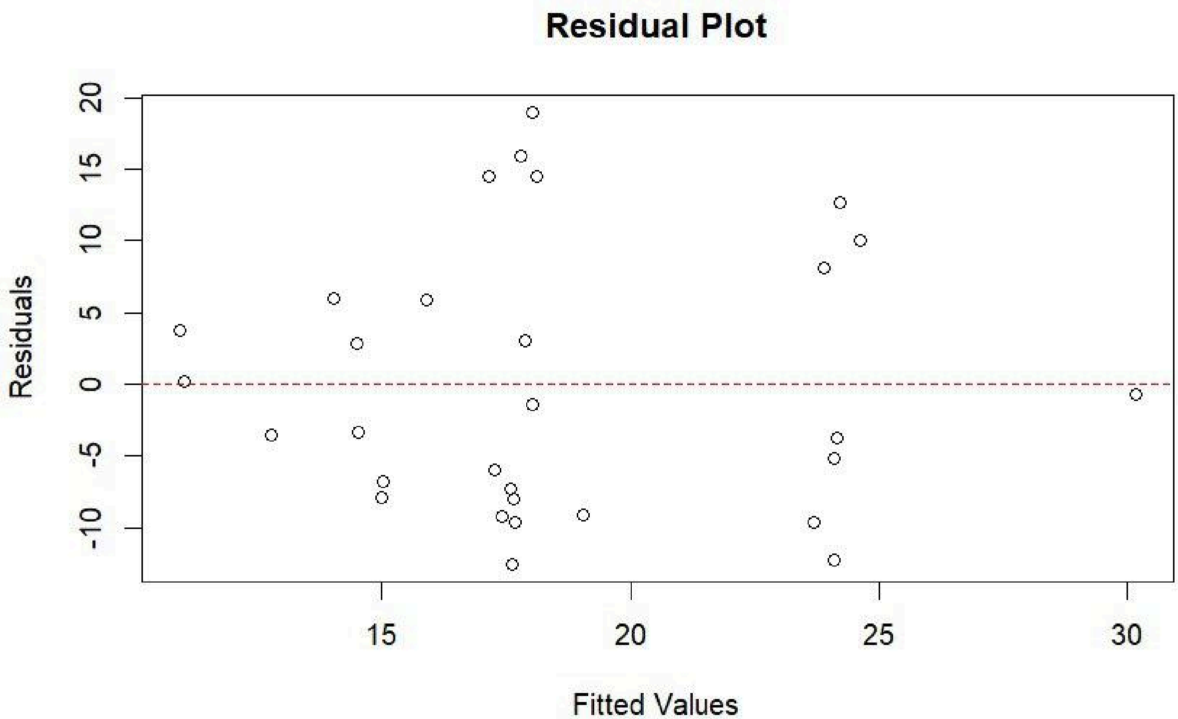
Key observations:

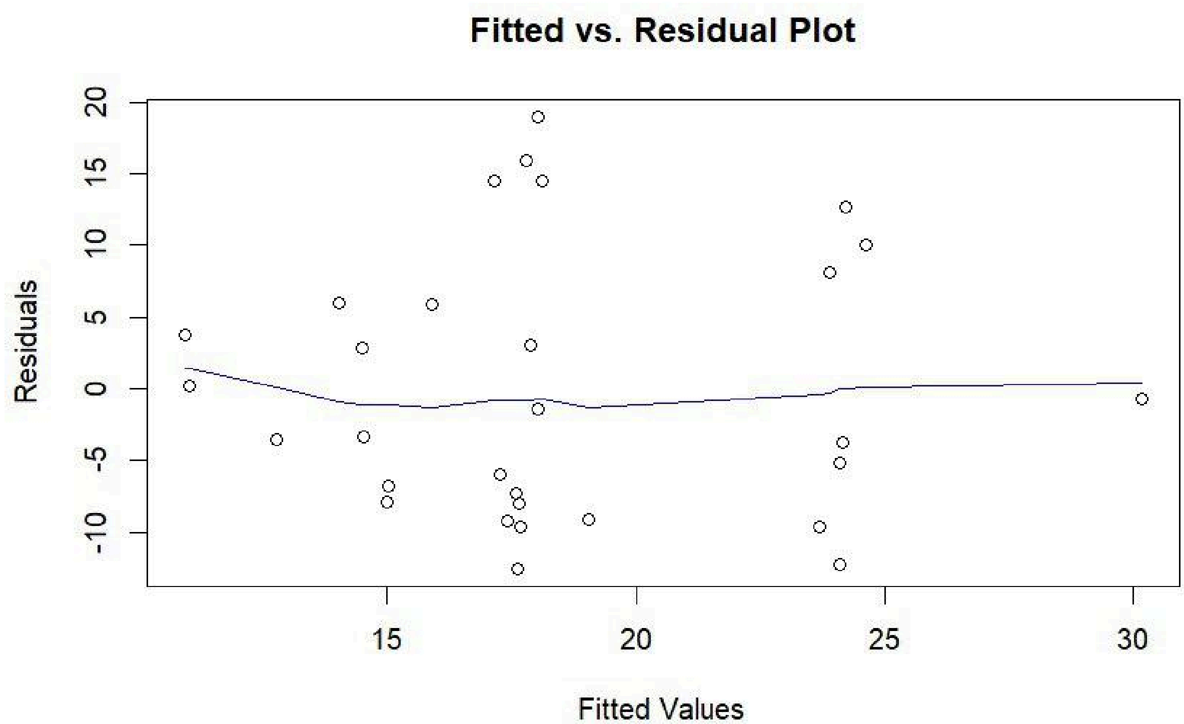
The linear regression model with Revenue_Deficit as the sole predictor variable for Poverty_Percentage has limitations, as indicated by the following observations from the model summary:

Coefficients:

- The intercept (constant) estimate is 17.67029823, which represents the expected value of Poverty_Percentage when Revenue_Deficit is zero.
- The coefficient estimate for Revenue_Deficit is -0.0001848, indicating a very small negative relationship between Revenue_Deficit and Poverty_Percentage, contrary to the expected positive relationship observed in the scatter plot.
- The p-value for Revenue_Deficit (0.0787) is greater than the typical significance level of 0.05, suggesting that the relationship is not statistically significant.

Residuals:


Residual Plot
Residuals
Fitted Values
20
15
10
5
0
-5
-10
15
20
25
30



- The residuals (differences between observed and predicted values) range from -12.54 to 18.927, with a median of 5.936.

Model Fit:

```
Call:
lm(formula = Poverty_Percentage ~ Revenue_Deficit + Is_North_East,
    data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-12.540  -7.770  -2.385   5.936  18.927

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)     17.6702833  2.0495070   8.622 3.09e-09 ***
Revenue_Deficit -0.0001848  0.0001011  -1.828   0.0787 .
Is_North_East    6.4550490  4.1423444   1.558   0.1308
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 9.528 on 27 degrees of freedom
Multiple R-squared:  0.1949,    Adjusted R-squared:  0.1353
F-statistic: 3.269 on 2 and 27 DF,  p-value: 0.05353
```

- The adjusted R-squared value of 0.1353 suggests a relatively poor model fit, implying that Revenue_Deficit alone may not be a strong predictor of Poverty_Percentage in this data set.
- The F-statistic (3.269) and its associated p-value (0.05353) indicate that the overall model is marginally statistically significant at the 0.05 level.

The residual plots (Images 1 and 2) were used to assess the assumptions of linearity and homoscedasticity (constant variance of residuals) in the linear regression model. Image 1 suggests potential non-linearity or heteroscedasticity, as the residuals appear to increase with higher fitted values. Image 2 shows the residuals scattered around the horizontal zero line, indicating no significant violation of the homoscedasticity assumption.

Dummy variable:

To create a dummy variable for the northeastern states using the provided data, I will add a new column named "Is_North_East" and assign a value of 1 for the northeastern states and 0 for the rest of the states/union territories.

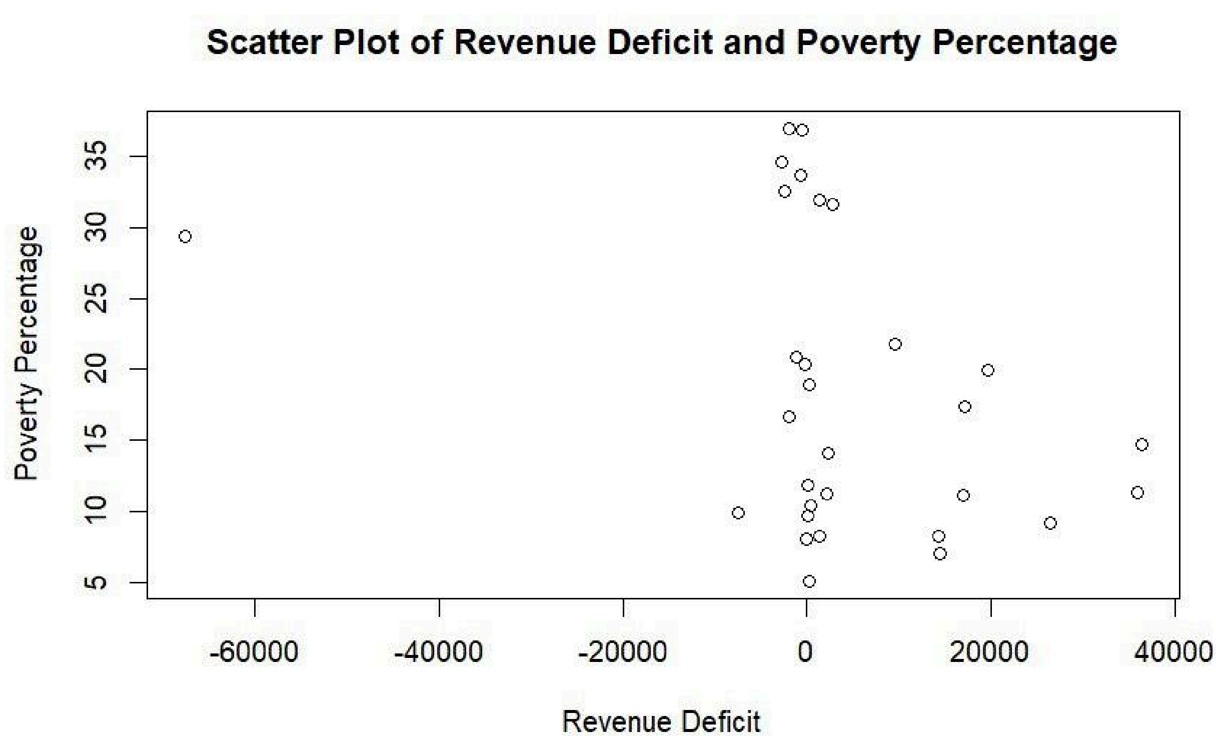
State/UT,2019-2020,,Is_North_East
,Revenue Surplus (-)/ Deficit (+),Percentage of population living below the national poverty line ,
,,,
1. Andhra Pradesh,"26,440.50", 9.2 ,0
2. Arunachal Pradesh,"-2,669.80", 34.67 ,1
3. Assam,"1,322.20", 31.98, 1
4. Bihar,-698.9, 33.74, 0
5. Chhattisgarh,"9,608.60", 21.81 ,0
6. Goa,218.8, 5.09, 0
7. Gujarat,"-1,944.80", 16.63, 0
8. Haryana,"16,990.10", 11.16, 0
9. Himachal Pradesh,-7.5, 8.06, 0
10. Jharkhand,"-1,961.00", 36.96, 0
11. Karnataka,"-1,185.40", 20.91, 0
12. Kerala,"14,495.30", 7.05, 0
13. Madhya Pradesh,"2,800.90" ,31.65 ,0
14. Maharashtra,"17,115.60", 17.35, 0
15. Manipur,-445.5, 36.89, 1
16. Meghalaya,151.6, 11.87, 1
17. Mizoram,-204.3, 20.4, 1
18. Nagaland,213.7, 18.88, 1
19. Odisha,"-2,430.40", 32.59, 0
20. Punjab,"14,284.90", 8.26, 0
21. Rajasthan,"36,371.30", 14.71, 0
22. Sikkim,"1,343.80", 8.19, 1

23. Tamil Nadu,"35,908.80" ,11.28, 0
24. Telangana,"6,254.10", -, 0
25. Tripura,"2,375.30", 14.05, 1
26. Uttar Pradesh,"-67,560.10", 29.43, 0
27. Uttarakhand,"2,136.20", 11.26, 0
28. West Bengal,"19,660.90", 19.98, 0
29. Jammu and Kashmir,354.1, 10.35, 0
30. NCT Delhi,"-7,498.80", 9.91, 0
31. Puducherry,54.5, 9.69, 0

The northeastern states (Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, and Tripura) have been assigned a value of 1 in the "Is_North_East" column, while the rest of the states/union territories have been assigned a value of 0.

This dummy variable can be used in the regression analysis to capture any potential differences between the northeastern states and the other states/union territories in the relationship between revenue deficit and poverty percentage.

Results:



The scatter plot reveals a generally positive relationship between revenue deficit and the percentage of the population living in poverty for the Indian states and union territories in the year 2019-2020. States and union territories with higher revenue deficits tend to have a larger proportion of their population living below the poverty line.

However, the linear regression model with Revenue_Deficit as the sole predictor variable may not adequately capture this relationship due to the lack of statistical significance, potential non-linearity, and the relatively poor model fit. Additional explanatory variables or a different model specification might be needed to better explain the variation in poverty percentage across Indian states and union territories.

Conclusion:

In conclusion, while the scatter plot indicated a positive relationship between revenue deficit and poverty percentage for Indian states and union territories in the year 2019-2020, the linear regression model with Revenue_Deficit as the sole predictor variable had several limitations in accurately capturing this relationship. The lack of statistical significance, potential non-linearity, and relatively poor model fit suggest that additional explanatory variables or a different model specification might be necessary to better explain the variation in poverty percentage across these regions.

Incorporating regional differences, such as the "Is_North_East" dummy variable created to distinguish northeastern states, could provide additional insights into the factors influencing poverty levels. However, it is crucial to recognize that poverty is a complex, multidimensional issue influenced by various economic, social, political, and geographical factors. Revenue deficit alone may not fully capture the intricacies and nuances associated with poverty levels across diverse states and union territories.

To gain a more comprehensive understanding of the determinants of poverty in Indian states and union territories, further research and analysis are warranted. This could involve incorporating additional relevant variables, considering potential non-linear relationships, and exploring alternative modeling approaches. Such efforts could yield valuable insights and inform policymakers in developing targeted strategies and interventions to effectively address poverty and promote inclusive economic development across the country.

It is imperative to approach the challenge of poverty alleviation with a holistic perspective, recognizing the complex interplay of factors that contribute to its persistence. By combining rigorous data analysis with a deep understanding of local contexts and conditions, policymakers and researchers can work towards designing evidence-based policies and programs that effectively tackle poverty and improve the overall well-being of the population.

GROUP(11)

SANTOSH(12738)

SHIVAM SHAURAV(12703)

TUSHAR DAGAR(12701)

Source:

1. Ministry of Finance, Government of India:
Website: https://www.indiabudget.gov.in/

2. Reserve Bank of India (RBI):
Website: https://www.rbi.org.in/

3. Ministry of Defence, Government of India:
Website: https://www.mod.gov.in/

4. Stockholm International Peace Research Institute (SIPRI): Website:
https://www.sipri.org/databases

SOFTWARE USED- R-STUDIO, CLAUDE AI, GOOGLE SHEETS.

PGDAV COLLEGE(M)

Introductory Econometrics by **Rimpy Kaushal**

Research Paper on:

Impact of Monetary Policy on Economic Variable

By:

Princess Verma, Anant Singh, Roshani Sharma

Abstract

In this research paper, we are writting about the monetary policy and economic variables and it’s effects that these economic variables has on the monetary policy.what are the implementation of monetary policy , challenges faced while making effective monetary policy and can one overcome it.

1. Introduction

1.1. Overview of Monetary Policy

Monetary policy refers to the policy that is adopted by a nation to control money supply and influence financial conditions so that there will be increase in employment and there will be price stability. In India monetary policy is regulated by Reserve Bank of India(RBI) which is also a central bank. Monetary policy is often used by developing countries so that there will be control on inflation and other macroeconomic variables. It is implemented by using various tools such as change in interest rate, exchange rate, changing the amount of cash circulating in economy, purchasing and selling of government securities etc. The primary target of the monetary policy is to control inflation or unemployment and maintenance of currency.

1.2. Importance of Economic Variable

Economic variables are nothing but an important factor in order to explain the production function and generally are affected by the ratio of given output to income. These variables are just indicators that help us to show the current trends in the economy and it is used to understand the forces of economic growth like if price increases then how inflation occurs etc.There are two types of variables like microeconomics and macroeconomics variable.

Microeconomic variables describe individual economic units: a family, a person or a company.It studies the behavior of individual economics . These variables revolves around an individual not a country as a whole.

Macroeconomic variables are associated withaggregate economics like country, region, population of a country, companies in a country. It studies the behavior of economic aggregates. These variables revolves around the country as a whole.

Example of variables are : Gross domestic product(GDP), Inflation, Exchange rate, unemployment etc.

2. Theortical Framework

2.1. Monetary Policy transmission mechanism

The monetary policy transmission happens when the increase in price of asset and general economic conditions are affected through the monetary policy implementation. The traditional monetary transmission occurs through interest rate channels, which affect interest rate, levels of physical involvement , cost of borrowing and aggregate demand.Moreover, the friction in credit view can also affect aggredate demand. The monetary transmission mechanism can be defined as the link between monetary policy and aggregate demand.

Since the interest rate channel has been categorized as traditional therefore, monetary policy(mp) affects real interest rate rather than nominal interest rate, which influence investments, spending on new housing, aggregate demand and consumer spending. If there is an ease in monetary policy in the traditional view then it leads to decrease in real interest rate, which therefore lowers the cost of borrowing resulting in more investment spending which involve an overall increase in aggregate demand.

The man transmission channel of monetary policy includes change in interest rate, which have several impacts. It has impact on money-markets and inflation. This transmission mechanism has long time lags which makes it difficult to predict the correct effect of monetary policy actions on price level and the economy.

2.2 Relationship between Monetary policy and Economic Variable

Monetary policy and economic variables are closely related as impact that monetary policy has on economy is very significant. It aims to influence employment, economic growth and financial conditions by adjusting the money supply and other instruments of monetary policy in the economy. It is a complex mechanism that involves several channels through which changes in monetary policy affect economy.The first stage of transmission focuses on how monetary policy influence interest rate in the economy. For example, the primary tool for monetary policy have a strong influence on the interest rate , such as deposits and lending rate for household and businesses. The RBI’s other tools, such as price and quantity targets for government bonds and the provision of low-cost fixed term funding to financial institutions, primarily affect longer-term interest rate in the economy.

The second stage for transmission is how these policy influence inflation and economic activity. For example, lower interest rate for business and households can increase aggregate demand because there will be increase in spending, which can led to upward shift in price and there will be temporary imbalance between demand and supply and in the higher interest rate the situation will be vice versa. Monetary policy can also influence inflation expectations, which are crucial for transmission of monetary policy as if the workers expect that there will be increase in inflation then they will demand increase in their salary which will led to higher inflation. By having an inflation target, central bank can anchor inflation expectations which will lead to the increase in the confidence of households and businesses in making decisions about saving and investment. The saving and investment channel affects consumption, housing investment and business investment by changing the incentives for saving and investment. The cash flow channel influences the decisions of households and businesses by changing the amount of cash they have available to spend on goods and services. The exchange rate channel affects the price of goods, asset price, exchange rate by influencing expectations about the future direction of economic activity and inflation.

In short, Monetary policy and economic variables are closely related, with monetary policy decisions significantly impacting various aspects of the economy. The transmission mechanism of monetary policy involve complex channels through which changes in monetary policy affect interest rate, economic activity and inflation.

2.3. Factors influencing the impact of Monetary Policy

The factors that influence monetary policy are exogenous and endogenous variables.

Exogenous variables are those variables which are done externally and those which are not determined and can be modified or eliminated like: Fiscal dominance, dollarization and global risk.

Endogenous variables are those variables which affect the monetary policy internally like: strategy, tactics and governance of the government.

3. Empirical Evidence

3.1. Study on the impact that monetary policy has on economic variable

The data collected for the research from rbi.gov and many more sites. The period which it covers is from 2014 to 2023. The variables that it includes are:

Interest rate: The central bank’s benchmark interest rate.

Money Supply: The total currency and other liquid assests in the economy.

Inflation Rate: The annual percentage change in consumer price index(CPI).

GDP growth rate: The annual percentage change in gross domestic product.

Unemployment rate: The percentage of the labor force that is unemployed.

3.2. Model Build

The linear regression model is:

GDP Growth Rate (%) = $\beta_0 + \beta_1 \times$ Interest Rate (%) + $\beta_2 \times$ Money Supply (USD) + $\beta_3 \times$ Inflation Rate (%) + $\beta_4 \times$ Unemployment Rate (%) + ε

Where β_0 is the intercept, β_1, β_2, β_3, and β_4 are the coefficients for the independent variables, and ε is the error term.

The model is fitted to the data using the lm() function in R, and diagnostic checks are performed to ensure that the assumptions of linear regression are met, such as linearity, normality of residuals, and homoscedasticity.

Year	Interest Rate (%)	Money Supply (USD)	Inflation Rate (%)	GDP Growth Rate (%)	Unemployment Rate (%)
2014	6.5	3.2 trillion	2.1	2.6	5.8
2015	5.8	3.5 trillion	1.8	3.1	5.3
2016	5.2	3.8 trillion	2.3	1.7	4.9
2017	4.7	4.1 trillion	1.9	2.4	4.4
2018	4.2	4.4 trillion	2.5	2.9	3.9
2019	3.8	4.7 trillion	1.7	2.2	3.7
2020	3.2	5.1 trillion	1.4	-3.5	8.1
2021	2.5	5.4 trillion	4.7	5.7	5.4
2022	3.1	5.8 trillion	3.2	3.9	3.6
2023	3.6	6.2 trillion	2.8	2.7	3.4

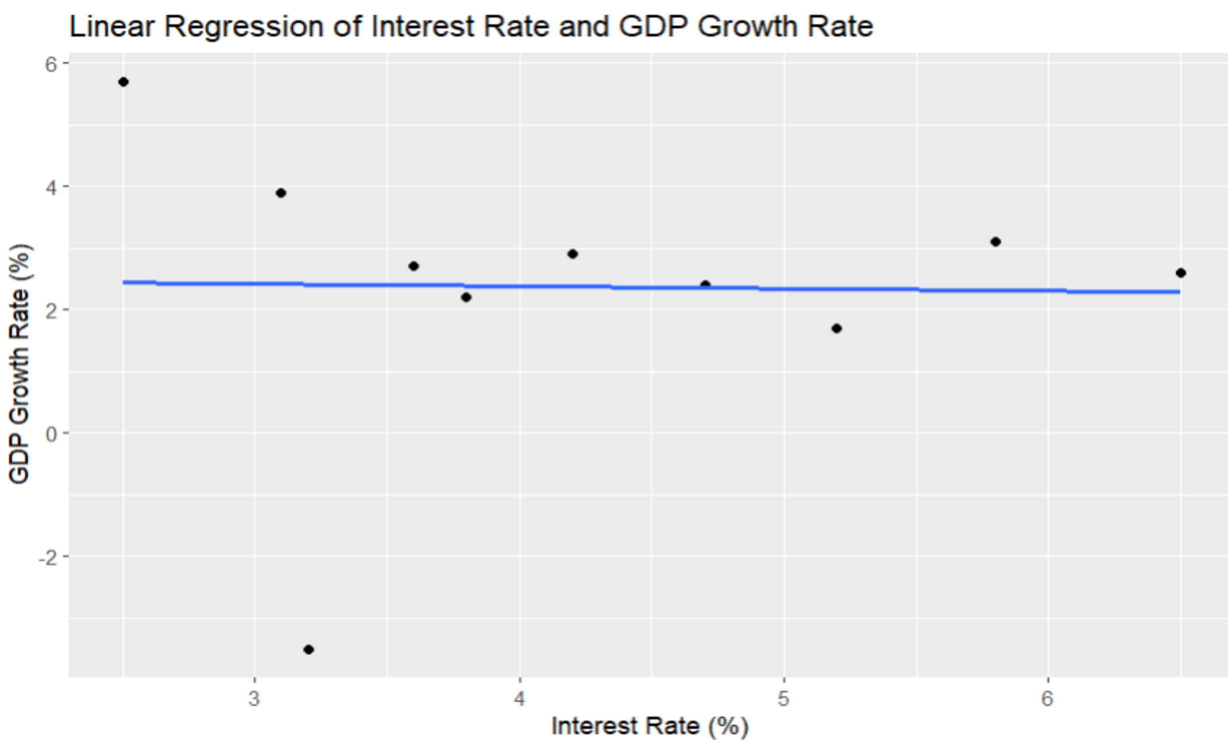
```
> summary(model)

Call:
lm(formula = data[[gdp_growth_rate_col]] ~ data[[interest_rate_col]])

Residuals:
    Min      1Q  Median      3Q     Max 
-5.9106 -0.1290  0.3102  0.7236  3.2626 

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                2.53304    2.86977   0.883    0.403
data[[interest_rate_col]] -0.03827    0.64795  -0.059    0.954

Residual standard error: 2.483 on 8 degrees of freedom
Multiple R-squared:  0.0004359, Adjusted R-squared:  -0.1245 
F-statistic: 0.003489 on 1 and 8 DF,  p-value: 0.9543
```

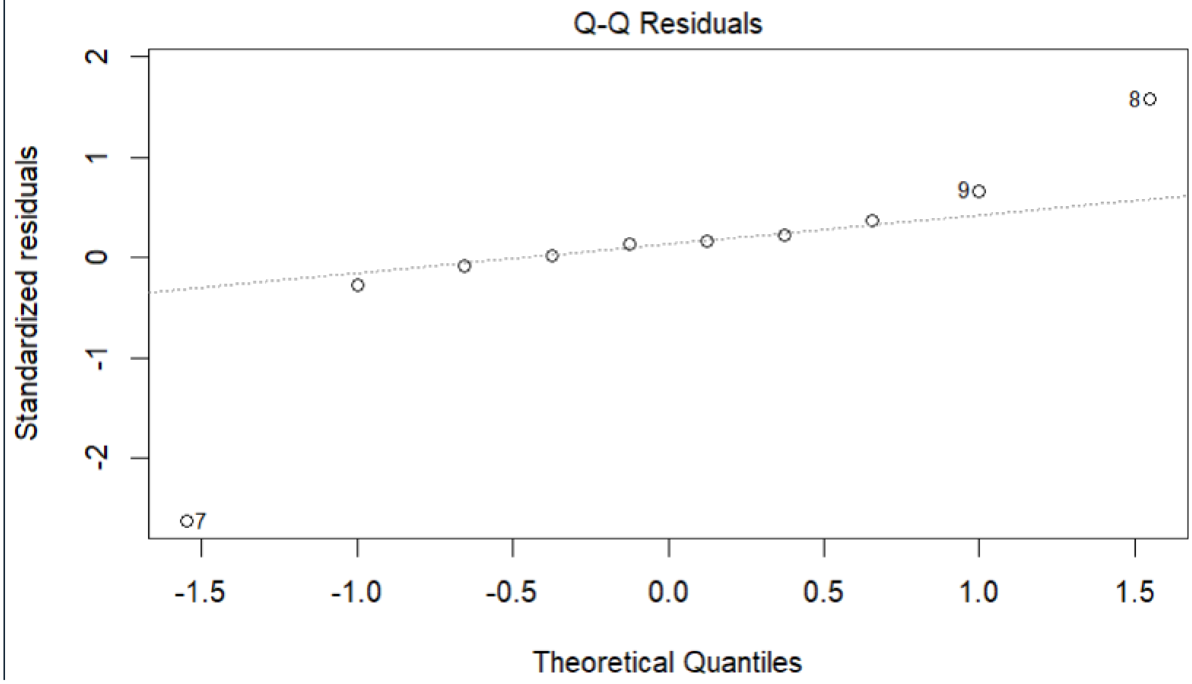
```
> predicted_gdp_growth
       1        2        3        4        5        6        7        8        9       10 
2.284269 2.311060 2.334023 2.353160 2.372296 2.387606 2.410569 2.437360 2.414397 2.395260 
```

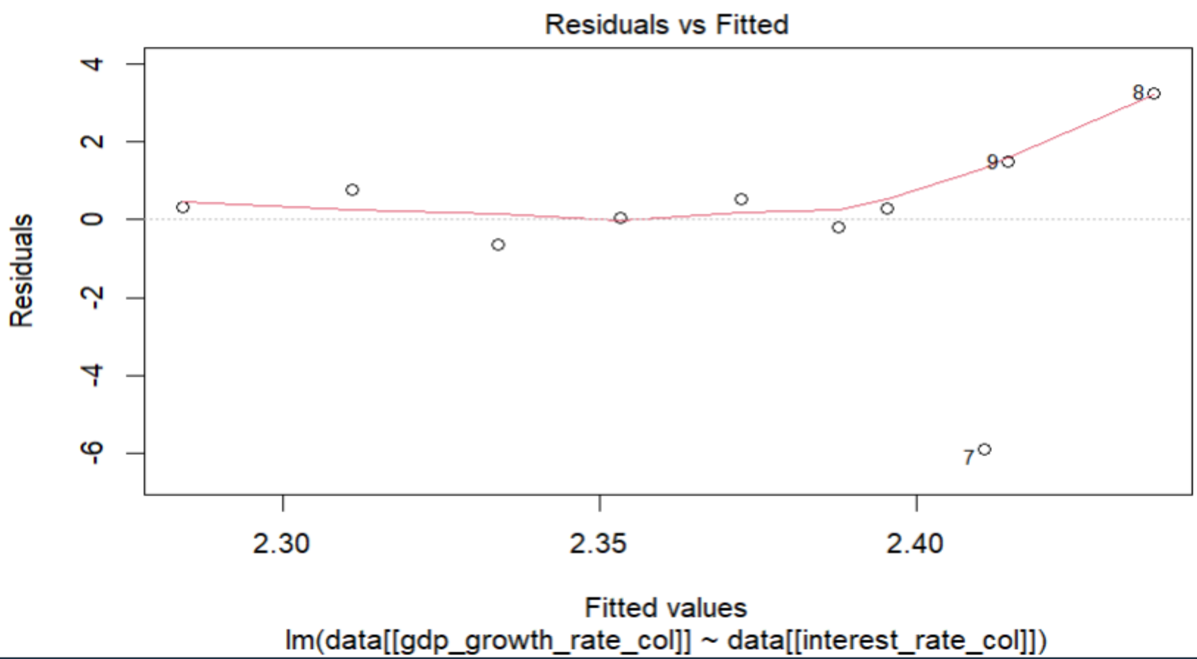
```
data = read.csv("effect of monetary policy on economic variables - Sheet1.csv")
str(data)
head(data)
install.packages("ggplot2)
```

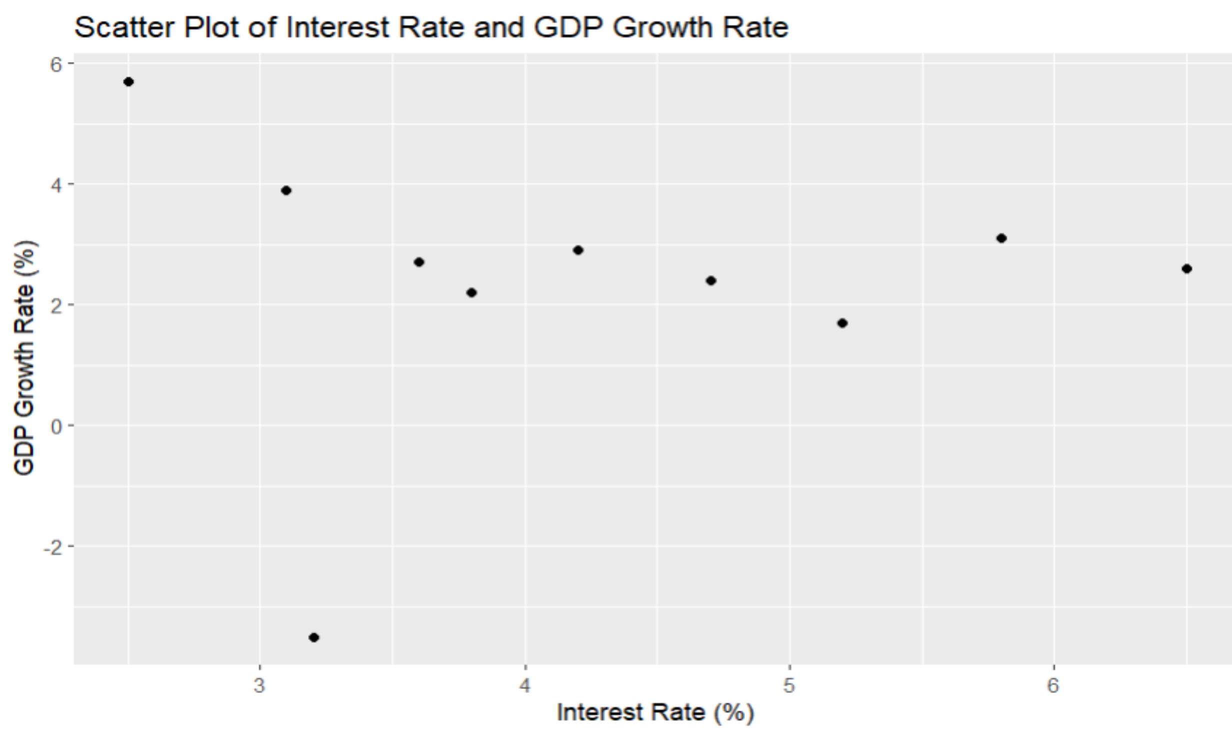
```
install.packages("ggplot2")
library(ggplot2)
interest_rate_col = names(data)[2]
gdp_growth_rate_col = names(data)[5]
ggplot(data, aes(x = data[, interest_rate_col], y = data[, gdp_growth_rate_col])) +
geom_point() +
labs(x = "Interest Rate (%)", y = "GDP Growth Rate (%)", title = "Scatter Plot of Interest Ra
and GDP Growth Rate")
```

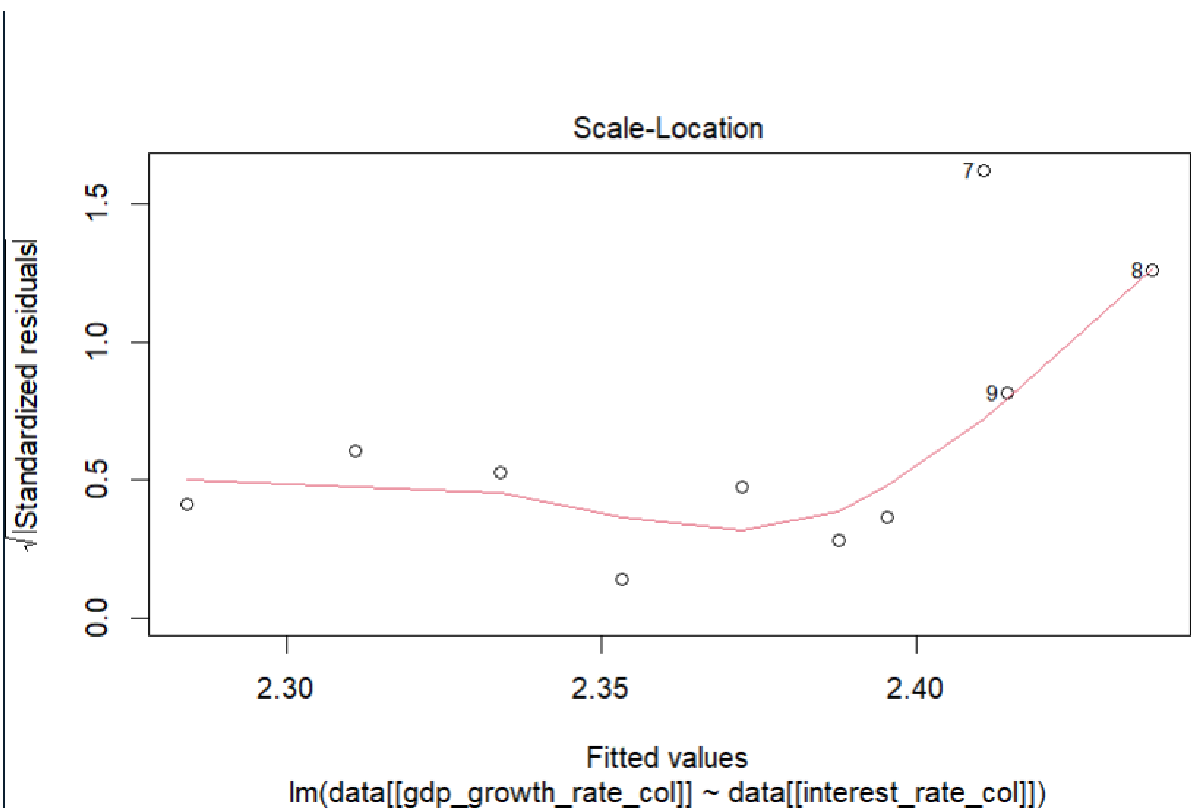
```
model <- with(data, lm(data[, gdp_growth_rate_col] ~ data[, interest_rate_col]))
summary(model)
plot(model, which = 1)
plot(model, which = 2)
plot(model, which = 3)
ggplot(data, aes(x = data[, interest_rate_col], y = data[, gdp_growth_rate_col])) +
geom_point() +
geom_smooth(method = "lm", se = FALSE) +
labs(x = "Interest Rate (%)", y = "GDP Growth Rate (%)", title = "Linear Regression of Interes
Rate and GDP Growth Rate")
```




Q-Q Residuals
Standardized residuals
Theoretical Quantiles
2
1
0
-1
-2
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
8
9
7


Residuals vs Fitted
Residuals
4
2
0
-2
-4
-6
2.30
2.35
2.40
8
9
7
Fitted values
lm(data[[gdp_growth_rate_col]] ~ data[[interest_rate_col]])


Scatter Plot of Interest Rate and GDP Growth Rate
GDP Growth Rate (%)
Interest Rate (%)
6
4
2
0
-2
3
4
5
6


Scale-Location
√|Standardized residuals|
Fitted values
lm(data[[gdp_growth_rate_col]] ~ data[[interest_rate_col]])
0.0
0.5
1.0
1.5
2.30
2.35
2.40
7
8
9

4.Policy Implications

4.1. Optimal Monetary Policy Strategies

Optimal monetary policy refers to the use of monetary tools and strategies so that the outcome is effective in minimal sources such as price stability, inflation etc. It has severe implications like it has severe economic environment and the presence of Various frictions such as habit formation and the zero lower bound on interest rates. It can influenced by the presence of multiple inefficiencies and multiple policymakers with various objectives.

4.2.Challenges in Implementing Effective Monetary Policy

To implement the effective monetary policy is very challenging because there is complexity in the transmission process, the weak and uncompetitive banking sector, clear communication etc. Capital flow can also pose challenges to effective monetary policy as there should be good flow of capital in the economy.

CONCLUSION

In this research paper we described the effect that economic variables has on the monetary policy and can RBI overcome or surpress these effects so that there will be no problems faced by the economy. We also focuses on the challenges that come while implementing the monetary policy. The graphs shows the effect of monetary policy on economic variables. The understanding of the impact of monetary policy on economic variables, policymakers can make informed decisions regarding interest rate adjustments and money supply management to achieve desired economic outcomes, such as promoting sustainable economic growth, maintaining price stability, and addressing unemployment concerns

Refrences

- Economic times
- Investopedia
- Corporate finance institute
- Cbn.gov
- Rba.gov
- Eestibank
- The monetary transmission bossonfed
- https://gov.com
- www.mospi.gov.in

RESEARCH PAPER

INCOME INEQUALITY AND ECONOMIC DEVELOPMENT

AUTHOR:DIVANSHI VIRMANI
RISHIKA MAKKAR
MANSHA

ABSTRACT

Income inequality is a pressing issue in many economies all across the world, with profound implications for economic development and social welfare. This research paper aims to examine the relationship between income inequalities and economic development, focusing on the theoretical underpinnings, empirical evidence, and policy implications. Through a comprehensive literature review and empirical analysis, this paper sheds light on the complex dynamics of income distribution and its effects on various aspects of economic growth and development.
The focus is on understanding that What happened to income inequality in the context of superior growth performance.

1. INTRODUCTION

1.1 Background:

Income inequality has emerged as a central concern in contemporary economics, attracting significant attention from policymakers, academics, and the public in recent years. With globalisation, technological advancements, and demographic shifts, income disparities have widened in many countries, raising questions about their implications for economic prosperity and social cohesion.

1.2 Overview:

Income inequality has become a critical issue in today's world, with the gap between the rich and poor continuing to widen. While economic development is essential for improving living standards, it is often associated with income inequality.
Income inequality refers to the unequal distribution of income among individuals or households within a society, where some people earn significantly higher incomes than others. While some degree of income inequality is inevitable in any market economy due to differences in skills, education, and entrepreneurial abilities, excessive inequality can have profound implications for economic development and social stability.
The effects of income inequality on economic development are complex and multifaceted. On one hand, moderate levels of inequality can serve as a boost for economic growth by investment, entrepreneurship, and innovation. Higher-income

individuals may channel their resources into productive activities, leading to increased capital formation and technological progress. Moreover, income disparities may foster competition, encouraging individuals to strive for excellence and contribute to overall productivity.

This Paper makes an attempt to understand and explore the complex relationship between income inequality and economic development, focusing on factors such as education, government policies, and labour market dynamics.

By examining these factors, we can better understand the causes of income inequality and develop strategies to reduce it.

Education plays a significant role in decreasing the income inequality of a nation. An individual can increase their earning potential by adding necessary skills and knowledge required in the labour market.Education can help to reduce income inequality by promoting social mobility and equal opportunities. But the access to quality education is limited by socio-economical factors such as poverty and discrimination. Thus, comes the need for government intervention to promote education access for all which is necessary to reduce income inequality.

The impact of education on income distribution is well documented. Studies have shown that countries with higher levels of education tend to have lower levels of income inequality.

For example, Initiatives like the Rashtriya Uchchatar Shiksha Abhiyan (RUSA) aim to improve the quality of higher education and increase enrollment rates, particularly among marginalised groups and rural populations. By investing in higher education, the government seeks to equip individuals with the skills and knowledge needed to access better-paying jobs and improve their economic prospects.

The income distribution can be significantly impacted by government policies. Tax policies can be used to redistribute the income from wealthy to poor. Initiating Progressive Taxation, where the rich are taxed at a higher rate than the poor, can help to reduce income inequality.

Likewisely, Social welfare policies, such as unemployment benefits and healthcare, can provide a safety net for the poor and reduce income inequality.

However, there are some government policies and actions which can make income inequality even greater or more severe.

For example, some policies that favour the wealthy, such as tax breaks for the rich, can widen the income gap. Similarly, policies that reduce social welfare spending can increase poverty and income inequality. Therefore, it is essential to analyse government policies to determine their impact on income distribution. Policies that promote social justice and equality should be prioritised, while policies that favour the wealthy should be reevaluated.

Also, Globalisation and technological change have had a significant impact on income distribution. Globalisation has led to increased competition in the labour market, which has put pressure on wages and thus increased income inequality.
Similarly, technological change has led to the displacement of workers in certain industries, leading to job losses and income inequality. However, labour market institutions can play a crucial role in reducing income inequality. Strong labour unions, for example, can negotiate and demand higher wages and better working conditions for workers, thereby reducing income inequality. Similarly, minimum wage laws can ensure that workers are paid a fair wage, reducing poverty and income inequality.

Addressing income inequality requires a multifaceted approach that combines policies to promote inclusive growth, such as investments in education, healthcare, and social protection, with measures to enhance labour market efficiency, ensure fair taxation, and reduce barriers to economic participation. By fostering greater equality of opportunity and redistributing resources to those in need, societies can not only mitigate the adverse effects of income inequality but also unleash the full potential of their human capital, driving sustainable economic development and shared prosperity.

1.3 Statement of problem:

This study aims to address the following research question:
What is the relationship between income inequalities and economic development? Specifically, how do variations in income distribution impact key economic indicators such as GDP growth, poverty rates, and human capital accumulation?

1.4 Importance of study:

Understanding the nexus between income inequality and economic development is necessary for formulating effective policy interventions aimed at promoting inclusive growth and reducing poverty. Highlighting the socio-economics impact of economic inequality. Explain why understanding the relationship between income and economic development is crucial. Emphasise how disparities in income distribution can affect key macroeconomics indicators such as GDP growth , investment and productivity. Introducing key concepts such as Gini Coefficient or measures of wealth distribution . This research contributes to the ongoing debate on equitable development strategies.

1.5 OUTLINE OF THE PAPER:

2. Literature Review
   2.1 Overview of Existing Research

2.Literature Review:

2.1 Overview of Existing Research:

Exploring the Relationship between Income Inequality and Economic Development: An Analysis of Global Trends and Regional Disparities.

Income inequality is a pervasive issue that impacts economic development on a global scale. Over the past few decades, there has been a noticeable trend towards an increasing concentration of income at the top of the income distribution worldwide, leading to significant disparities in economic progress . This trend is further compounded by the interplay between within-country and between-country inequality, along with the varying growth rates of emerging and advanced economies, which have influenced the dynamics of the middle class globally since 1980 . The concept of the "elephant curve" visually represents this phenomenon, showcasing a growing middle class in emerging economies juxtaposed with a squeezed middle class in wealthier nations, underscoring the profound impact of income inequality on economic development. Notably, emerging economies like China and India have made strides in narrowing the income gap with advanced economies, while within-country inequality has been on the rise in recent years, accounting for a significant portion of global inequality in 2020 compared to 1980. This disparity in income distribution not only impedes individual progress but also poses a threat to the economic stability and growth of developing nations, emphasising the urgency for inclusive policies and social programs to address income inequality for sustainable development .

Income inequality is a multifaceted phenomenon influenced by various factors related to economic development. Moller et al. Birčiaková et al. (2013) suggest that income inequality is intricately linked to economic development, educational expansion, racial and ethnic structure, urbanisation, as well as political and institutional factors. Kuznets' work Huynh et al. (2023) highlighted that economic development initially widens income inequality but eventually leads to its reduction as the economy progresses. This concept is further supported by Paukert's findings

Greenwood & Jovanovic (1990) that intra-country income inequality tends to rise and then decline with economic development.

Moreover, the relationship between economic freedom and income inequality has been extensively studied. Apergis & Cooray (2015) discuss how economic freedom can impact income inequality, indicating a trade-off between economic freedom and income equality. Similarly, Bergh & Bjørnskov (2021) delve into the effects of economic freedom on growth and income equality, building on earlier research. Additionally, 's empirical note Carter (2006) emphasises a positive but relatively inelastic relationship between economic freedom and income inequality.

Furthermore, the impact of tourism on income inequality has been explored by (Subramaniam et al., 2022), who reference the Kuznets curve to explain the inverted-U relationship between income and income inequality during economic development. Fang et al. (2020) also contribute to this discussion by highlighting the positive impact of tourism growth on income inequality in both developing and developed economies.

In conclusion, income equality and economic development are intricately intertwined, with various factors such as economic freedom, educational expansion, and tourism playing significant roles in shaping income distribution patterns within societies. Understanding these relationships is crucial for policymakers aiming to foster sustainable economic growth while addressing income inequality.

https://doi.org/10.1108/jbsed-07-2021-0102

2.2Key Concepts and Definitions:

We will discuss some concepts which give us a good understanding of income inequality.
They provide valuable insight into distribution of income with a population, highlighting disparities in wealth distribution.

GINI COEFFICIENT,Italian statistician Corrado Gini (1884–1965) created the GINI COEFFICIENT, which bears his name. GINI COEFFICIENT also referred as GINI INDEX is the most commonly used measure of inequality.
It is used as a measure of *income* inequality, but it can be used to measure the inequality of any distribution like the distribution of wealth.

It measures inequality on a scale from 0 to 1, where higher values indicate higher inequality. This can sometimes be shown as a percentage from 0 to 100%, this is then called the ‘Gini Index’ Perfect inequality is denoted by a value of 1, implying that no one gets any money and only one person gets all of it.This method can be used to tell us the difference we expect to find between the income of any two individuals related to the mean.

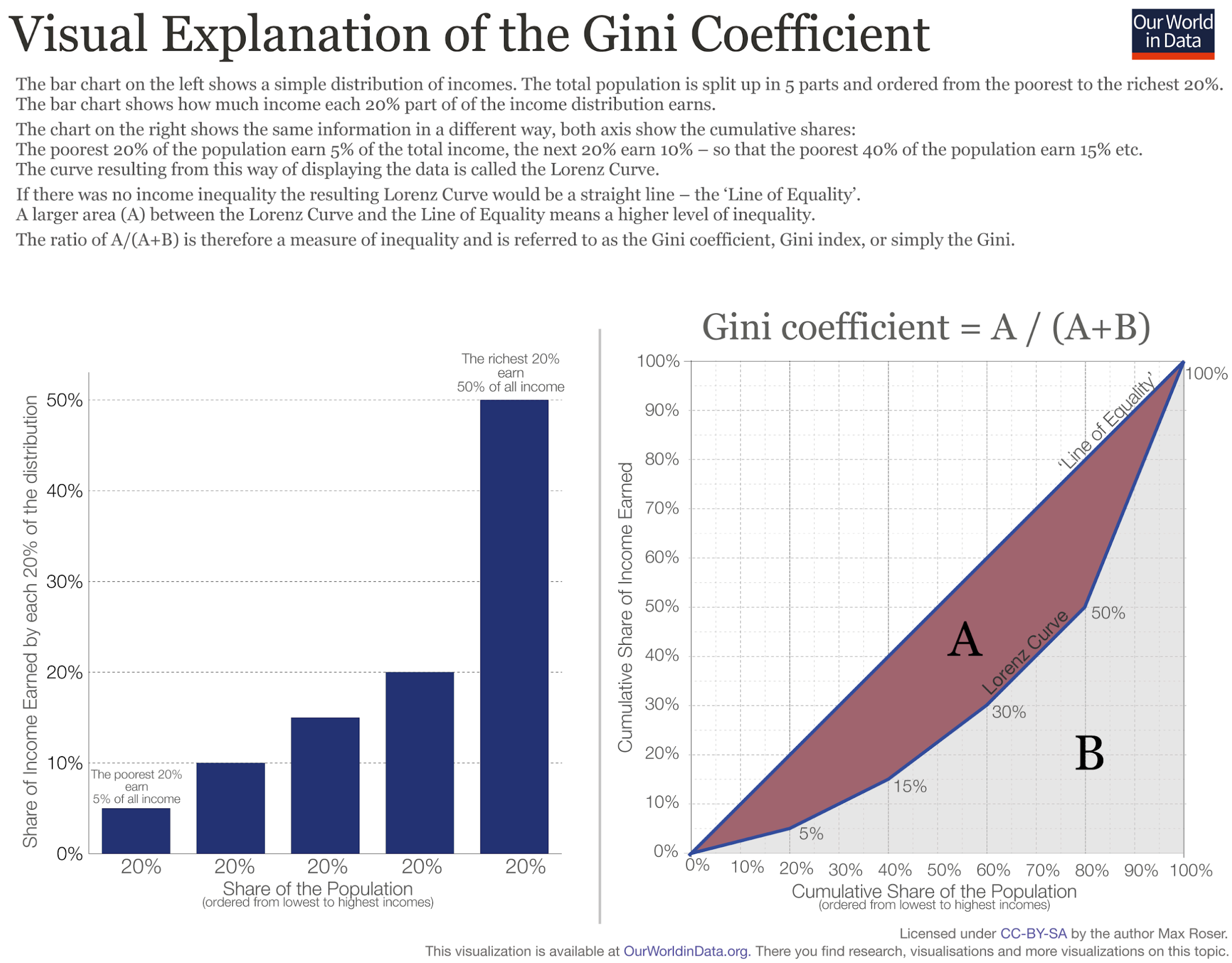
We will perform an experiment to understand this.

Imagine two strangers meeting at random on the street.They compare their incomes and find out how much richer one person is than the other.How much of a difference would we expect?

This expected gap between two randomly chosen people is what the Gini coefficient measures.It is calculated by summing the gaps between each pair of individuals.The difference in income between two randomly chosen people should be minimal in a situation where incomes are allocated equally. Where inequality is high, we would expect the gap to be large.If measured in absolute terms, however, this will also depend on how rich or poor the population is generally. The absolute difference in earnings cannot be large in a society where even the wealthiest individuals in society live in poverty. Conversely, where incomes are generally high, even very small relative differences between people's incomes can result in large absolute gaps.

Because of this, the predicted absolute difference between an individual's income and the population mean income is expressed by the Gini coefficient.Specifically, it is computed as the anticipated gap as a share of twice the mean income. Twice the mean income is the highest possible value for the average gap – a situation of perfect inequality, where one person has all the income and everyone else has none. So in this case of maximum inequality, the Gini coefficient is 1.

Since everyone makes the same amount of money, there can never be a scenario of complete equality, where the average disparity between any two pairs of individuals is equal to zero. The Gini coefficient in this instance is 0.The figure illustrates a second, visual definition of the Gini coefficient.

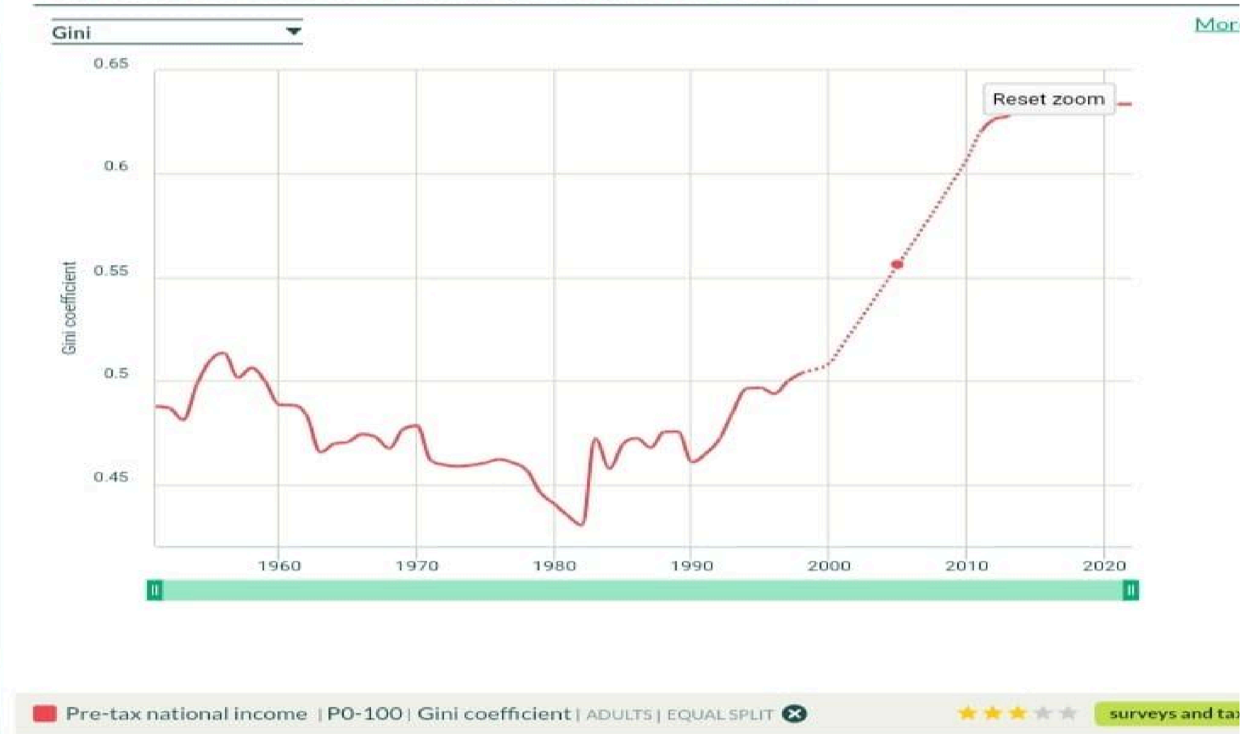


The left panel shows the share of income received by each fifth of a hypothetical population.The cumulative plot of this data is displayed in the right panel.This is known as a 'Lorenz curve'.

In a population where income is shared perfectly equally, the Lorenz curve would be a straight diagonal line: 10% of the population would earn 10% of the total income, 20% would earn 20% of the total income, and so on. This is shown in the chart as the 'line of equality'.

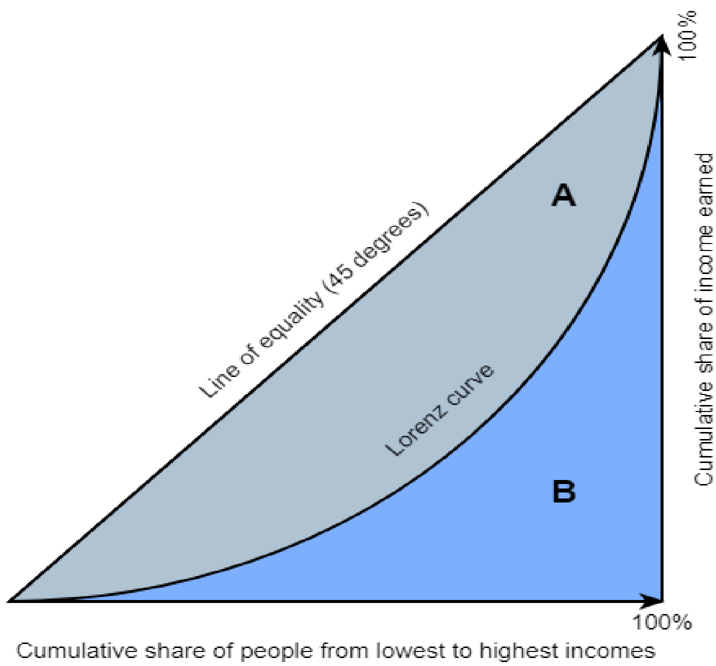
In the hypothetical population shown in the chart, though, incomes are not distributed equally.

Gini index of national income, India, 1952-2022



INCOME INEQUALITY AND LORENZ CURVE

The LORENZ CURVE was taken up by the American statistician MAX LORENZ to study the income problem.The number of consumer units in each income group and their income are assessed initially in the Lorenz curve approach. The consumer units are then ranked from lowest to greatest income by dividing them into percentages. The group's % contribution to the national income is displayed opposite each percentage. Ultimately, a table is created with the consumer units and the national income as percentages, and the information is plotted in a diagram. The Lorenz curve is the diagrammatic expression that emerges from this procedure. The percentage of consumer units arranged cumulatively from the lowest income to the highest income on the horizontal axis is displayed on this curve If the percentages that different groups receive from income are marked on the diagram and the marked points are connected, a curve showing the income distribution of the economy is obtained. In the curve diagram showing this income distribution, the inequality in income distribution decreases as it approaches the line drawn by forming an angle of 45 degrees, and the inequality in income distribution increases as it gets farther away.



We have another conceptual theory to discuss **<u>"TOP 1% VS BOTTOM 99%"</u>** ***"We are the 99%"*** is the slogan of the Anti-Wall Street campaign in the US. This slogan is highlighting the sheer inequality of income and wealth which exists in America. According to the papers, the top 1% of income holders in the US receive over 20% of all income and own about 35% of all America's wealth in 2007, the top 1% of wealth holders had net worth of nearly $22trn, or 33.8% of all net wealth in America ($65trn). In addition, he calculated that the richest 1% of earners collected 21.4% of all income in 2006, or $2 trillion annually out of a total of $9 trillion. If we add in the next 4% of wealth holders, then the top 5% have over 60% of all wealth and 37% of all income.

It is increasingly fashionable to argue that inequality is the cause of the crisis in capitalism.The cause of recessions and instability in capitalism in the 1970s was not assigned to inequality of income or wealth. Indeed, many mainstream and heterodox economists argued the opposite, namely that it was caused by wages rising to squeeze profits in overall national income.

But now, many Marxist economists argue that this current crisis is a product of wages being too low and profits too high. This leads to low wage earners being forced to borrow more and thus eventually causing a credit crisis. So it seems that the underlying cause of the capitalist crisis can vary. This eclectic approach has the drawback of making it difficult to determine what is causing the economic crisis—is it wages driving down profits, as they did in the 1970s or is it low wages leading to a collapse of demand in the noughties?

Inequality of wealth and income may not be the cause of capitalist crises, but it is certainly a product of capitalism.The private profit society is increasingly inefficient in delivering our needs and increasingly unstable. However, it's always been unfair and uneven. Actually, inequality of incomes and wealth in a society is a feature of all class societies, whether it is slavery, Asian absolutism, priestly castes, feudalism or capitalism. By definition, inequality accompanies a class society. After all, why would anybody want to be a member of the ruling elite if they did not enjoy the fruits, namely extreme wealth and income as well as power and status. Indeed, even priestly castes, supposedly engaged in waiting for the rewards of the ‘afterlife’ and telling their flock that they must also wait, were not slow in coming forward to benefit from the material life? Just consider the Vatican and other Christian cathedrals, the Middle Eastern mosques, the synagogues, and even further back to the ancient priesthood kings of Egypt, of the Aztecs and Incas of South America. Class society means that the ruling class controls the surplus generated by the labour time of the non-rulers, in whatever form. In other words, inequality is not a unique characteristic of capitalism that explains its ongoing booms and busts, but rather a result of all class societies.

How unequal is wealth and income distribution currently around the world (not just the US)? Well, Branco Milanovic at the World Bank has studied global inequality over the years carefully in a series of books and papers. His measure of inequality is not a measure between nations but between individuals, even though inequality is decided more by which country you live in than by inequality within a country. Global elites (the top 10%) take 57% of all the income created in the world and this ratio has hardly moved in 100 years. As Milanovic explains, this flies in the face of the predictions of mainstream economics that argue inequality should decline as countries get richer.

Simon Kuznets developed what is called the Kuznets curve in the mid-1950s.This supported the theory that inequality is low in pre-industrial societies where poverty is universal. When industrialisation takes place, as in the UK in the early 19th century and in the rest of major capitalist economies later in that century, then inequality grows between the urban and rural populations. Then, as an economy matures, the urban-rural gap narrows and the welfare state kicks in and inequality falls. So the Kuznets curve of inequality is an upside down U. The other mainstream theory is that, as global trade expands, the demand for low-skill labour in poor countries rises and so their incomes rise relative to those in mature economies. Inequality should decline.

Well, the evidence refutes both those theories. Global inequality of income – i.e. between world citizens – is high and has stayed high. Globally, the Gini coefficient, which calculates the ratio of wealth or income possessed by cohorts of people, has stayed extremely high, hovering around 70. It has risen in the US and the UK over the last 30 years and it has also jumped in China.   So the richest 1% of the world's population now receive nearly 15% of all the world's income, while the poorest 20% receive only just over 1%!

Almost all of the main capitalist economies have seen an increase in income inequality during the 1980s, but this has been especially true of the so-called Anglo-Saxon, deregulated "free market" nations like the US and the UK (the top 1% had about 10% in the early 1980s and now have nearer 20%). In Continental Europe, inequality has been static – the top 1% had about 10% in the 1980s and now have much the same ratio. Between 1993 and 2008, average real incomes in the US increased at a 1.3% annual pace; however, if the top 1% is taken out of the equation, the growth rate was only 0.75%. The incomes of the top 1% achieved a 3.9% a year rise, taking 58% of all the increase in real incomes between 1993 and 2008. Indeed, in the great US booms of 1993-2000 and 2001-2007, the top 1% achieved annual growth in their incomes of over 10%, while the bottom 99% achieved only 1.3-2.7% a year.

The top 10% of income earners in the US, with incomes over $110,00 a year, now receive half of all the income each year. The top 1% with incomes over $400,000 a year received nearly 25% of all income in 2007. This 1% constitutes 1.5m earners out of 150m earners in America. One of the stats that the authors reveal is that, in the Great Recession, the bottom 99% of income earners in the US suffered a 7% fall in their incomes, the largest drop since the Great Depression.   The other discovery is that it is the top 1% and even more, the top 0.1% of earners, who gained the most, compared with the top 5%. The inequalities are rising within the rich to create a super-rich elite. Whereas the top 1% took 9% of incomes in 1978 and the top 5% took the next 12% , to make 21% in all; by 2008, the top 1% took 21% while the top 5% took another 15% (making 36% in all). So the top 1% have reaped the biggest gains in the last 30 years. Also the top 0.1% took only 1% of incomes in 1978 but now take 6%!

The main reason for this higher concentration is the massive rise in the incomes going to the very top chief executives in the banks and big corporations. It's the same story in the UK. The UK's High Pay Commission found that bosses' salaries rose by 63% since 2002, but total pay packages for top company executives had gone up by 700% since 2002. In contrast, pay levels for the average worker in Britain rose only 27% before inflation and taxes.

These extremes of inequality and the particularly fast rise in that inequality in the last 30 years is now beginning to worry the strategists of capital in an environment of depression in the mature capitalist economies. It is not that they think inequality causes crises (although some are beginning to argue that), it is because they fear a social backlash from the 99% towards the 1%. And how justified that concern seems to be now, considering the worldwide movement emerging against the ultra-wealthy. "My worst fears about the potential loss of confidence in our leaders, institutions, and capitalism itself are realising," expressed City economist Jeremy Grantham. This is a hole that we have been digging for a while. We ought to be sincere in our endeavours to revive the fortunes of the typical labourer. Isn't it preferable for us to consciously choose the income distribution that strikes the optimal balance between social justice and work incentives on our own? *How about going back to the levels of income equality that existed under the Presidency of that notable Pinko, Dwight Eisenhower? And don't think for a second that this more equal income distribution somehow interfered with economic growth: the 50s and 60s were the heyday of sustained U.S. economic gains."*

2.3 Theoretical Framework:

Measuring income inequality is critical for several reasons, spanning social, economic, and political realms. Understanding the distribution of income within a society can help inform policies, highlight social issues, and guide economic decisions
Here are some key reasons why measuring income inequality is important:

1.Informing Public Policy and Redistribution:
Measuring income inequality helps policymakers understand the extent of disparities in income distribution. This knowledge can drive the creation of targeted economic policies, such as progressive taxation, welfare programs, and social security measures aimed at redistributing income more equitably across a society.

2.Assessing Economic Health and Stability:
High levels of income inequality can indicate potential instability within an economic system. Economies with severe inequality might experience less effective demand, as a significant portion of the population might not have sufficient income to consume goods and services. This can lead to slower growth and business cycles characterised by boom and bust phases.

3.Promoting Social Cohesion and Stability:

Large income disparities can lead to social unrest and tension. Communities with less inequality generally report higher levels of social cohesion and trust, which are vital for the stable and peaceful functioning of society. Conversely, high inequality can erode trust in institutions and increase the risk of social unrest.

4.Encouraging Sustainable Development:
Addressing income inequality is key to sustainable development. The United Nations' Sustainable Development Goals (SDGs), particularly Goal 10, focus on reducing inequality within and among countries. Sustainable development is more achievable in environments where economic resources are distributed fairly, enabling broader access to health, education, and other essential services.

5.Understanding Economic Mobility:
Measuring inequality helps in analysing economic mobility, or the ability for individuals or families to move between economic brackets. High inequality often corresponds with low mobility, suggesting that it is harder for people born into poverty to improve their economic status. This can perpetuate cycles of poverty.

6.Guiding International Comparisons and Investment Decisions:
Investors and businesses use measures of income inequality, like the Gini coefficient, to assess country risk and to make decisions about where to allocate resources. Countries with high inequality might be seen as riskier or less stable, potentially affecting foreign investment decisions.

7.Highlighting Gender and Ethnic Disparities:

Income inequality metrics can also be used to identify and address income disparities between different genders and ethnic or racial groups. This aids in crafting specific interventions aimed at promoting equality across different segments of the population.

In conclusion, measuring income inequality provides essential insights that help shape effective governance, promote equitable growth, and ensure social harmony.

INCOME SEGREGATION AND INCOME INEQUALITY:
INCOME SEGREGATION refers to the spatial separation of households with different income levels within a given area. This phenomenon occurs when people of similar income level cluster together in certain areas, leading to neighbourhoods that are predominantly composed of either high income or low income residents.
Contrarily, income inequality describes the unequal distribution of income across people or families within a society. It's a measure of how much wealth and income are concentrated in the hands of a few compared to the rest of the population .

The prerequisite for income segregation is INCOME INEQUALITY. If there were no income inequality, neighbourhoods would possess the same opportunities and conditions and thus no income segregation would likely occur. However, it is important to recognize that the relationship can be affected by many different factors and that it is not necessarily a one-to-one relationship.

According to studies, the existence of residential choices that are connected with income rather than income inequality alone leads to income segregation; hence, an income-based housing market and/or housing policy are essential.

Income segregation increased in the US between 1970 and 1990 as a result of rising inequality. Also, an increase in income segregation in economic school segregation is highly dependent on income inequality (as well as the role of educational policies).Income segregation (as well as racial segregation) is stronger in the public educational system than in the private system, especially in primary education. This is usually explained by the neighbourhood (metropolitan area) where schools are located. This can be due to the fact that regardless of one's financial background, anyone can attend a public school; therefore, a broader range of people will study at a public school. The relationship between income segregation and income inequality is complex: its results can be influenced by local political, economic, and city planning agendas, which can either work for or against the fight with income segregation and income inequality

Income segregation and income inequality are closely related and often reinforce each other in a vicious cycle. Here's how:

Spatial concentration of wealth:

Income segregation leads to the spatial concentration of wealth in certain areas. High-income individuals tend to live in neighbourhoods with better amenities, schools, and services, while low-income individuals are often relegated to areas with fewer resources and opportunities. This concentration of wealth exacerbates income inequality by concentrating economic resources in the hands of a few.

Limited Social Mobility:

Income segregation can also limit social mobility, the ability of individuals to move up or down the income ladder. When low-income individuals are isolated in neighbourhoods with limited economic opportunities, they may face barriers to accessing quality education, jobs, healthcare, and other essential services. This perpetuates a cycle of poverty and prevents upward mobility, contributing to persistent income inequality.

Unequal Access to Opportunities: Neighbourhoods with high levels of income segregation often have unequal access to opportunities. For example, affluent neighbourhoods may have better-funded schools, safer environments, and more job opportunities, while disadvantaged neighbourhoods may lack essential resources and infrastructure. This unequal distribution of opportunities further widens the gap between the rich and the poor.

Political and Economic Segregation: Income segregation can also lead to political and economic segregation, where affluent communities have more influence over local policies and decision-making processes. This can result in policies that further exacerbate income inequality, such as tax breaks for the wealthy or cuts to social welfare programs that benefit low-income individuals.

3.Methodology

3.1 Empirical topics in income inequality

Exploring empirical topics in income inequality can provide valuable insights into economic disparities, their causes, and potential solutions.
Here are several compelling empirical topics that you could consider for a research paper:
Impact of Globalization on Income Inequality: Investigate how global economic integration has affected income distribution within and between countries. Analyse trade liberalisation, foreign direct investment, and the outsourcing of jobs to lower-wage countries.

Technology and Wage Polarisation: Examine how technological advancements, particularly in AI and automation, have impacted jobs and wages across different sectors. Consider how technology-driven changes have contributed to the expansion of high-wage and low-wage employment, while diminishing middle-wage jobs.

Effects of Tax Policies on Income Inequality: Study the role of progressive versus regressive tax systems and how different tax policies affect income distribution. Include analysis of capital gains taxes, income taxes, and corporate taxes.

Educational Attainment and Income Inequality: Explore the relationship between education levels and income disparity. Assess how educational opportunities are distributed across different socio-economic groups and how this affects income potential.

Health Disparities and Income Inequality: Investigate how differences in income levels correlate with health outcomes and access to healthcare services. Explore the impact of public health policies and insurance systems on these disparities.

Gender and Racial Income Disparities: Focus on how income inequality manifests across different genders and races.Analyse wage gaps, employment opportunities, and the glass ceiling in various industries.

Income Mobility and Inequality: Study the dynamics of income mobility across generations. Examine factors that contribute to or hinder upward mobility, such as education, family background, and social networks.

The Impact of Minimum Wage on Income Inequality: Analyse empirical data to assess how changes in the minimum wage have impacted income distribution and poverty rates. Consider different regional economic contexts for a nuanced analysis.

Gig Economy and Income Inequality: Examine the role of the gig economy in income inequality, focusing on job security, wage levels, and worker rights. Consider how gig work compares to traditional employment in terms of income stability and social benefits.

Wealth Concentration and Economic Power: Explore how wealth concentration among the ultra-rich affects economic decision-making and policy influence. Study the implications of such concentration for average citizens and overall economic health.

3.2 Problems in analysing income inequality

Analysing income inequality presents several challenges, due to the complexity and multifaceted nature of the factors that influence it. Some key problems encountered in such analyses include:

1. Data Availability and Quality

Incompleteness: In many regions, especially in less developed countries, income data can be sparse, irregularly collected, or non-existent.

Inaccuracy: Income is often underreported, particularly at the extremes of the wealth spectrum. Wealthy individuals may hide income to avoid taxes, while low-income groups may not be captured accurately in surveys.

Inconsistency: Different countries may have different standards and methods of data collection, making comparisons challenging.

2. Complex Causality and Interdependencies

Multiple Factors: Income inequality is influenced by a complex interplay of economic, political, social, and cultural factors. Disentangling the effects of individual variables can be challenging.

Feedback Loops: Changes in income inequality can feed back into the factors that affect it, such as education and political power, creating cycles that are difficult to break and analyse.

3. Measurement Challenges

Choosing Metrics: There are various metrics to measure income inequality, such as the Gini coefficient, the Palma ratio, and income shares across percentiles. Each metric offers different insights and may tell different stories about the level of inequality.

Temporal Dynamics: Inequality may change over different time scales. Short-term fluctuations might mask long-term trends, requiring careful analysis to understand underlying movements.

4. Global vs. Local Dynamics

Global Influences: Global economic trends, such as international trade policies or financial crises, can impact local income distributions. Analysing these effects requires understanding both local contexts and global dynamics.

Regional Variations: Inequality must often be analysed both at a national and a local level, as economic conditions can vary greatly within countries.

5. Theoretical and Ideological Biases

Economic Models: Different economic theories provide different explanations and solutions for inequality, which can influence the interpretation of data.

Political Biases: Research into inequality can be politically charged. Studies might be skewed by the ideological positions of those funding or conducting the research.

6. Ethical and Privacy Concerns

Sensitive Data: Gathering detailed income data often raises privacy concerns and ethical issues, potentially limiting the scope of research.

Stigmatisation: Reporting on income inequality risks stigmatising certain groups or regions, which must be handled with sensitivity to avoid exacerbating social tensions.

Overall, analysing income inequality requires a nuanced approach that considers a variety of data sources, acknowledges theoretical and methodological limitations, and remains alert to the broader socio-political implications of the findings.

3.3 statistical Methods

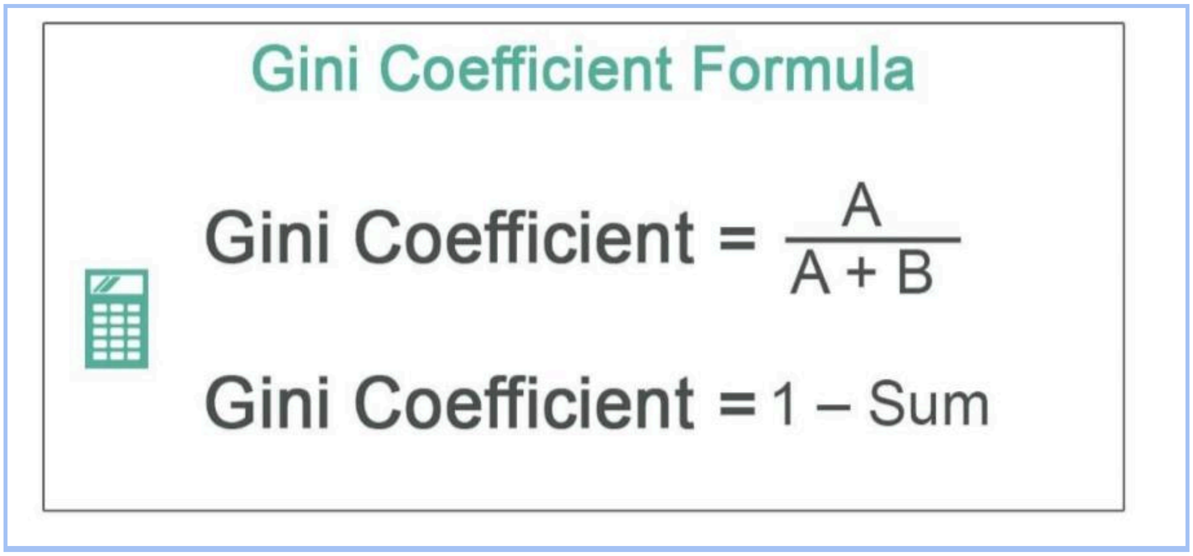
Measuring inequality is crucial in economics, social sciences, public policy, and health studies to understand how resources such as income, wealth, or opportunities are distributed among individuals or groups within a society. Several statistical methods are used to measure and analyse inequality:

Gini Coefficient:

The Gini coefficient is one of the most widely used measures of inequality. It calculates the extent to which the distribution of income (or wealth) among individuals or households deviates from a perfectly equal distribution. A Gini index of 0 represents perfect equality, while a Gini index of 1 indicates maximum inequality.

The coefficient is often calculated based on the Lorenz curve, which plots the cumulative percentages of total income received against the cumulative number of recipients, starting with the poorest individual or household.

Let us understand the formula for us to find the value through the Gini coefficient calculator from the explanation below.



If A=0, the Lorenz curve is the line of equality. When A=0, the Gini index is 0. The Gini coefficient is large in case A is very large and B is small. It indicates there is huge income/wealth inequality.

How To Calculate?

Gini coefficient countries take the following steps to find the value that indicates the inequality among their citizens' incomes is as follows:

1.Organise the data into a table with the category head mentioned below.

Fraction of Income	Fraction of Population	% of Population that is richer	Score

It is important to note that all the rows must organise from the poorest to the richest. For instance, if it states that the bottom 10% of the population earns 3% income, write 0.03 in the 'Fraction of Income' column. Next, write 0.10 in the 'Fraction of Population' column. Similarly, fill these two columns with other percentages given.

2.Fill '% of Population that is richer' column by adding all terms in 'Fraction of Population' below that row.

Fraction of Income	Fraction of Population	% of Population that is richer	Score
0.03	0.1	0.9	
0.15	0.5	0.4	
0.85	0.4	0	

For instance, we fill the first row in the '% of Population that is richer' column. Then, we will add 0.50 and 0.40, the rows in 'Fraction of population' below it. Hence, we get 0.90.

3.Calculate the Score for each of the rows. The formula for Score is:

Score = Fraction of Income * (Fraction of Population + 2 * % of Population that is richer).

For instance, score for the 1st row is 0.03*(0.10+2*0.90) = 0.057

4. Next, add all the terms in the 'Score' column. Let us call it 'Sum.'
Calculate the Gini coefficient using the formula: = 1 – Sum

Example #1

The Gini coefficient of 2 countries based on citizens' income is as under.

	2010	2011	2012	2013	2014	2015
Country A	0.4	0.43	0.48	0.5	0.52	0.57
Country B	0.38	0.37	0.35	0.34	0.3	0.29

- **Interpret the trend of income inequality in the two countries**
- **Which country has higher income inequality in 2015?**

Solution:

a) The Gini coefficient of Country A has shown a rising trend from 0.40 in 2010 to 0.57 in 2015. Hence, income inequality in Country A has risen in these years. On the other hand, the coefficient of Country B has fallen from 0.38 in 2010 to 0.29 in 2015. Therefore, income inequality in Country B has declined over these years.

b) The coefficient of Country A (0.57) is more than that of Country B (0.29). Hence, Country A had higher income inequality in 2015.

Example #2

In a particular country, the lowest 10% of the earners make 2% of all wages. The next 40% of earners make 13% of wages. The following 40% of earners make up 45% of all wages. Moreover, the highest 10% of all earners make 40% of all wages. Calculate the Gini coefficient of the country.

Solution:

Use the following data for the calculation.

	A	B
1	Fraction of Income	Fraction of Population
2	0.02	0.1
3	0.13	0.4
4	0.45	0.4
5	0.4	0.1

Let us compile the above information in table format. But, first, the information has to be compiled by organizing the rows from the poorest to the richest.

	A	B	C	D	E
1	Fraction of Income	Fraction of Population	% of Population that is Richer	Score	Score Formula
2	0.02	0.1	0.9	0.038	=A2*(B2+2*C2)
3	0.13	0.4	0.5	0.182	=A3*(B3+2*C3)
4	0.45	0.4	0.1	0.27	=A4*(B4+2*C4)
5	0.4	0.1	0	0.04	=A5*(B5+2*C5)

=1-D7

Gini Coefficient = 1 - Sum

	A	B	C	D	E
1	Fraction of Income	Fraction of Population	% of Population that is Richer	Score	Score Formula
2	0.02	0.1	0.9	0.038	=A2*(B2+2*C2)
3	0.13	0.4	0.5	0.182	=A3*(B3+2*C3)
4	0.45	0.4	0.1	0.27	=A4*(B4+2*C4)
5	0.4	0.1	0	0.04	=A5*(B5+2*C5)
6					
7		Sum of Scores		0.53	
8		Gini Coefficient		=1-D7	

Sum of Scores = 0.038+0.182+0.27+0.04 =0.53

The coefficient will be –

D8 =1-D7

	A	B	C	D	E
1	Fraction of Income	Fraction of Population	% of Population that is Richer	Score	Score Formula
2	0.02	0.1	0.9	0.038	=A2*(B2+2*C2)
3	0.13	0.4	0.5	0.182	=A3*(B3+2*C3)
4	0.45	0.4	0.1	0.27	=A4*(B4+2*C4)
5	0.4	0.1	0	0.04	=A5*(B5+2*C5)
6					
7		Sum of Scores		0.53	
8		Gini Coefficient		0.47	

Frequently Asked Questions (FAQs)

What does a low Gini coefficient mean?
A Gini coefficient of zero indicates there is equal income distribution. Whereas a number near one shows greater inequality. Thus, the lower the Gini coefficient, the more similar the society.

What are the Lorenz curve and Gini coefficients?
The Lorenz curve is the focal point in determining the Gini coefficient, a mathematical portrayal of inequality levels. Still, since Lorenz curves are mathematical calculations depending on fitting a continuous curve to incomplete and discontinuous data, they may be imperfect evaluations of true inequality.

What is the U.S.e US Gini coefficient?
In 2021, as per the Gini coefficient, household income distribution in the United States was 0.49. In addition, in 1990, this figure was at 0.43, which shows an increase in income inequality in the U.S. over the last 30 years.

What does a negative Gini coefficient mean?
The Gini coefficient can range from 0, complete equality, to 1 (absolute inequality). Often, it is expressed as a percentage ranging between 0 and 100. However, if negative values are possible, the Gini coefficient can be more than 1.

Lorenz Curve:

This is a graphical representation of the distribution of income or wealth. The curve shows the proportion of total income earned by the bottom x% of the population. The further the curve is from the diagonal line (the line of equality), the greater the inequality.

Example #1

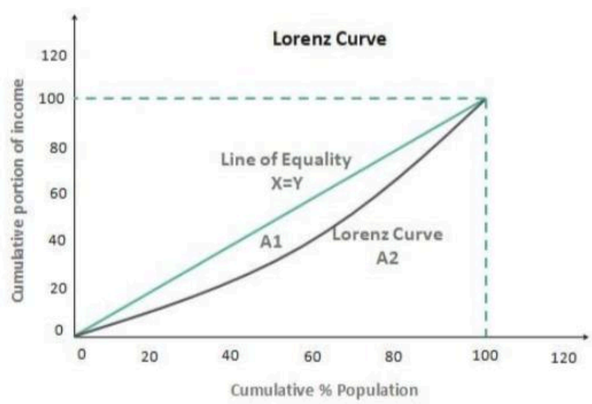
Let us consider an economy with the following population and income statistics: –

Population	Income Portion %
0	0
20	10
40	20
60	35
80	60
100	100

And for the line of perfect equality, let us consider this table: –

Population	Income Portion %
0	0
20	20
40	40
80	80
100	100

Let us now see how a graph for this data looks: –



As we can see, there are two lines in the graph of the Lorenz curve, the curved red line, and the straight black line. The black line represents the imaginary line of equality, i.e., the ideal graph when income or wealth is equally distributed amongst the population. But on the other hand, the red curve, the Lorenz curve, which we have been discussing, represents the distribution of wealth among the people.

Hence, we can say that the Lorenz curve is the graphical method of studying dispersion. Gini coefficient, also known as the Gini index, can be computed as follows. Let us assume in the graph area that A1 represents the Lorenz curve and the line, and A2 is the line below the curve. So,

Gini coefficient = A1/ (A1+ A2)

The Gini coefficient lies between 0 and 1, 0 being the instance where there is perfect equality and 1 being the instance where there is perfect inequality. Therefore, the higher the area enclosed between the two lines represents higher economic inequality.

We can say that there are two indicators for measuring income inequality:

- The Lorenz curve is the visual indicator
- The Gini coefficient is a mathematical indicator

Income inequality is a pressing issue across the world. So,**what are the reasons for inequality in an economy ?**
Corruption
Education
Tax
Gender differences
Culture
Race and caste discrimination
The difference in preferences of leisure and risks.

Reasons for income inequality

1.The distribution of economic characteristics across the population should be considered.

2. Analysing how the differences give rise to different outcomes in terms of income.

3. A country may have a high degree of inequality because of:

i A great disparity in these characteristics across the population.

ii These characteristics generate huge effects on the amount of income a person earns.

4. ANALYSIS AND RESULT

4.1Descriptive Statistics

Table 1 reports the descriptive statistics for the variables. Middle-income countries show the higher average income with Gini net income of 42.06 and Gini market income of 47.14 and the higher Standard deviation of Gini net income of 8.60 and Gini market income of 8.17 in comparison to high and low-income countries. High-income countries provide higher GDP per capita of 32576.59 and mean GDP per capita growth of 1.90 with a standard deviation of 17476.96 and 3.07 in comparison with middle and low-income countries. Nonetheless, high-income countries show a higher average financial development index of 0.54 with a standard deviation of 0.20 as compared to the middle and low-income countries.

Variables	Mean	Std Dev	Minm	Maxm
Panel A: High-Income Countries				
Gini Net (GINIn)	30.70	6.69	16.60	51.41
Gini Market (GINIm)	46.37	4.67	29.96	49.97
Per capita GDP (ED)	32576.59	17476.96	4266.11	91593.63
Per capita GDP growth(g)	1.90	3.07	−14.55	13.08
Financial Development Index (FD)	0.54	0.20	0.10	1.00
Dependency Ratio (AGE)	50.79	4.53	40.31	70.82
Govt. Expenditure (EXP)	19.59	4.10	9.95	36.26
Inflation (INF)	5.69	13.18	−9.68	259.99
Trade (TRD)	75.00	34.31	16.57	188.97
Urban Population (POP)	77.55	11.12	32.06	97.77
Panel B: Middle-Income Countries				
Gini Net (GINIn)	42.06	8.60	18.07	67.21
Gini Market (GINIm)	47.14	8.17	23.23	76.88
Per capita GDP (ED)	3847.82	2521.39	331.97	11797.45
Per capita GDP growth(g)	2.51	4.97	−30.71	33.03
Financial Development Index (FD)	0.24	0.13	0.002	0.68
Dependency Ratio (AGE)	64.12	14.29	35.33	102.61
Govt. Expenditure (EXP)	13.78	5.26	2.97	30.12
Inflation (INF)	65.45	358.67	−26.29	6261.24
Trade (TRD)	74.47	40.03	12.008	220.40
Urban Population (POP)	53.95	18.19	16.946	91.45
Panel C: Low-Income Countries				
Gini Net (GINIn)	41.10	6.05	21.85	61.46
Gini Market (GINIm)	47.35	7.00	25.42	70.93
Per capita GDP (ED)	542.96	355.91	202.43	2363.67
Per capita GDP growth(g)	1.47	6.00	−47.72	37.12
Financial Development Index (FD)	0.08	0.02	0.02	0.19
Dependency Ratio (AGE)	94.55	7.70	72.84	110.92
Govt. Expenditure (EXP)	13.00	4.33	4.83	31.55
Inflation (INF)	14.79	20.07	−10.73	128.76
Trade (TRD)	49.31	14.13	19.68	91.37
Urban Population (POP)	23.88	10.47	5.05	44.36

Note: Std Dev indicates the STANDARD DEVIATION , Minm denotes the minimum and Maximum represents the maximum

Finally, other control variables including Dependency Ratio, Govt. Expenditure, Inflation, Trade and Urban Population present considerable variations among high, middle and low-income countries. Low-income countries show a higher rate of Dependency Ratio with a mean value of 94.55 and a standard deviation of 7.70 than the high income and middle-income countries. Inflation shows the significant variations in middle-income countries with an average value of 65.45 and a standard deviation of 358.67 than in high income and low-income countries. Moreover, Government expenditure (average of 19.59 and standard deviation of 4.10), Trade (average of 75.00 and standard deviation of 34.31) and Urban population (average of 77.55 and standard deviation of 11.12) show comparatively higher deviations in high-income countries than in middle and low-income countries.

4.2Regression Analysis

Static linear regression models

Assume the standard regression model

$$y_i = \boldsymbol{x}_i \boldsymbol{\beta} + u_i$$

for $i = 1, \ldots, N$, where y_i and u_i are scalars, $\boldsymbol{x}_i$ is $1 \times K$ and $\boldsymbol{\beta}$ is $K \times 1$. Note that the above expression defines a generic regression model, so y_i does not need to be related to income or income inequality. Stacking the $\boldsymbol{x}_i$:s into one matrix gives $\boldsymbol{X}$, which is $N \times K$. Moreover, let $\boldsymbol{y} = (y_1, \ldots, y_N)'$. The expression for the least squares estimate of $\boldsymbol{\beta}$ is then

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$$

An estimator $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ is said to be *asymptotically consistent* if $\lim \hat{\boldsymbol{\beta}} = \boldsymbol{\beta}$ as $N \to \infty$. By rewriting (12) and applying Slutsky's theorem, it is possible to show that the limit condition stated above is satisfied if and only if $\mathbb{E}[\boldsymbol{x}_i | u_i] = \mathbf{0}$ so that $\mathbb{E}[\boldsymbol{x}_i u_i] = \mathbf{0}$. In practice, this means that an observation $\mathbf{x}_i$ must not be correlated with the random error term u_i. However, in econometrics, it is often the case that there is some sort of dependence between one or several of the independent variables and the error term. An independent variable that is correlated with the error term is said to be *endogenous*, while a variable uncorrelated with the error term is called *exogenous*. In the case of endogenity, $\hat{\boldsymbol{\beta}}$ is not a consistent estimator of $\boldsymbol{\beta}$. To remedy this problem, the well-known concept of *instrumental variables* will be briefly reviewed.

The empirical analysis tried to capture the fundamental linkage between economic and financial development on income inequality applying the three primary hypotheses, namely, Kuznets hypothesis, financial curve hypothesis, and financial Kuznets curve hypothesis. Following Beck et al., theoretical predictions were tested using Dynamic Panel Generalised Method of Moments (GMM) developed by Holtz-Eakin, Newey ,and Rossen ,Arellano and Bond and Arellano and Bover . GMM overcomes the problems of potential endogeneity and unobserved country-specific effects in the data ensuring the best fit for panel data.

The dynamic specifications of each panel model are as follows.

Kuznets curve hypothesis;

$$GINI_{it} = \alpha + \beta_0 GINI_{it-1} + \beta_1 \ln ED_{it} + \beta_2 \ln ED_{it}^2 + \delta \ln Z_{it} + \varepsilon_{it}$$

Financial curve hypothesis;

$$GINI_{it} = \alpha + \beta_0 GINI_{it-1} + \beta_1 \ln FD_{it} + \beta_2 \ln FD_{it}^2 + \delta \ln Z_{it} + \varepsilon_{it}$$

Financial Kuznets curve hypothesis;

$$GINI_{it} = \alpha + \beta_0 GINI_{it-1} + \beta_1 \ln EDG_{it} + \beta_2 G_{it} + \beta_3 \ln FDG_{it} + \delta \ln Z_{it} + \varepsilon_{it}$$

where $GINI_{it}$ captures Gini net/gross income, $GINI_{it-1}$ captures lagged value of the dependent variable, ED_{it} captures economic development, FD_{it} captures financial development, and G_{it} captures GDP_{it} per capita growth. EDG_{it} is the product of economic development and GDP growth rate, FDG_{it} is the product of financial development and per capita GDP growth, while Z_{it} captures other control variables such as trade openness, inflation, government expenditure, dependency ratio, and urban population.

5.CONCLUSION

The study of income inequality and its implication for economic development reveals factors that shape the socio-economic status of nations. Through a paper review of theoretical frameworks, empirical evidence, historical perspectives, and policy implications, several key insights emerge.

Firstly, it is evident that income inequality is not merely a distributional issue but a multifaceted phenomenon with profound implications for overall economic growth and societal well-being. The theoretical data provide diverse perspectives on how income distribution affects investment, savings, human capital accumulation, consumption patterns, social mobility, and political stability. Empirical studies showcase these theories, demonstrating that excessive income inequality can hinder economic development by weakening investment, reducing aggregate demand, limiting opportunities for human capital development, and undermining social cohesion.

Policy interventions such as progressive taxation, social safety nets, education and skill development programs, labour market reforms, and wealth redistribution policies can play a crucial role in mitigating the adverse effects of income inequality and promoting equitable growth.

Furthermore, the role of globalisation and technological change in shaping income distribution dynamics cannot be overlooked. Globalisation has led to both opportunities and challenges, contributing to both rising and falling income inequality depending on the context.

In conclusion, the study of income inequality and its impact on economic development underscores the need for a holistic approach that considers the interplay of economic, social, political, and institutional factors. Addressing income inequality is not only a moral imperative but also a pragmatic strategy for promoting sustainable development and fostering a more prosperous and equitable society. By implementing targeted policy interventions and fostering inclusive growth strategies, policymakers can work towards creating a future where economic opportunities are accessible to all, irrespective of their socio-economic background.

6.REFERENCES

https://doi.org/10.1146/annurev.so.13.080187.001525
https://thenextrecession.wordpress.com/
https://www.wallstreetmojo.com

https://wid.world/
https://en.wikipedia.org/
https://link.springer.com/
https://www.worldbank.org/

Economies | Free Full-Text | Globalisation and Income Inequality in Developing Economies: A Comprehensive Analysis. (n.d.) retrieved April 15, 2024, from www.mdpi.com/2227-7099/12/1/23

RESEARCH PAPER
INCOME INEQUALITY AND ECONOMIC DEVELOPMENT

AUTHOR:DIVANSHI VIRMANI
RISHIKA MAKKAR
MANSHA

ABSTRACT

Income inequality is a pressing issue in many economies all across the world, with profound implications for economic development and social welfare. This research paper aims to examine the relationship between income inequalities and economic development, focusing on the theoretical underpinnings, empirical evidence, and policy implications. Through a comprehensive literature review and empirical analysis, this paper sheds light on the complex dynamics of income distribution and its effects on various aspects of economic growth and development.
The focus is on understanding that What happened to income inequality in the context of superior growth performance.

1. INTRODUCTION

1.1 Background:

Income inequality has emerged as a central concern in contemporary economics, attracting significant attention from policymakers, academics, and the public in recent years. With globalisation, technological advancements, and demographic shifts, income disparities have widened in many countries, raising questions about their implications for economic prosperity and social cohesion.

1.2 Overview:

Income inequality has become a critical issue in today's world, with the gap between the rich and poor continuing to widen. While economic development is essential for improving living standards, it is often associated with income inequality.
Income inequality refers to the unequal distribution of income among individuals or households within a society, where some people earn significantly higher incomes than others. While some degree of income inequality is inevitable in any market economy due to differences in skills, education, and entrepreneurial abilities, excessive inequality can have profound implications for economic development and social stability.
The effects of income inequality on economic development are complex and multifaceted. On one hand, moderate levels of inequality can serve as a boost for economic growth by investment, entrepreneurship, and innovation. Higher-income

individuals may channel their resources into productive activities, leading to increased capital formation and technological progress. Moreover, income disparities may foster competition, encouraging individuals to strive for excellence and contribute to overall productivity.

This Paper makes an attempt to understand and explore the complex relationship between income inequality and economic development, focusing on factors such as education, government policies, and labour market dynamics.
By examining these factors, we can better understand the causes of income inequality and develop strategies to reduce it.

Education plays a significant role in decreasing the income inequality of a nation. An individual can increase their earning potential by adding necessary skills and knowledge required in the labour market.Education can help to reduce income inequality by promoting social mobility and equal opportunities. But the access to quality education is limited by socio-economical factors such as poverty and discrimination. Thus, comes the need for government intervention to promote education access for all which is necessary to reduce income inequality.
The impact of education on income distribution is well documented. Studies have shown that countries with higher levels of education tend to have lower levels of income inequality.
For example, Initiatives like the Rashtriya Uchchatar Shiksha Abhiyan (RUSA) aim to improve the quality of higher education and increase enrollment rates, particularly among marginalised groups and rural populations. By investing in higher education, the government seeks to equip individuals with the skills and knowledge needed to access better-paying jobs and improve their economic prospects.

The income distribution can be significantly impacted by government policies. Tax policies can be used to redistribute the income from wealthy to poor. Initiating Progressive Taxation, where the rich are taxed at a higher rate than the poor, can help to reduce income inequality.

Likewisely, Social welfare policies, such as unemployment benefits and healthcare, can provide a safety net for the poor and reduce income inequality.
However, there are some government policies and actions which can make income inequality even greater or more severe.
For example, some policies that favour the wealthy, such as tax breaks for the rich, can widen the income gap. Similarly, policies that reduce social welfare spending can increase poverty and income inequality. Therefore, it is essential to analyse government policies to determine their impact on income distribution. Policies that promote social

justice and equality should be prioritised, while policies that favour the wealthy should be reevaluated.

Also, Globalisation and technological change have had a significant impact on income distribution. Globalisation has led to increased competition in the labour market, which has put pressure on wages and thus increased income inequality.
Similarly, technological change has led to the displacement of workers in certain industries, leading to job losses and income inequality. However, labour market institutions can play a crucial role in reducing income inequality. Strong labour unions, for example, can negotiate and demand higher wages and better working conditions for workers, thereby reducing income inequality. Similarly, minimum wage laws can ensure that workers are paid a fair wage, reducing poverty and income inequality.

Addressing income inequality requires a multifaceted approach that combines policies to promote inclusive growth, such as investments in education, healthcare, and social protection, with measures to enhance labour market efficiency, ensure fair taxation, and reduce barriers to economic participation. By fostering greater equality of opportunity and redistributing resources to those in need, societies can not only mitigate the adverse effects of income inequality but also unleash the full potential of their human capital, driving sustainable economic development and shared prosperity.

1.3 Statement of problem:

This study aims to address the following research question:
What is the relationship between income inequalities and economic development? Specifically, how do variations in income distribution impact key economic indicators such as GDP growth, poverty rates, and human capital accumulation?

1.4 Importance of study:

Understanding the nexus between income inequality and economic development is necessary for formulating effective policy interventions aimed at promoting inclusive growth and reducing poverty. Highlighting the socio-economics impact of economic inequality. Explain why understanding the relationship between income and economic development is crucial. Emphasise how disparities in income distribution can affect key macroeconomics indicators such as GDP growth , investment and productivity. Introducing key concepts such as Gini Coefficient or measures of wealth distribution . This research contributes to the ongoing debate on equitable development strategies.

1.5 OUTLINE OF THE PAPER:

2. Literature Review
   2.1 Overview of Existing Research

2.Literature Review:

2.1 Overview of Existing Research:

Exploring the Relationship between Income Inequality and Economic Development: An Analysis of Global Trends and Regional Disparities.

Income inequality is a pervasive issue that impacts economic development on a global scale. Over the past few decades, there has been a noticeable trend towards an increasing concentration of income at the top of the income distribution worldwide, leading to significant disparities in economic progress . This trend is further compounded by the interplay between within-country and between-country inequality, along with the varying growth rates of emerging and advanced economies, which have influenced the dynamics of the middle class globally since 1980 . The concept of the "elephant curve" visually represents this phenomenon, showcasing a growing middle class in emerging economies juxtaposed with a squeezed middle class in wealthier nations, underscoring the profound impact of income inequality on economic development. Notably, emerging economies like China and India have made strides in narrowing the income gap with advanced economies, while within-country inequality has been on the rise in recent years, accounting for a significant portion of global inequality in 2020 compared to 1980. This disparity in income distribution not only impedes individual progress but also poses a threat to the economic stability and growth of developing nations, emphasising the urgency for inclusive policies and social programs to address income inequality for sustainable development .

Income inequality is a multifaceted phenomenon influenced by various factors related to economic development. Moller et al. Birčiaková et al. (2013) suggest that income inequality is intricately linked to economic development, educational expansion, racial and ethnic structure, urbanisation, as well as political and institutional factors. Kuznets' work Huynh et al. (2023) highlighted that economic development initially widens income inequality but eventually leads to its reduction as the economy progresses. This concept is further supported by Paukert's findings

Greenwood & Jovanovic (1990) that intra-country income inequality tends to rise and then decline with economic development.

Moreover, the relationship between economic freedom and income inequality has been extensively studied. Apergis & Cooray (2015) discuss how economic freedom can impact income inequality, indicating a trade-off between economic freedom and income equality. Similarly, Bergh & Bjørnskov (2021) delve into the effects of economic freedom on growth and income equality, building on earlier research. Additionally, 's empirical note Carter (2006) emphasises a positive but relatively inelastic relationship between economic freedom and income inequality.

Furthermore, the impact of tourism on income inequality has been explored by (Subramaniam et al., 2022), who reference the Kuznets curve to explain the inverted-U relationship between income and income inequality during economic development. Fang et al. (2020) also contribute to this discussion by highlighting the positive impact of tourism growth on income inequality in both developing and developed economies.

In conclusion, income equality and economic development are intricately intertwined, with various factors such as economic freedom, educational expansion, and tourism playing significant roles in shaping income distribution patterns within societies. Understanding these relationships is crucial for policymakers aiming to foster sustainable economic growth while addressing income inequality.

https://doi.org/10.1108/jbsed-07-2021-0102

2.2Key Concepts and Definitions:

We will discuss some concepts which give us a good understanding of income inequality.
They provide valuable insight into distribution of income with a population, highlighting disparities in wealth distribution.

GINI COEFFICIENT,Italian statistician Corrado Gini (1884–1965) created the GINI COEFFICIENT, which bears his name. GINI COEFFICIENT also referred as GINI INDEX is the most commonly used measure of inequality.
It is used as a measure of *income* inequality, but it can be used to measure the inequality of any distribution like the distribution of wealth.

It measures inequality on a scale from 0 to 1, where higher values indicate higher inequality. This can sometimes be shown as a percentage from 0 to 100%, this is then called the ‘Gini Index’ Perfect inequality is denoted by a value of 1, implying that no one gets any money and only one person gets all of it.This method can be used to tell us the difference we expect to find between the income of any two individuals related to the mean.

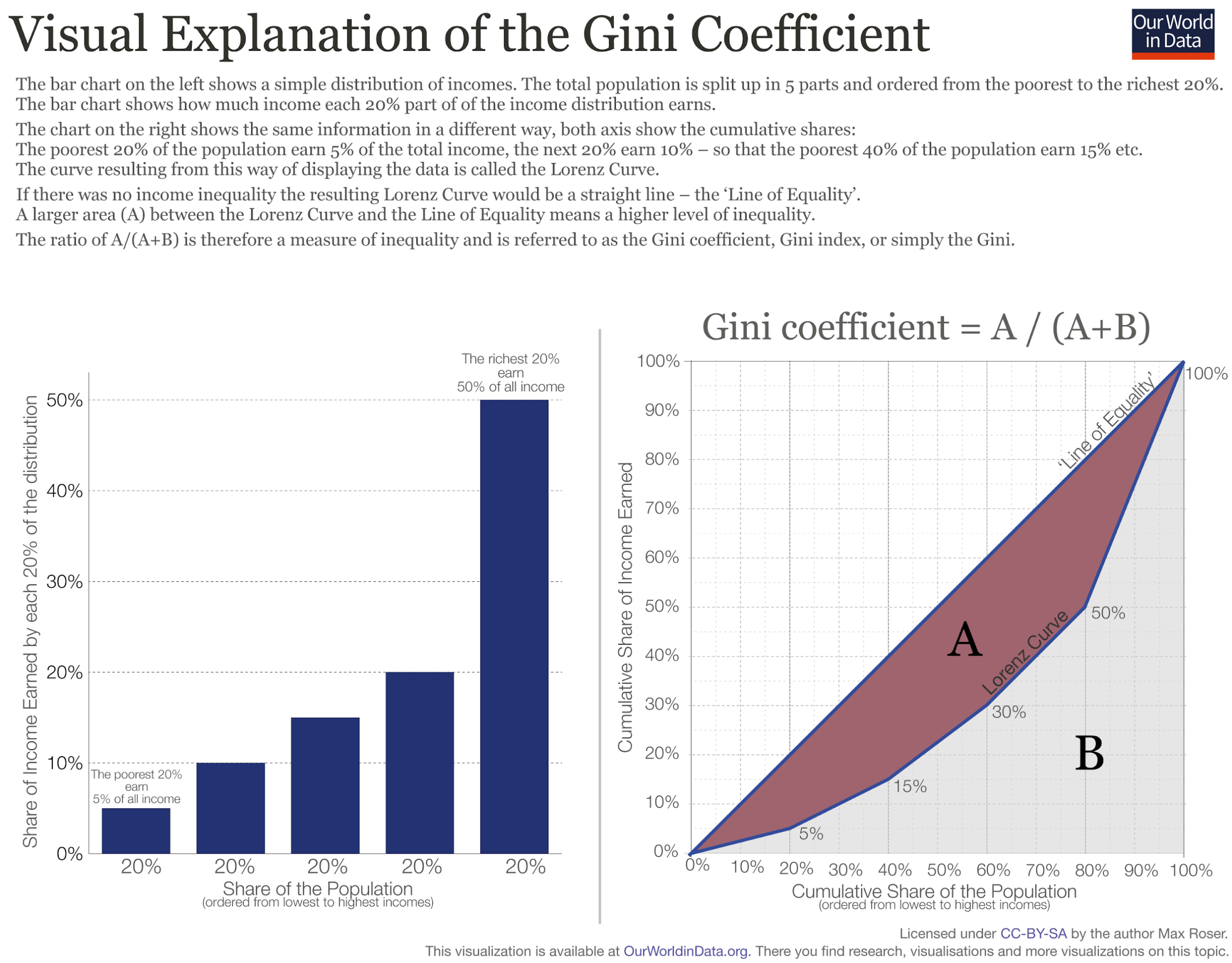
We will perform an experiment to understand this.

Imagine two strangers meeting at random on the street.They compare their incomes and find out how much richer one person is than the other.How much of a difference would we expect?

This expected gap between two randomly chosen people is what the Gini coefficient measures.It is calculated by summing the gaps between each pair of individuals.The difference in income between two randomly chosen people should be minimal in a situation where incomes are allocated equally. Where inequality is high, we would expect the gap to be large.If measured in absolute terms, however, this will also depend on how rich or poor the population is generally. The absolute difference in earnings cannot be large in a society where even the wealthiest individuals in society live in poverty. Conversely, where incomes are generally high, even very small relative differences between people's incomes can result in large absolute gaps.

Because of this, the predicted absolute difference between an individual's income and the population mean income is expressed by the Gini coefficient.Specifically, it is computed as the anticipated gap as a share of twice the mean income. Twice the mean income is the highest possible value for the average gap – a situation of perfect inequality, where one person has all the income and everyone else has none. So in this case of maximum inequality, the Gini coefficient is 1.

Since everyone makes the same amount of money, there can never be a scenario of complete equality, where the average disparity between any two pairs of individuals is equal to zero. The Gini coefficient in this instance is 0.The figure illustrates a second, visual definition of the Gini coefficient.



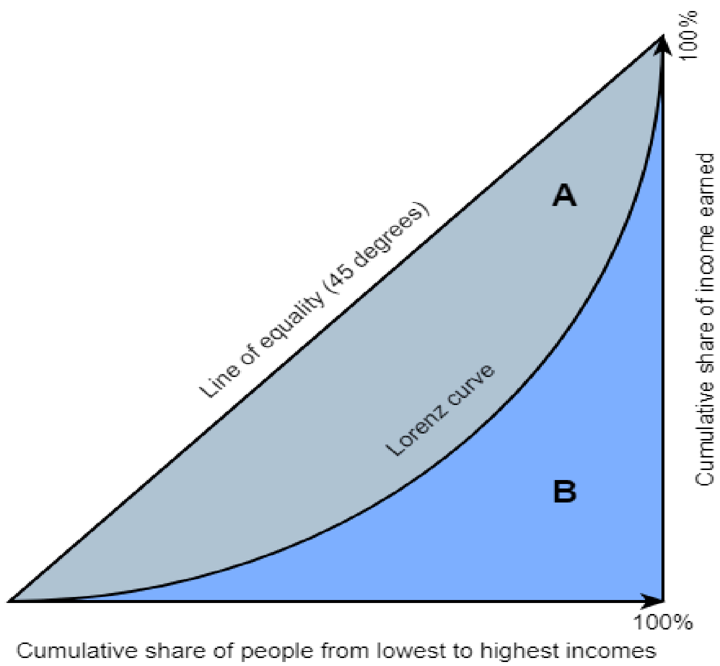
The left panel shows the share of income received by each fifth of a hypothetical population.The cumulative plot of this data is displayed in the right panel.This is known as a 'Lorenz curve'.

In a population where income is shared perfectly equally, the Lorenz curve would be a straight diagonal line: 10% of the population would earn 10% of the total income, 20% would earn 20% of the total income, and so on. This is shown in the chart as the 'line of equality'.

In the hypothetical population shown in the chart, though, incomes are not distributed equally.

INCOME INEQUALITY AND LORENZ CURVE

The LORENZ CURVE was taken up by the American statistician MAX LORENZ to study the income problem.The number of consumer units in each income group and their income are assessed initially in the Lorenz curve approach. The consumer units are then ranked from lowest to greatest income by dividing them into percentages. The group's % contribution to the national income is displayed opposite each percentage. Ultimately, a table is created with the consumer units and the national income as percentages, and the information is plotted in a diagram. The Lorenz curve is the diagrammatic expression that emerges from this procedure. The percentage of consumer units arranged cumulatively from the lowest income to the highest income on the horizontal axis is displayed on this curve If the percentages that different groups receive from income are marked on the diagram and the marked points are connected, a curve showing the income distribution of the economy is obtained. In the curve diagram showing this income distribution, the inequality in income distribution decreases as it approaches the line drawn by forming an angle of 45 degrees, and the inequality in income distribution increases as it gets farther away.



We have another conceptual theory to discuss **<u>"TOP 1% VS BOTTOM 99%"</u>** ***"We are the 99%"*** is the slogan of the Anti-Wall Street campaign in the US. This slogan is highlighting the sheer inequality of income and wealth which exists in America. According to the papers, the top 1% of income holders in the US receive over 20% of all income and own about 35% of all America's wealth in 2007, the top 1% of wealth holders had net worth of nearly $22trn, or 33.8% of all net wealth in America ($65trn). In addition, he calculated that the richest 1% of earners collected 21.4% of all income in 2006, or $2 trillion annually out of a total of $9 trillion. If we add in the next 4% of wealth holders, then the top 5% have over 60% of all wealth and 37% of all income.

It is increasingly fashionable to argue that inequality is the cause of the crisis in capitalism.The cause of recessions and instability in capitalism in the 1970s was not assigned to inequality of income or wealth. Indeed, many mainstream and heterodox economists argued the opposite, namely that it was caused by wages rising to squeeze profits in overall national income.

But now, many Marxist economists argue that this current crisis is a product of wages being too low and profits too high. This leads to low wage earners being forced to borrow more and thus eventually causing a credit crisis. So it seems that the underlying cause of the capitalist crisis can vary. This eclectic approach has the drawback of making it difficult to determine what is causing the economic crisis—is it wages driving down profits, as they did in the 1970s or is it low wages leading to a collapse of demand in the noughties?

Inequality of wealth and income may not be the cause of capitalist crises, but it is certainly a product of capitalism.The private profit society is increasingly inefficient in delivering our needs and increasingly unstable. However, it's always been unfair and uneven. Actually, inequality of incomes and wealth in a society is a feature of all class societies, whether it is slavery, Asian absolutism, priestly castes, feudalism or capitalism. By definition, inequality accompanies a class society. After all, why would anybody want to be a member of the ruling elite if they did not enjoy the fruits, namely extreme wealth and income as well as power and status. Indeed, even priestly castes, supposedly engaged in waiting for the rewards of the 'afterlife' and telling their flock that they must also wait, were not slow in coming forward to benefit from the material life? Just consider the Vatican and other Christian cathedrals, the Middle Eastern mosques, the synagogues, and even further back to the ancient priesthood kings of Egypt, of the Aztecs and Incas of South America. Class society means that the ruling class controls the surplus generated by the labour time of the non-rulers, in whatever form. In other words, inequality is not a unique characteristic of capitalism that explains its ongoing booms and busts, but rather a result of all class societies.

How unequal is wealth and income distribution currently around the world (not just the US)? Well, Branco Milanovic at the World Bank has studied global inequality over the years carefully in a series of books and papers. His measure of inequality is not a measure between nations but between individuals, even though inequality is decided more by which country you live in than by inequality within a country. Global elites (the top 10%) take 57% of all the income created in the world and this ratio has hardly moved in 100 years. As Milanovic explains, this flies in the face of the predictions of mainstream economics that argue inequality should decline as countries get richer.

Simon Kuznets developed what is called the Kuznets curve in the mid-1950s.This supported the theory that inequality is low in pre-industrial societies where poverty is universal. When industrialisation takes place, as in the UK in the early 19th century and in the rest of major capitalist economies later in that century, then inequality grows between the urban and rural populations. Then, as an economy matures, the urban-rural gap narrows and the welfare state kicks in and inequality falls. So the Kuznets curve of inequality is an upside down U. The other mainstream theory is that, as global trade expands, the demand for low-skill labour in poor countries rises and so their incomes rise relative to those in mature economies. Inequality should decline.

Well, the evidence refutes both those theories. Global inequality of income – i.e. between world citizens – is high and has stayed high. Globally, the Gini coefficient, which calculates the ratio of wealth or income possessed by cohorts of people, has stayed extremely high, hovering around 70. It has risen in the US and the UK over the last 30 years and it has also jumped in China. So the richest 1% of the world's population now receive nearly 15% of all the world's income, while the poorest 20% receive only just over 1%!

Almost all of the main capitalist economies have seen an increase in income inequality during the 1980s, but this has been especially true of the so-called Anglo-Saxon, deregulated "free market" nations like the US and the UK (the top 1% had about 10% in the early 1980s and now have nearer 20%). In Continental Europe, inequality has been static – the top 1% had about 10% in the 1980s and now have much the same ratio. Between 1993 and 2008, average real incomes in the US increased at a 1.3% annual pace; however, if the top 1% is taken out of the equation, the growth rate was only 0.75%. The incomes of the top 1% achieved a 3.9% a year rise, taking 58% of all the increase in real incomes between 1993 and 2008. Indeed, in the great US booms of 1993-2000 and 2001-2007, the top 1% achieved annual growth in their incomes of over 10%, while the bottom 99% achieved only 1.3-2.7% a year.

The top 10% of income earners in the US, with incomes over $110,00 a year, now receive half of all the income each year. The top 1% with incomes over $400,000 a year received nearly 25% of all income in 2007. This 1% constitutes 1.5m earners out of 150m earners in America. One of the stats that the authors reveal is that, in the Great Recession, the bottom 99% of income earners in the US suffered a 7% fall in their incomes, the largest drop since the Great Depression. The other discovery is that it is the top 1% and even more, the top 0.1% of earners, who gained the most, compared with the top 5%. The inequalities are rising within the rich to create a super-rich elite. Whereas the top 1% took 9% of incomes in 1978 and the top 5% took the next 12% , to make 21% in all; by 2008, the top 1% took 21% while the top 5% took another 15% (making 36% in all). So the top 1% have reaped the biggest gains in the last 30 years. Also the top 0.1% took only 1% of incomes in 1978 but now take 6%!

The main reason for this higher concentration is the massive rise in the incomes going to the very top chief executives in the banks and big corporations. It's the same story in the UK. The UK's High Pay Commission found that bosses' salaries rose by 63% since 2002, but total pay packages for top company executives had gone up by 700% since 2002. In contrast, pay levels for the average worker in Britain rose only 27% before inflation and taxes.

These extremes of inequality and the particularly fast rise in that inequality in the last 30 years is now beginning to worry the strategists of capital in an environment of depression in the mature capitalist economies. It is not that they think inequality causes crises (although some are beginning to argue that), it is because they fear a social backlash from the 99% towards the 1%. And how justified that concern seems to be now, considering the worldwide movement emerging against the ultra-wealthy. "My worst fears about the potential loss of confidence in our leaders, institutions, and capitalism itself are realising," expressed City economist Jeremy Grantham. This is a hole that we have been digging for a while. We ought to be sincere in our endeavours to revive the fortunes of the typical labourer. Isn't it preferable for us to consciously choose the income distribution that strikes the optimal balance between social justice and work incentives on our own? *How about going back to the levels of income equality that existed under the Presidency of that notable Pinko, Dwight Eisenhower? And don't think for a second that this more equal income distribution somehow interfered with economic growth: the 50s and 60s were the heyday of sustained U.S. economic gains."*

2.3 Theoretical Framework:

Measuring income inequality is critical for several reasons, spanning social, economic, and political realms. Understanding the distribution of income within a society can help inform policies, highlight social issues, and guide economic decisions
Here are some key reasons why measuring income inequality is important:

1.Informing Public Policy and Redistribution:
Measuring income inequality helps policymakers understand the extent of disparities in income distribution. This knowledge can drive the creation of targeted economic policies, such as progressive taxation, welfare programs, and social security measures aimed at redistributing income more equitably across a society.

2.Assessing Economic Health and Stability:
High levels of income inequality can indicate potential instability within an economic system. Economies with severe inequality might experience less effective demand, as a significant portion of the population might not have sufficient income to consume goods and services. This can lead to slower growth and business cycles characterised by boom and bust phases.

3.Promoting Social Cohesion and Stability:

Large income disparities can lead to social unrest and tension. Communities with less inequality generally report higher levels of social cohesion and trust, which are vital for the stable and peaceful functioning of society. Conversely, high inequality can erode trust in institutions and increase the risk of social unrest.

4.Encouraging Sustainable Development:
Addressing income inequality is key to sustainable development. The United Nations' Sustainable Development Goals (SDGs), particularly Goal 10, focus on reducing inequality within and among countries. Sustainable development is more achievable in environments where economic resources are distributed fairly, enabling broader access to health, education, and other essential services.

5.Understanding Economic Mobility:
Measuring inequality helps in analysing economic mobility, or the ability for individuals or families to move between economic brackets. High inequality often corresponds with low mobility, suggesting that it is harder for people born into poverty to improve their economic status. This can perpetuate cycles of poverty.

6.Guiding International Comparisons and Investment Decisions:
Investors and businesses use measures of income inequality, like the Gini coefficient, to assess country risk and to make decisions about where to allocate resources. Countries with high inequality might be seen as riskier or less stable, potentially affecting foreign investment decisions.

7.Highlighting Gender and Ethnic Disparities:

Income inequality metrics can also be used to identify and address income disparities between different genders and ethnic or racial groups. This aids in crafting specific interventions aimed at promoting equality across different segments of the population.

In conclusion, measuring income inequality provides essential insights that help shape effective governance, promote equitable growth, and ensure social harmony.

INCOME SEGREGATION AND INCOME INEQUALITY:
INCOME SEGREGATION refers to the spatial separation of households with different income levels within a given area. This phenomenon occurs when people of similar income level cluster together in certain areas, leading to neighbourhoods that are predominantly composed of either high income or low income residents.
Contrarily, income inequality describes the unequal distribution of income across people or families within a society. It's a measure of how much wealth and income are concentrated in the hands of a few compared to the rest of the population .

The prerequisite for income segregation is INCOME INEQUALITY. If there were no income inequality, neighbourhoods would possess the same opportunities and conditions and thus no income segregation would likely occur. However, it is important to recognize that the relationship can be affected by many different factors and that it is not necessarily a one-to-one relationship.

According to studies, the existence of residential choices that are connected with income rather than income inequality alone leads to income segregation; hence, an income-based housing market and/or housing policy are essential.

Income segregation increased in the US between 1970 and 1990 as a result of rising inequality. Also, an increase in income segregation in economic school segregation is highly dependent on income inequality (as well as the role of educational policies).Income segregation (as well as racial segregation) is stronger in the public educational system than in the private system, especially in primary education. This is usually explained by the neighbourhood (metropolitan area) where schools are located. This can be due to the fact that regardless of one's financial background, anyone can attend a public school; therefore, a broader range of people will study at a public school. The relationship between income segregation and income inequality is complex: its results can be influenced by local political, economic, and city planning agendas, which can either work for or against the fight with income segregation and income inequality

Income segregation and income inequality are closely related and often reinforce each other in a vicious cycle. Here's how:

<u>Spatial concentration of wealth:</u>

Income segregation leads to the spatial concentration of wealth in certain areas. High-income individuals tend to live in neighbourhoods with better amenities, schools, and services, while low-income individuals are often relegated to areas with fewer resources and opportunities. This concentration of wealth exacerbates income inequality by concentrating economic resources in the hands of a few.

<u>Limited Social Mobility:</u>

Income segregation can also limit social mobility, the ability of individuals to move up or down the income ladder. When low-income individuals are isolated in neighbourhoods with limited economic opportunities, they may face barriers to accessing quality education, jobs, healthcare, and other essential services. This perpetuates a cycle of poverty and prevents upward mobility, contributing to persistent income inequality.

<u>Unequal Access to Opportunities:</u> Neighbourhoods with high levels of income segregation often have unequal access to opportunities. For example, affluent neighbourhoods may have better-funded schools, safer environments, and more job opportunities, while disadvantaged neighbourhoods may lack essential resources and infrastructure. This unequal distribution of opportunities further widens the gap between the rich and the poor.

<u>Political and Economic Segregation:</u> Income segregation can also lead to political and economic segregation, where affluent communities have more influence over local policies and decision-making processes. This can result in policies that further exacerbate income inequality, such as tax breaks for the wealthy or cuts to social welfare programs that benefit low-income individuals.

3.Methodology

3.1 Empirical topics in income inequality

Exploring empirical topics in income inequality can provide valuable insights into economic disparities, their causes, and potential solutions.
Here are several compelling empirical topics that you could consider for a research paper:
Impact of Globalization on Income Inequality: Investigate how global economic integration has affected income distribution within and between countries. Analyse trade liberalisation, foreign direct investment, and the outsourcing of jobs to lower-wage countries.

Technology and Wage Polarisation: Examine how technological advancements, particularly in AI and automation, have impacted jobs and wages across different sectors. Consider how technology-driven changes have contributed to the expansion of high-wage and low-wage employment, while diminishing middle-wage jobs.

Effects of Tax Policies on Income Inequality: Study the role of progressive versus regressive tax systems and how different tax policies affect income distribution. Include analysis of capital gains taxes, income taxes, and corporate taxes.

Educational Attainment and Income Inequality: Explore the relationship between education levels and income disparity. Assess how educational opportunities are distributed across different socio-economic groups and how this affects income potential.

Health Disparities and Income Inequality: Investigate how differences in income levels correlate with health outcomes and access to healthcare services. Explore the impact of public health policies and insurance systems on these disparities.

Gender and Racial Income Disparities: Focus on how income inequality manifests across different genders and races.Analyse wage gaps, employment opportunities, and the glass ceiling in various industries.

Income Mobility and Inequality: Study the dynamics of income mobility across generations. Examine factors that contribute to or hinder upward mobility, such as education, family background, and social networks.

The Impact of Minimum Wage on Income Inequality: Analyse empirical data to assess how changes in the minimum wage have impacted income distribution and poverty rates. Consider different regional economic contexts for a nuanced analysis.

Gig Economy and Income Inequality: Examine the role of the gig economy in income inequality, focusing on job security, wage levels, and worker rights. Consider how gig work compares to traditional employment in terms of income stability and social benefits.

Wealth Concentration and Economic Power: Explore how wealth concentration among the ultra-rich affects economic decision-making and policy influence. Study the implications of such concentration for average citizens and overall economic health.

3.2 Problems in analysing income inequality

Analysing income inequality presents several challenges, due to the complexity and multifaceted nature of the factors that influence it. Some key problems encountered in such analyses include:

1. Data Availability and Quality

Incompleteness: In many regions, especially in less developed countries, income data can be sparse, irregularly collected, or non-existent.

Inaccuracy: Income is often underreported, particularly at the extremes of the wealth spectrum. Wealthy individuals may hide income to avoid taxes, while low-income groups may not be captured accurately in surveys.

Inconsistency: Different countries may have different standards and methods of data collection, making comparisons challenging.

2. Complex Causality and Interdependencies

Multiple Factors: Income inequality is influenced by a complex interplay of economic, political, social, and cultural factors. Disentangling the effects of individual variables can be challenging.

Feedback Loops: Changes in income inequality can feed back into the factors that affect it, such as education and political power, creating cycles that are difficult to break and analyse.

3. Measurement Challenges

Choosing Metrics: There are various metrics to measure income inequality, such as the Gini coefficient, the Palma ratio, and income shares across percentiles. Each metric offers different insights and may tell different stories about the level of inequality.

Temporal Dynamics: Inequality may change over different time scales. Short-term fluctuations might mask long-term trends, requiring careful analysis to understand underlying movements.

4. Global vs. Local Dynamics

Global Influences: Global economic trends, such as international trade policies or financial crises, can impact local income distributions. Analysing these effects requires understanding both local contexts and global dynamics.

Regional Variations: Inequality must often be analysed both at a national and a local level, as economic conditions can vary greatly within countries.

5. Theoretical and Ideological Biases

Economic Models: Different economic theories provide different explanations and solutions for inequality, which can influence the interpretation of data.

Political Biases: Research into inequality can be politically charged. Studies might be skewed by the ideological positions of those funding or conducting the research.

6. Ethical and Privacy Concerns

Sensitive Data: Gathering detailed income data often raises privacy concerns and ethical issues, potentially limiting the scope of research.

Stigmatisation: Reporting on income inequality risks stigmatising certain groups or regions, which must be handled with sensitivity to avoid exacerbating social tensions.

Overall, analysing income inequality requires a nuanced approach that considers a variety of data sources, acknowledges theoretical and methodological limitations, and remains alert to the broader socio-political implications of the findings.

3.3 statistical Methods

Measuring inequality is crucial in economics, social sciences, public policy, and health studies to understand how resources such as income, wealth, or opportunities are distributed among individuals or groups within a society. Several statistical methods are used to measure and analyse inequality:

Gini Coefficient:

The Gini coefficient is one of the most widely used measures of inequality. It calculates the extent to which the distribution of income (or wealth) among individuals or households deviates from a perfectly equal distribution. A Gini index of 0 represents perfect equality, while a Gini index of 1 indicates maximum inequality.

The coefficient is often calculated based on the Lorenz curve, which plots the cumulative percentages of total income received against the cumulative number of recipients, starting with the poorest individual or household.

Let us understand the formula for us to find the value through the Gini coefficient calculator from the explanation below.

Gini Coefficient Formula

$$\text{Gini Coefficient} = \frac{A}{A+B}$$

$$\text{Gini Coefficient} = 1 - \text{Sum}$$

If A=0, the Lorenz curve is the line of equality. When A=0, the Gini index is 0. The Gini coefficient is large in case A is very large and B is small. It indicates there is huge income/wealth inequality.

How To Calculate?

Gini coefficient countries take the following steps to find the value that indicates the inequality among their citizens' incomes is as follows:

1.Organise the data into a table with the category head mentioned below.

Fraction of Income	Fraction of Population	% of Population that is richer	Score

It is important to note that all the rows must organise from the poorest to the richest. For instance, if it states that the bottom 10% of the population earns 3% income, write 0.03 in the 'Fraction of Income' column. Next, write 0.10 in the 'Fraction of Population' column. Similarly, fill these two columns with other percentages given.

2.Fill '% of Population that is richer' column by adding all terms in 'Fraction of Population' below that row.

Fraction of Income	Fraction of Population	% of Population that is richer	Score
0.03	0.1	0.9	
0.15	0.5	0.4	
0.85	0.4	0	

For instance, we fill the first row in the '% of Population that is richer' column. Then, we will add 0.50 and 0.40, the rows in 'Fraction of population' below it. Hence, we get 0.90.

3.Calculate the Score for each of the rows. The formula for Score is:

Score = Fraction of Income * (Fraction of Population + 2 * % of Population that is richer).

For instance, score for the 1st row is 0.03*(0.10+2*0.90) = 0.057

4. Next, add all the terms in the 'Score' column. Let us call it 'Sum.'
Calculate the Gini coefficient using the formula: = 1 – Sum

Example #1

The Gini coefficient of 2 countries based on citizens' income is as under.

	2010	2011	2012	2013	2014	2015
Country A	0.4	0.43	0.48	0.5	0.52	0.57
Country B	0.38	0.37	0.35	0.34	0.3	0.29

- **Interpret the trend of income inequality in the two countries**
- **Which country has higher income inequality in 2015?**

Solution:

a) The Gini coefficient of Country A has shown a rising trend from 0.40 in 2010 to 0.57 in 2015. Hence, income inequality in Country A has risen in these years. On the other hand, the coefficient of Country B has fallen from 0.38 in 2010 to 0.29 in 2015. Therefore, income inequality in Country B has declined over these years.

b) The coefficient of Country A (0.57) is more than that of Country B (0.29). Hence, Country A had higher income inequality in 2015.

Example #2

In a particular country, the lowest 10% of the earners make 2% of all wages. The next 40% of earners make 13% of wages. The following 40% of earners make up 45% of all wages. Moreover, the highest 10% of all earners make 40% of all wages. Calculate the Gini coefficient of the country.

Solution:

Use the following data for the calculation.

Fraction of Income	Fraction of Population
0.02	0.1
0.13	0.4
0.45	0.4
0.4	0.1

Let us compile the above information in table format. But, first, the information has to be compiled by organizing the rows from the poorest to the richest.

Fraction of Income	Fraction of Population	% of Population that is Richer	Score	Score Formula
0.02	0.1	0.9	0.038	=A2*(B2+2*C2)
0.13	0.4	0.5	0.182	=A3*(B3+2*C3)
0.45	0.4	0.1	0.27	=A4*(B4+2*C4)
0.4	0.1	0	0.04	=A5*(B5+2*C5)

=1-D7

Gini Coefficient = 1 - Sum

Fraction of Income	Fraction of Population	% of Population that is Richer	Score	Score Formula
0.02	0.1	0.9	0.038	=A2*(B2+2*C2)
0.13	0.4	0.5	0.182	=A3*(B3+2*C3)
0.45	0.4	0.1	0.27	=A4*(B4+2*C4)
0.4	0.1	0	0.04	=A5*(B5+2*C5)
Sum of Scores			0.53	
Gini Coefficient			=1-D7	

Sum of Scores = 0.038+0.182+0.27+0.04 =0.53

The coefficient will be –

D8 =1-D7

Fraction of Income	Fraction of Population	% of Population that is Richer	Score	Score Formula
0.02	0.1	0.9	0.038	=A2*(B2+2*C2)
0.13	0.4	0.5	0.182	=A3*(B3+2*C3)
0.45	0.4	0.1	0.27	=A4*(B4+2*C4)
0.4	0.1	0	0.04	=A5*(B5+2*C5)
Sum of Scores			0.53	
Gini Coefficient			0.47	

Frequently Asked Questions (FAQs)

What does a low Gini coefficient mean?
A Gini coefficient of zero indicates there is equal income distribution. Whereas a number near one shows greater inequality. Thus, the lower the Gini coefficient, the more similar the society.

What are the Lorenz curve and Gini coefficients?
The Lorenz curve is the focal point in determining the Gini coefficient, a mathematical portrayal of inequality levels. Still, since Lorenz curves are mathematical calculations depending on fitting a continuous curve to incomplete and discontinuous data, they may be imperfect evaluations of true inequality.

What is the U.S.e US Gini coefficient?
In 2021, as per the Gini coefficient, household income distribution in the United States was 0.49. In addition, in 1990, this figure was at 0.43, which shows an increase in income inequality in the U.S. over the last 30 years.

What does a negative Gini coefficient mean?
The Gini coefficient can range from 0, complete equality, to 1 (absolute inequality). Often, it is expressed as a percentage ranging between 0 and 100. However, if negative values are possible, the Gini coefficient can be more than 1.

Lorenz Curve:

This is a graphical representation of the distribution of income or wealth. The curve shows the proportion of total income earned by the bottom x% of the population. The further the curve is from the diagonal line (the line of equality), the greater the inequality.

Example #1

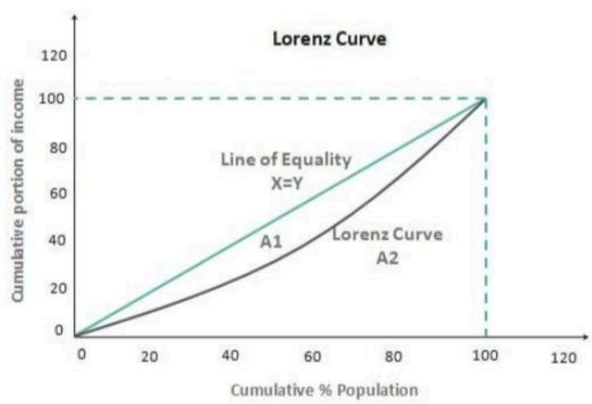
Let us consider an economy with the following population and income statistics: –

Population	Income Portion %
0	0
20	10
40	20
60	35
80	60
100	100

And for the line of perfect equality, let us consider this table: –

Population	Income Portion %
0	0
20	20
40	40
80	80
100	100

Let us now see how a graph for this data looks: –



As we can see, there are two lines in the graph of the Lorenz curve, the curved red line, and the straight black line. The black line represents the imaginary line of equality, i.e., the ideal graph when income or wealth is equally distributed amongst the population. But on the other hand, the red curve, the Lorenz curve, which we have been discussing, represents the distribution of wealth among the people.

Hence, we can say that the Lorenz curve is the graphical method of studying dispersion. Gini coefficient, also known as the Gini index, can be computed as follows. Let us assume in the graph area that A1 represents the Lorenz curve and the line, and A2 is the line below the curve. So,

Gini coefficient = A1/ (A1+ A2)

The Gini coefficient lies between 0 and 1, 0 being the instance where there is perfect equality and 1 being the instance where there is perfect inequality. Therefore, the higher the area enclosed between the two lines represents higher economic inequality.

We can say that there are two indicators for measuring income inequality:
- The Lorenz curve is the visual indicator
- The Gini coefficient is a mathematical indicator

Income inequality is a pressing issue across the world. So,**what are the reasons for inequality in an economy ?**
Corruption
Education
Tax
Gender differences
Culture
Race and caste discrimination
The difference in preferences of leisure and risks.

Reasons for income inequality

1.The distribution of economic characteristics across the population should be considered.

2. Analysing how the differences give rise to different outcomes in terms of income.

3. A country may have a high degree of inequality because of:

i A great disparity in these characteristics across the population.

ii These characteristics generate huge effects on the amount of income a person earns.

4. ANALYSIS AND RESULT

4.1Descriptive Statistics

Table 1 reports the descriptive statistics for the variables. Middle-income countries show the higher average income with Gini net income of 42.06 and Gini market income of 47.14 and the higher Standard deviation of Gini net income of 8.60 and Gini market income of 8.17 in comparison to high and low-income countries. High-income countries provide higher GDP per capita of 32576.59 and mean GDP per capita growth of 1.90 with a standard deviation of 17476.96 and 3.07 in comparison with middle and low-income countries. Nonetheless, high-income countries show a higher average financial development index of 0.54 with a standard deviation of 0.20 as compared to the middle and low-income countries.

Variables	Mean	Std Dev	Minm	Maxm
Panel A: High-Income Countries				
Gini Net (GINIn)	30.70	6.69	16.60	51.41
Gini Market (GINIm)	46.37	4.67	29.96	49.97
Per capita GDP (ED)	32576.59	17476.96	4266.11	91593.63
Per capita GDP growth(g)	1.90	3.07	−14.55	13.08
Financial Development Index (FD)	0.54	0.20	0.10	1.00
Dependency Ratio (AGE)	50.79	4.53	40.31	70.82
Govt. Expenditure (EXP)	19.59	4.10	9.95	36.26
Inflation (INF)	5.69	13.18	−9.68	259.99
Trade (TRD)	75.00	34.31	16.57	188.97
Urban Population (POP)	77.55	11.12	32.06	97.77
Panel B: Middle-Income Countries				
Gini Net (GINIn)	42.06	8.60	18.07	67.21
Gini Market (GINIm)	47.14	8.17	23.23	76.88
Per capita GDP (ED)	3847.82	2521.39	331.97	11797.45
Per capita GDP growth(g)	2.51	4.97	−30.71	33.03
Financial Development Index (FD)	0.24	0.13	0.002	0.68
Dependency Ratio (AGE)	64.12	14.29	35.33	102.61
Govt. Expenditure (EXP)	13.78	5.26	2.97	30.12
Inflation (INF)	65.45	358.67	−26.29	6261.24
Trade (TRD)	74.47	40.03	12.008	220.40
Urban Population (POP)	53.95	18.19	16.946	91.45
Panel C: Low-Income Countries				
Gini Net (GINIn)	41.10	6.05	21.85	61.46
Gini Market (GINIm)	47.35	7.00	25.42	70.93
Per capita GDP (ED)	542.96	355.91	202.43	2363.67
Per capita GDP growth(g)	1.47	6.00	−47.72	37.12
Financial Development Index (FD)	0.08	0.02	0.02	0.19
Dependency Ratio (AGE)	94.55	7.70	72.84	110.92
Govt. Expenditure (EXP)	13.00	4.33	4.83	31.55
Inflation (INF)	14.79	20.07	−10.73	128.76
Trade (TRD)	49.31	14.13	19.68	91.37
Urban Population (POP)	23.88	10.47	5.05	44.36

Note: Std Dev indicates the STANDARD DEVIATION , Minm denotes the minimum and Maximum represents the maximum

Finally, other control variables including Dependency Ratio, Govt. Expenditure, Inflation, Trade and Urban Population present considerable variations among high, middle and low-income countries. Low-income countries show a higher rate of Dependency Ratio with a mean value of 94.55 and a standard deviation of 7.70 than the high income and middle-income countries. Inflation shows the significant variations in middle-income countries with an average value of 65.45 and a standard deviation of 358.67 than in high income and low-income countries. Moreover, Government expenditure (average of 19.59 and standard deviation of 4.10), Trade (average of 75.00 and standard deviation of 34.31) and Urban population (average of 77.55 and standard deviation of 11.12) show comparatively higher deviations in high-income countries than in middle and low-income countries.

4.2Regression Analysis

Static linear regression models

Assume the standard regression model

$$y_i = \boldsymbol{x}_i\boldsymbol{\beta} + u_i$$

for $i = 1, \ldots, N$, where y_i and u_i are scalars, $\boldsymbol{x}_i$ is $1 \times K$ and $\boldsymbol{\beta}$ is $K \times 1$. Note that the above expression defines a generic regression model, so y_i does not need to be related to income or income inequality. Stacking the $\boldsymbol{x}_i$:s into one matrix gives $\boldsymbol{X}$, which is $N \times K$. Moreover, let $\boldsymbol{y} = (y_1, \ldots, y_N)'$. The expression for the least squares estimate of $\boldsymbol{\beta}$ is then

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$$

An estimator $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ is said to be *asymptotically consistent* if $\lim\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}$ as $N \to \infty$. By rewriting (12) and applying Slutsky's theorem, it is possible to show that the limit condition stated above is satisfied if and only if $\mathbb{E}[\boldsymbol{x}_i|u_i] = \mathbf{0}$ so that $\mathbb{E}[\boldsymbol{x}_i u_i] = \mathbf{0}$. In practice, this means that an observation $\mathbf{x}_i$ must not be correlated with the random error term u_i. However, in econometrics, it is often the case that there is some sort of dependence between one or several of the independent variables and the error term. An independent variable that is correlated with the error term is said to be *endogenous*, while a variable uncorrelated with the error term is called *exogenous*. In the case of endogenity, $\hat{\boldsymbol{\beta}}$ is not a consistent estimator of $\boldsymbol{\beta}$. To remedy this problem, the well-known concept of *instrumental variables* will be briefly reviewed.

The empirical analysis tried to capture the fundamental linkage between economic and financial development on income inequality applying the three primary hypotheses, namely, Kuznets hypothesis, financial curve hypothesis, and financial Kuznets curve hypothesis. Following Beck et al., theoretical predictions were tested using Dynamic Panel Generalised Method of Moments (GMM) developed by Holtz-Eakin, Newey ,and Rossen ,Arellano and Bond and Arellano and Bover . GMM overcomes the problems of potential endogeneity and unobserved country-specific effects in the data ensuring the best fit for panel data.

The dynamic specifications of each panel model are as follows.

Kuznets curve hypothesis;

$$GINI_{it} = \alpha + \beta_0 GINI_{it-1} + \beta_1 \ \ln \ ED_{it} + \beta_2 \ \ln \ ED_{it}^2 + \delta \ \ln Z_{it} + \varepsilon_{it}$$

Financial curve hypothesis;

$$GINI_{it} = \alpha + \beta_0 GINI_{it-1} + \beta_1 \ \ln \ FD_{it} + \beta_2 \ \ln \ FD_{it}^2 + \delta \ \ln Z_{it} + \varepsilon_{it}$$

Financial Kuznets curve hypothesis;

$$GINI_{it} = \alpha + \beta_0 GINI_{it-1} + \beta_1 \ \ln \ EDG_{it} + \beta_2 G_{it} + \beta_3 \ \ln \ FDG_{it} \\ + \delta \ \ln Z_{it} + \varepsilon_{it}$$

where $GINI_{it}$ captures Gini net/gross income, $GINI_{it-1}$ captures lagged value of the dependent variable, ED_{it} captures economic development, FD_{it} captures financial development, and G_{it} captures GDP_{it} per capita growth. EDG_{it} is the product of economic development and GDP growth rate, FDG_{it} is the product of financial development and per capita GDP growth, while Z_{it} captures other control variables such as trade openness, inflation, government expenditure, dependency ratio, and urban population.

5.CONCLUSION

The study of income inequality and its implication for economic development reveals factors that shape the socio-economic status of nations. Through a paper review of theoretical frameworks, empirical evidence, historical perspectives, and policy implications, several key insights emerge.

Firstly, it is evident that income inequality is not merely a distributional issue but a multifaceted phenomenon with profound implications for overall economic growth and societal well-being. The theoretical data provide diverse perspectives on how income distribution affects investment, savings, human capital accumulation, consumption patterns, social mobility, and political stability. Empirical studies showcase these theories, demonstrating that excessive income inequality can hinder economic development by weakening investment, reducing aggregate demand, limiting opportunities for human capital development, and undermining social cohesion.

Policy interventions such as progressive taxation, social safety nets, education and skill development programs, labour market reforms, and wealth redistribution policies can play a crucial role in mitigating the adverse effects of income inequality and promoting equitable growth.

Furthermore, the role of globalisation and technological change in shaping income distribution dynamics cannot be overlooked. Globalisation has led to both opportunities and challenges, contributing to both rising and falling income inequality depending on the context.

In conclusion, the study of income inequality and its impact on economic development underscores the need for a holistic approach that considers the interplay of economic, social, political, and institutional factors. Addressing income inequality is not only a moral imperative but also a pragmatic strategy for promoting sustainable development and fostering a more prosperous and equitable society. By implementing targeted policy interventions and fostering inclusive growth strategies, policymakers can work towards creating a future where economic opportunities are accessible to all, irrespective of their socio-economic background.

6.REFERENCES

https://doi.org/10.1146/annurev.so.13.080187.001525
https://thenextrecession.wordpress.com/
https://www.wallstreetmojo.com

https://wid.world/
https://en.wikipedia.org/
https://link.springer.com/
https://www.worldbank.org/

Economies | Free Full-Text | Globalisation and Income Inequality in Developing Economies: A Comprehensive Analysis. (n.d.) retrieved April 15, 2024, from www.mdpi.com/2227-7099/12/1/23

P.G.D.A.V COLLEGE

****

RESEARCH PROJECT OF INTRODUCTORY ECONOMETRICS

****

TOPIC: - DO ECONOMIC CONDITIONS AFFECT PEOPLE'S DECISION TO ENTER THE LABOUR FORCE.

****

SUBMITTED BY: -

DEVENDRA MACHRA 12737

RAHUL RAY 12704

VIYOM RANA 12725

BA(HONS) ECONOMICS

INDEX

INTRODUCTION

The decision to enter the labor force represents a critical juncture in individuals' economic lives, influenced by a myriad of factors ranging from personal aspirations to broader socio-economic conditions. At the heart of this decision-making process lies the intricate relationship between individuals and the economy they inhabit. Economic conditions, characterized by fluctuations in growth, employment opportunities, and wage levels, exert a profound impact on individuals' calculus when considering labor force participation.

Historically, scholars and policymakers have recognized the significant role played by economic conditions in shaping labor market dynamics. Recessions, for instance, often coincide with spikes in unemployment rates and declining job opportunities, leading to heightened uncertainty and risk aversion among potential entrants into the labor force. Conversely, periods of economic expansion tend to be accompanied by robust job creation, higher wages, and greater optimism about future career prospects, incentivizing individuals to seek employment opportunities actively.

However, the relationship between economic conditions and labor force participation is far from deterministic. Various socio-economic factors, including educational attainment, family structure, and individual preferences, interact with macroeconomic indicators to shape individuals' decisions regarding labor force entry. For example, highly educated individuals may exhibit greater resilience to economic downturns, leveraging their skills and qualifications to navigate turbulent job markets more effectively. Similarly, cultural norms and social expectations regarding gender roles, caregiving responsibilities, and retirement age can influence individuals' labor force participation decisions in complex ways.

Against this backdrop, a rich body of literature has emerged, seeking to unravel the intricate mechanisms through which economic conditions influence labor force participation. From empirical studies analyzing the impact of business cycles on job search behavior to theoretical frameworks exploring the role of human capital accumulation in shaping individuals' labor market outcomes, researchers have explored diverse dimensions of this multifaceted relationship.

This literature review endeavors to synthesize and critically evaluate existing research on the nexus between economic conditions and labor force participation. By delving into the nuances of this relationship, we aim to identify key insights, highlight areas of consensus, and pinpoint gaps in knowledge. Moreover, we seek to elucidate the policy implications arising from our understanding of how economic conditions shape individuals' decisions to enter the labor force. Ultimately, by deepening our understanding of these dynamics, we can inform evidence-based interventions aimed at fostering a more inclusive, resilient, and equitable labor market for all.

Over the years, numerous studies have explored the relationship between economic conditions and labor force participation. These studies delve into various aspects, including the impact of recessions on job prospects, the role of educational attainment in navigating economic downturns, and the dynamics of labor market entry among different demographic groups. By examining historical trends, analyzing empirical data, and developing theoretical frameworks, researchers have shed light on the complex interplay between economic factors and individuals' decisions regarding labor force participation.

Labor force participation: The unemployment rate is closely linked to the labor force participation rate, which measures the proportion of the working-age population that is either employed or actively seeking employment. Changes in the labor force participation rate can affect the unemployment rate.

The unemployment rate is a key economic indicator that measures the percentage of the labor force that is unemployed and actively seeking employment. It is an important metric that reflects the overall health and strength of an economy, as well as the efficiency of labor markets

The GDP growth rate refers to the percentage increase in a country's Gross Domestic Product (GDP) from one period to the next, typically measured quarterly or annually. It's a key indicator of the economic health and performance of a nation. Positive GDP growth indicates that the economy is expanding, while negative growth suggests it's contracting. Various factors such as consumer spending, investment, government spending, and net exports contribute to GDP growth. Typically, governments and economists monitor GDP growth closely as it reflects overall economic activity and influences policy decisions.

Discouraged Worker Hypothesis: The discouraged worker hypothesis suggests that during economic downturns or periods of high unemployment, some individuals may become discouraged and stop actively seeking employment, effectively dropping out of the labor force. This phenomenon can lead to a decline in the labor force participation rate.

A Review Of Existing Literature…

1. Autor, D., & Dorn, D. (2013). "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market."

In their seminal paper, Autor and Dorn delve into the transformative changes occurring within the US labor market, particularly the proliferation of low-skill service jobs. Their research examines how these structural shifts influence individuals' decisions regarding labor force participation. By analyzing the dynamics of job creation and destruction across different sectors, the authors uncover a pattern of polarization within the labor market, characterized by the simultaneous expansion of high-skill and low-skill occupations at the expense of middle-skill jobs. This polarization phenomenon has significant implications for individuals contemplating entry into the labor force. With the growth of low-skill service jobs outpacing other sectors, individuals may be more inclined to enter or exit the labor force based on the availability and nature of these employment opportunities. Moreover, the decline in middle-skill occupations may limit the options available to workers, leading to increased competition for low-skill jobs and potentially influencing labor force participation decisions.

In conclusion, Autor and Dorn's research highlights the significant impact of structural shifts in the labor market, particularly the growth of low-skill service jobs, on individuals' decisions regarding labor force participation. The polarization of the labor market, with an expansion of both high-skill and low-skill occupations at the expense of middle-skill jobs, has profound implications for workers across various sectors. As the economy evolves, policymakers and employers must consider the changing nature of job opportunities and tailor interventions to address the diverse needs of the workforce.

2. Kahn, L. (2010). "The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy."

Kahn's research focuses on the enduring repercussions of graduating from college during economic downturns. By examining the long-term labor market outcomes of individuals who entered the workforce during recessions, Kahn sheds light on the challenges they face in terms of career advancement and earnings potential. Graduating during a recession not only exposes individuals to a more competitive job market with fewer opportunities but also sets them on a trajectory of lower wages and diminished job prospects that persist over time. The adverse effects of entering the labor force during economic downturns are particularly pronounced for recent college graduates, who may struggle to secure employment commensurate with their skills and education level. As a result, these individuals may experience prolonged spells of unemployment and slower career progression compared to their counterparts who graduated during periods of economic prosperity. Kahn's findings underscore the importance of understanding the long-term implications of early career experiences on individuals' labor market outcomes and highlight the need for targeted policies to support the economic integration of young workers during recessions.

Kahn's research underscores the enduring challenges faced by individuals who entered the labor force during economic downturns, particularly recent college graduates. The long-term consequences of graduating during a recession include lower wages, diminished job prospects, and slower career advancement, highlighting the need for targeted policies to support early-career workers during periods of economic uncertainty. By understanding the lasting impact of early career experiences on labor market outcomes, policymakers can develop strategies to mitigate disparities and promote economic mobility for all workers.

3. Ore. P., von Wachter, T., & Heisz, A. (2012). "The Short- and Long-Term Career Effects of Graduating in a Recession."

Ore. von Wachter, and Heisz explore the short- and long-term career trajectories of individuals who entered the labor market during recessions. Their research reveals the profound impact of economic conditions at the time of graduation on individuals' labor market outcomes. Graduating during a recession not only reduces job opportunities and initial earnings but also leads to slower career advancement and lower wages in the long run. These effects persist even after controlling for factors such as education level and field of study, indicating the enduring nature of the challenges faced by individuals who entered the labor force during economic downturns. The study underscores the importance of considering the timing of labor market entry in understanding disparities in career outcomes and highlights the need for targeted interventions to mitigate the adverse effects of recessions on early-career workers.

Ores., von Wachter, and Heisz's research emphasizes the persistent effects of economic conditions at the time of graduation on individuals' labor market outcomes. Graduating during a recession not only hampers initial job prospects and earnings but also leads to long-term disadvantages in terms of career advancement and wages. The findings underscore the importance of addressing the challenges faced by early-career workers during economic downturns and implementing policies to support their economic integration and upward mobility.

4. Fujita, S., & Ramey, G. (2009). "The Cyclicality of Separation and Job Finding Rates." Fujita and Ramey's research focuses on the cyclical patterns of separation and job finding rates in the labor market. Their analysis reveals the dynamic relationship between economic conditions and individuals' decisions regarding labor force participation. During economic downturns, individuals are less likely to voluntarily leave their jobs (separation rates decrease) due to heightened job security concerns and a scarcity of alternative employment opportunities. Simultaneously, job finding rates decline as businesses reduce hiring activity and job openings become scarce. These cyclical fluctuations in separation and job finding rates underscore the interconnectedness of economic conditions and labor market dynamics. Moreover, they have important implications for policymakers seeking to understand the impact of business cycles on employment outcomes and design effective interventions to support workers during economic downturns.

Fujita and Ramey's analysis highlights the dynamic nature of separation and job finding rates in response to changes in economic conditions. Understanding the cyclical patterns of labor market dynamics is essential for policymakers seeking to design effective interventions to support workers during economic downturns. By considering the interplay between economic cycles and employment outcomes, policymakers can develop targeted strategies to mitigate the adverse effects of recessions on workers and promote labor market resilience.

5.Farber, H. (2011). "Job Loss in the Great Recession: Historical Perspective from the Displaced Workers Survey, 1984-2010."

Farber's study provides a comprehensive analysis of job displacement during the Great Recession, drawing on historical data from the Displaced Workers Survey spanning the period from 1984 to 2010. By examining the prevalence and characteristics of job loss during economic downturns, Farber sheds light on the challenges faced by displaced workers and the factors influencing their decisions regarding labor force participation. The study highlights the heterogeneity of job displacement experiences across different demographic groups and industries, with certain sectors experiencing higher rates of layoffs than others. Moreover, it underscores the importance of understanding the historical context of job displacement in informing policy responses to mitigate the adverse effects of economic downturns on workers and promote labor market resilience. Farber's research contributes valuable insights into the dynamics of job loss during recessions and the role of policy interventions in supporting affected workers and facilitating their reintegration into the labor force.

Farber's research provides valuable insights into the prevalence and characteristics of job displacement during the Great Recession. The study underscores the need for policymakers to understand the historical context of job loss and develop targeted interventions to support affected workers and facilitate their reintegration into the labor force. By addressing the challenges faced by displaced workers during economic downturns, policymakers can promote labor market stability and resilience, ensuring that workers have the support they need to navigate transitions and thrive in a dynamic economy.

A review of existing literature on whether economic conditions affect people's decisions to enter the labor force would likely cover the following key points:

1.Labor force participation rate and economic conditions:

The labor force participation rate measures the proportion of the working-age population that is either employed or actively seeking employment.

Economic conditions, such as economic growth, unemployment rates, and wage levels, can influence an individual's decision to enter or exit the labor force.

2.Cyclical fluctuations and labor force participation:

During economic expansions, when job opportunities are more abundant and wages are generally higher, the labor force participation rate tends to increase as more individuals are incentivized to join the workforce.

Conversely, during economic downturns or recessions, when job prospects are scarce and wages may stagnate or decline, some individuals may choose to temporarily or permanently leave the labor force.

3.Discouraged worker effect:

The discouraged worker effect refers to the phenomenon where individuals who have been unemployed for an extended period may become discouraged and stop actively seeking employment, effectively dropping out of the labor force. This effect can be more pronounced during economic downturns, leading to a decline in the labor force participation rate.

4.Demographic factors:

Economic conditions can have varying impacts on different demographic groups, such as age, gender, and educational attainment levels.

For example, economic downturns may disproportionately affect certain industries or occupations, leading to changes in labor force participation rates for specific demographic groups.

5.Wealth effects and labor supply:

Changes in household wealth, driven by factors like asset prices or income levels, can influence labor supply decisions.

During periods of economic prosperity and rising wealth, some individuals may choose to retire early or reduce their labor force participation, while others may increase their participation to take advantage of higher potential earnings.

6.Government policies and labor force participation:

Government policies, such as unemployment benefits, tax policies, and retirement programs, can also impact labor force participation rates by affecting the incentives and constraints faced by individuals.

The literature review would likely explore these and other factors, drawing evidence from empirical studies, economic theories, and historical data to analyze the complex interplay between economic conditions and labor force participation decisions.

Data and Methodology

This analysis will look into the economics condition affect people's decision to enter the LABOUR force across the states of India, hopefully providing insight into the Indian LABOUR force, unemployment rate and GDP growth rate of past 15 years.

This section outlines the research design, data collection methods, and data analysis techniques, demonstrating the RIGOUR and validity of the research methods employed. Moreover, any specifications required will be provided in this section.

symbol	variable
Yi	Labour force participate rate
X2	Unemployment rate
X3	GDP growth rate
ε	Random error term

After careful thought and discussion, we decided that the most suitable functional form for this analysis would be the Multiple Linear Regression Model. The Model itself would look like the following:

$\hat{Y} = \beta 1 - \beta 2.\ x2 + \beta 3.\ x3 + \varepsilon$

Where,- Y^= labour force participate rate

- B1 is the intercept cooefficient

- Xi's are independent variable

- ε is the error term.

Data for report

Years	unemployment rate	labor force participate rate	GDP growth rate
2005	8.7	55.1	7.92
2006	8.63	55.28	8.06
2007	8.54	55.32	7.66
2008	8.35	55.47	3.09
2009	8.38	55.76	7.86
2010	8.32	55.9	8.5
2011	8.17	54	5.24
2012	8.1	53.2	5.46
2013	8.04	53	6.39
2014	7.98	52.8	7.41
2015	7.92	52.72	8
2016	7.84	52.4	8.26
2017	7.13	52.1	6.8
2018	7.65	51.9	6.45
2019	6.51	51.5	3.87
2020	10.2	50.3	5.83
2021	7.71	51.2	9.05
2022	7.33	52.2	7.24
2023	8.003	52.4	6.7

Source: Ministry of Labor And Employment

R code for the model

CONSOLE RESULTS:-

```
names(econometrics_part2)
[1] "Years"                "unemployment rate"         "labor force participate rate"
[4] "Gdp growth rate"
> reg1<-lm(`labor force participate rate`~`unemployment rate`+`Gdp growth rate`,data=econometrics_part2)
> summary(reg1)

Call:

lm(formula = `labor force participate rate` ~ `unemployment rate` +
    `Gdp growth rate`, data = econometrics_part2)

Residuals:
```

```
    Min      1Q  Median      3Q     Max 
-3.9904 -0.7923 -0.4985  1.4790  2.4107 

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)    
(Intercept)         48.42777    4.70405  10.295 1.83e-08 ***
`unemployment rate`  0.51777    0.56004   0.925    0.369    
`Gdp growth rate`    0.09972    0.26486   0.377    0.711    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.758 on 16 degrees of freedom
Multiple R-squared:  0.06351,	Adjusted R-squared:  -0.05355 
F-statistic: 0.5425 on 2 and 16 DF,  p-value: 0.5916
```

This output shows the summary of a multiple linear regression model fitted to the econometrics_part2 dataset. The model attempts to explain the labor force participate rate using two predictor variables: unemployment rate and Gdp growth rate.

Here's an interpretation of the key elements:

1. **Residuals**: The residuals section provides a summary of the model residuals, which are the differences between the observed values and the predicted values from the model. The minimum, 1st quartile, median, 3rd quartile, and maximum values of the residuals are given.
2. **Coefficients**:
    - **Intercept**: The estimated value of the intercept term is 48.42777, which is highly significant with a p-value < 0.001.
    - **unemployment rate**: The estimated coefficient for the unemployment rate variable is 0.51777, but it is not statistically significant (p-value = 0.369).
    - **GDP growth rate**: The estimated coefficient for the Gdp growth rate variable is 0.09972, but it is also not statistically significant (p-value = 0.711).
3. **Residual standard error**: The residual standard error is 1.758, which measures the average amount that the observed values deviate from the predicted values.
4. **Multiple R-squared**: The R-squared value is 0.06351, indicating that the model explains only about 6.35% of the variability in the labor force participate rate.
5. **Adjusted R-squared**: The adjusted R-squared is -0.05355, which is a modification of the R-squared value that accounts for the number of predictors in the model. The negative value suggests a poor fit.
6. **F-statistic**: The F-statistic is 0.5425 with a p-value of 0.5916, which is not statistically significant. This means that the overall model is not significantly better than a model with no predictors.

Based on the output, the multiple linear regression model does not appear to be a good fit for the data. Neither the unemployment rate nor the G growth rate are statistically significant predictors of the labor force participate rate. The low R-squared values and the insignificant F-statistic suggest that the model has limited explanatory power for the variation in the labor force participation rate.

This analysis indicates that additional predictors or different model specifications may be needed to better explain the labor force participation rate using the available variables in the econometrics_part2 dataset.

CONCLUSION: -

DATA IS NOT SUPPORTED TO DISCURAGED WORKERS THEORY THAT MEANS ADDED WORKERS THEORY IS APPROPIATE.

Based on the output, the conclusion is that the multiple linear regression model attempting to explain the "labor force participation rate" using the "unemployment rate" and "GDP growth rate" as predictors is not a good fit for the data.

The key points leading to this conclusion are:

1. Neither the "unemployment rate" nor the "GDP growth rate" are statistically significant predictors of the "labor force participation rate" in this model (p-values of 0.369 and 0.711, respectively).
2. The multiple R-squared value of 0.06351 and the adjusted R-squared value of -0.05355 indicate that the model explains very little of the variation in the "labor force participation rate".
3. The F-statistic of 0.5425 with a p-value of 0.5916 is not statistically significant, meaning that the overall model is not significantly better than a model with no predictors.
4. The residual standard error of 1.758 suggests that the model has a relatively high level of unexplained variation in the "labor force participation rate".

Therefore, the conclusion is that this multiple linear regression model, with only the "unemployment rate" and "GDP growth rate" as predictors, does not adequately explain or predict the "labor force participation rate" in this dataset. Additional predictors, different model specifications, or alternative modeling approaches may be needed to better capture the factors influencing the labor force participation rate.

REFERENCE: -

1. Darby, J., Hart, R. A., & Vecchi, M. (2001). Labour force participation and the business cycle: A comparative analysis of Europe, Japan and the United States. Japan and the World Economy, 13(2), 113-133.
2. Bredtmann, J., Otten, S., & Rulff, C. (2018). Husband's unemployment and wife's labor supply: The added worker effect across Europe. ILR Review, 71(5), 1201-1231.
3. Tuzemen, D. (2018). Why are prime-age men vanishing from the labor force? Economic Review, 103(1).
4. Blau, F. D., & Kahn, L. M. (2007). Changes in the labor supply behavior of married women: 1980-2000. Journal of Labor Economics, 25(3), 393-438.
5. Stephens Jr, M. (2002). Worker displacement and the added worker effect. Journal of Labor Economics, 20(3), 504-537.
6. Campolieti, M. (2018). Trade shocks and the added worker effect in married households. Labour Economics, 51, 156-171.

ECONOMIC GROWTH IN INDIA, IS IT A JOBLESS GROWTH?

AN EMPIRICAL EXAMINATION USING OKUN'S LAW

1.RAJU KUMAR (12747)

2. MANISH KUMAR (12726)

3. YOGESH KUMAR (12749)

Introduction

One of the cardinal objectives of macroeconomic policy is to achieve economic growth. Whyt is, why are all government policies designed in such a way that they will be consistent with the growth objective of the government? But the question is why is the government so concerned about economic growth? The answer to this question is multi-dimensional, but the most important one is to get political popularity. This is how the trend is: when there is an output expansion, the level of employment will rise, and the reduction in unemployment makes the government popular and this leads to winning elections. Therefore, the government will always seek output expansion even at the cost of inflation to get re-election by the public. The leading question to this episode is that does output expansion always lead to increased employment? There are three answers to the above question: when output expands rapidly, unemployment will reduce. Secondly, when output grows slowly, unemployment will rise, and finally, when output growth is equal to the potential output, unemployment will be constant.

At a theoretical level, Okun's law was the first effort made to describe the relationship that exists between output and unemployment. The whole idea started with the Keynes notion that at any time the economy is operating below full employment equilibrium, so that whenever there is an increase in aggregate demand above the present level of output, firms will hire more workers to increase the supply of output that will match the existing demand. In the process of doing so, aggregate employment will increase. This will continue until output reaches its optimal level. Therefore, to Keynes, the government can stimulate the economy through either an increase in spending or reduction in tax to boost aggregate demand which will imply output expansion and employment. At the empirical level, a lot of efforts exist trying to verify empirically the existence of Okun's law to different economies. See, for example, the works of Lancaster and Tulip (2015), Haggis (2011), Wen and Chen (2012), and Al-delaine (2016), Blinov (2014), Ball et al. (2012), and Lal (2010) among others. These researchers were conducted with the aim of determining whether the deviation of output from its potential level leads to unemployment. In other words, does an increase in output always lead to economic growth? The findings of these researchers vary across countries as some studies provide evidence that an increase in economic growth is followed by an increase in employment, whereas to others, an increase in output must be reasonable enough before it can lead to an increase in employment. The general conclusion in the literature is that economic growth does not in all the time imply a reduction in unemployment.

The main objective of this work is to analyze the relationship between economic growth and unemployment in India using Okun's law. Although the law has existed for a long period of time, the literature on its application to examine the relationship between employment and economic growth is very scanty; to the limited search of the literature, we only found the work of An et al. (2016) which fits the Okun's law for the low- and lower-middle-income countries like India inclusive. Therefore, the study intends to investigate whether economic growth in India always leads to a reduction in the level of unemployment' this will include the use of both linear and nonlinear econometric models so that the robustness of the estimates with several these approaches can be ascertained.

KEYWORDS: *unemployment, Okun's law, output*[i]

---

i rajukumareco22@pgdav.du.ac.in
Rajukumar1, Manishkumar2, Yogeshkumar3
PGDAV(M) college, Ring Road

Abstract

This study empirically examines the relationship between unemployment and output in India by fitting the Okun's law. To achieve this objective, an annual time series data for unemployment and output were collected from the World Bank and Reserve Bank of India. The empirical estimation starts with testing the unit root evidence using Perron (1997) and Elliot et al. (Econometric 64(4):813–836, 1996) DF-GLS, and the impact models were estimated using linear and nonlinear

econometric models. The estimates show that quantitatively, there is no significant difference between the two modelling approaches. The evidence indicates that the relationship between unemployment and output for the Indian economy is consistent with Okun's law. We find most of the coefficients to be negative, less than unity and statistically significant and this conforms to the theoretical expectation. The study also finds that based on the estimated evidence the 11.75% nominal GDP growth rate as targeted by the Indian government will only result in 0.52% decrease in unemployment which is insignificant for the Indian population. The study found that, to get a 1% decline in unemployment, 25 % nominal GDP growth rate is required which is twice the targeted value. Therefore, the study concludes that although Indian output is growing, the growth is jobless because it is not up to the threshold level that ensures a decline in unemployment.

Keywords: Unemployment, Okun's law, Output

Theoretical model

The existence of a meaningful relevance between economic growth rate of GDP and unemployment rate has been extensively studied in the economic literature (Misbah Akram, 2014). In 1962, Okun, in a paper, described two simple empirical relationships between unemployment and real production. In the first model, the seasonal changes in the unemployment rate Y (expressed as a percentage) were related to the seasonal percentage of changes in GDP X. This relationship was estimated by Okun using 55 observations, from the second season of 1947 to the fourth season of 1960 for America, as below:

$$Y=0.3-0.3X \qquad R^2=0.79 \qquad (1)$$

According to Okun's estimate, if the real GNP (X) doesn't change, the rate of unemployment (Y) will rise 0.3 cent (3%) from season to another season. Also, the unemployment rate will fall 0.3 percent for a one percent increment in the GNP. In alternative words, one percent (1%) higher unemployment rate means 3.3 percent lower GNP. The second Okun model, called the Gap Model, links the unemployment rate to the gap between actual and potential GDP and is expressed as follows:

$$U=a + b(gap) \qquad (2)$$

Where in this relation potential GDP "The amount of GDP at 4% unemployment rate (U = 4)" is considered. The b parameter for this equation is estimated to range between 0.28 and about 0.38 (Okun, 1962).

This equation was estimated in 1961 by the Board of Economic Advisers of the US Joint Economic Committee using the seasonal data for the period 1953-1960 as below:

U=3.72 +0.36 gap $\qquad$ R^2=0.93 $\qquad$ (3)

According to this equation, 1 percent increase in the unemployment would reduce the gap between real and potential GDP by 0.28 percent.Also, the estimated unemployment rate for a zero gap is 3.72 per cent, which is not far from the ideal 4 per cent level for full employment (Okun, 1962). According to Okun's interpretation, the parameter (a) in the above equation can be expound as the unemployment rate correlated with full employment. Also, since the actual rate of production is expected to be lower than its potential value, the sign of the coefficient (b)is positive.

The problem with this model is the use of full employment conditions and the amount of potential output that neither of these two data sets are directly visible and available in the macro economy. However, Okun was able to generate a series of data on potential outputs, assuming a four percent unemployment rate for full employment. But by changing the assumption of unemployment rates of full employment, different amounts of potential output are obtained. Okun also argues that simplifying these equations can be potentially problematic (Edward S. Knotek, 2007). The existence of these problems forced economists to make some changes to Okun's core relationships. Of course, these relationships and models are still called Okun's law, despite the considerable differences in their initial form of equations (Roma & Valde, 2017). One of these models is the "dynamic model", built on the assumption that in Okun's law some variables eliminated from the right side of the model; And more recently, many economists have used it in their studies (Knote, II.E.S., 2007) (Boulton, T., 2010) and Mossa, I. (2008). In the current dynamic model form, current and past years real production growth and past years unemployment rates as the right variables of the equality and the current changes in the unemployment rate are on the left equality, (Knut, 2007) Moses (2008) in his study has estimated this relationship.

$$U_t^c = \alpha + \sum_{i=1}^{m} \beta_i U_{t-i}^c + \sum_{i=0}^{n} \gamma_i Y_{t-i}^c + \upsilon_t$$

Although this model has some similarities to the differential model, the problem is that it is not as easily interpreted as the original differential model of Okun law. Another problem that is commonplace in the law of Okun (in all cases) is that labor unemployment is considered a representative of all idle resources. However, unemployment is only one of the determinants of the amount of labor, and other factors such as Population, the fraction of the population that is part of the labor force, and the number of hours worked by workers not included. These problems and deficiencies in Okun's law led to the creation of a "Production-Function Version". This model is derived by combining a theoretical production function with Okun's law-based slit model. In the theoretical production function used, products are derived from a combination of labor, capital and technology (Knote, II.E.S., 2007). According to (Freeman, 2001) study of total production (written in the form of a natural logarithm) it is stated as follows:

$$Y_t = \tau_t + \alpha K_t + \beta N_t$$

$Y_t = {}_t + \alpha K_t + B\ N_t$ These include Yt: product, Kt: volume of capital used, Net: employment, and t: technical progress. In this respect, according to (Paldam, M., 1987) and (Parktown, M. F. G., 1993) unemployment is defined as U=L-N, where L represents the total labor force. The " Okun’s law gap model” is thus defined (* sign indicates the equilibrium values of the variables).

$$Y_t - Y_t^* = \alpha(K_t - K_t^*) + \beta(L_t - L_t^*) - \beta(U_t - U_t^*)$$

It is necessary to explain that the technical progress at all-time points is assumed to be equilibrium (*). The Gap Model 6, by adding a random (*t*) component to account for productive shocks, can be specified and estimated for country (i) during the (t) period (Freeman, D. G. 2001) as below:

$$Y_{it}^{'} = \beta_{ui} U_{it}^{'} + \beta_{li} L_{it}^{'} + \beta_{ki} K_{it}^{'} + \varepsilon_{it}$$

Where the " ' " signifies the gaps of the variables or their deviation from the equilibrium values. This model, unlike earlier models, was more empirical; it had a theoretical structure that was considered an advantage, but the criticism of this model was difficult to measure in institutions such as capital reserves (Knote, II.E.S., 2007).

Although the existence of a stable empirical relationship such as the Okun relationship may be important for political modeling, few macroeconomic theories have modeled the association of GDP and unemployment. According to the conservative Keynesian view, it is very easy to interpret and explain Okun’s law. Due to changes in aggregate demand, companies are changing their production plans, which results in changes in labor demand and thus affecting the unemployment rate. But this kind of conception and logic only becomes problematic when prices and wages are implicitly assumed to be fixed. Neo-Keynesian economics attempts to overcome these problems by assuming the nominal and true inflexibility of these variables, for example if we consider, as in (Blanchard, O. and Kiyota, N.1987), a monopolistic competition model by introducing a cost list (or the same). Nominal Flexibility in the Commodity Market and Flexibility in the Labor Market It can easily be shown that changes in aggregate demand will affect production and employment and thus unemployment (SOGNER , L. and Stiassny, A., 2000).

In 1994, (Aghion, P. and Howitt, P.) and Hewitt examined the long-term effects of growth on unemployment. Their analysis revealed two rivalry effects of GDP growth on unemployment: The first effect was called investment, whereby increase in growth, through job creation, increases investment return and thus reduces the equilibrium unemployment rate. The second is called Creative Destruction, whereby increases in growth, due to increased demand for labor, increase the duration of work, thereby increasing the equilibrium level of unemployment. Lee, J. (2000) estimates that the eccentricity of Japan is 6.12 percent, indicating that this anomaly represents a significant institutional inflexibility in the Japanese labor market, especially given the level of job security available. Stephen, J. N. (1997) attributes the high unemployment in many European countries to the high rigidity in the labor market in these countries due to the high penetration power of unions, poor educational standards and the benefits of long-term unemployment. On the other hand, there has been an asymmetric relationship of unemployment rate and economic growth, according to studies conducted in recent years by some researchers, such as Cuaresma, J.C. (2003), Chuan Huang, H. and Chang, Y. K. (2005), Jardin, M. and Stephan, G. (2011). It shows the asymmetry of Okun's law. Courtney, H.G. (1991) is the first to propose that the coefficient of Okun may be different during times of boom and stagnation. Taking the "aggregate production function" approach, he has attributed the asymmetry of Okun's law to the succession of factors over periods, Multi-Factor Productivity Fluctuations and Changes of the Distribution in Sectoral Growth Rates (Harris, R. & Silverstone, 2001). Also, the results of some studies such as Fouquet, J. (2005) and Isthmian, M. (2010) indicate the instability of Okun's law over time.

Blanchard. (1999) has assumed that the stability of the Okun coefficient decreases over the time. change in GDP, we will always see a stronger impact on unemployment. He described the reasons for this as strong international competitiveness, less legal protection for workers, and a general shift in firm spending toward a reduction in the labor force (Sanger antialien, 2000). Isthmian, M. (2010) considers the Okun's coefficient as a reduced form of several structural parameters. According to his analysis, the Okun's coefficient tends to change over time. He argues that this coefficient of instability is primarily due to temporary changes in the parameters of the structural relationships between supply-side and labor demand variables, which are also likely to be due to changes in institutional legal characteristics and other related features. The labor market is a commodity. In fact, Isthmian, M. (2010) shows that Okun's law is inherently inclined to change over time, especially in response to structural changes in the legal-institutional characteristics and other characteristics of the labor and commodity markets. However, despite all the above considerations, Mossa, I. (2008) concluded that unemployment and production were irrelevant in four Arab countries. The reasons for this lack of communication include:

1. Unemployment is not cyclical. In some countries unemployment is structural or frictional, not periodic. The results in Structural unemployment from the mismatch of changes in the economy with changes of education. This means that unemployment is not because the economy is in recession, but because there is no need to learn the skills necessary to do the job. Frictional unemployment also stems from the mismatch of job vacancies with unemployed people. That is, people may have the skills to do specific tasks, but their lack of awareness of the positions that correspond to their skills has made them unemployed. According to some definitions of these two types of unemployment, increased production in such countries cannot reduce such unemployment.

2. Inflexible in the labor market. The flexibility of the labor market (for example, employer's freedom to hire or fire employee and the absence in labor laws like minimum wage) makes unemployment more affected by production. Therefore, in countries where the government plays a dominant role in the labor markets, the labor markets will be inflexible, which will lead to a lack of relation between production and unemployment, and

3. In economies dominated by the government or only one sector (such as, the oil sector), growth in this sector (which generates growth in this sector (which generates growth in this sector (which generates growth in this sector (which generates growth of economic) cannot decrease unemployment if most of the workforce is not concentrated.

Data and definition of variables

To estimate the relationship between unemployment and output for the Indian economy, annual time series data from 1991 to 2022 were collected from World Bank and Reserve bank of India . Two measures of output and unemployment were used, that is, unemployment (*U*) as a percentage of the total labor force which was sourced from the World Bank and youth unemployment which was sourced from the Reserve Bank of India. The two measures of output are Gross national income. (GNI) and GDP growth rate for output. The study chooses the sample size due to the unavailability of data for a long period. The data of unemployment that exist for Indian economy in World Bank and Reserve Bank of India.

Table: 1

Results of Stationarity analysis (Augmented Dickey Fuller Unit Root Tests) of GDP Growth rate

H0 Hypothesis: GDP growth has a unit root (Non-Stationary of variable)

		t-Statistic	Prob.*
ADF test statistic		-5.302579	0.0002
Test critical values:	1% level	-3.699871	
	5% level	-2.976263	
	10% level	-2.627420	

Source: Authors Calculations

According to Table 1, the (ADF) test results show that all variables are Stationary at 99%, 90%, and 95% confidence level. It means that our H_0 hypothesis (GDP growth has a unit root) is rejected, and our alternative hypothesis is confirmed (When prop value (0.0002) is less than 0.05 (in case 95% of confidence level) the Null Hypothesis is rejected and vice versa.), meaning that the growth rate of GDP is a Stationary variable.

Table 2- Hypothesis analysis table

confidence levels	critical values	Hypotheses
99% of confidence level	5.302579>3.699871	❖ H0 (Null) Hypothesis is accepted ❖ H1 (alternative) hypothesis is rejected
95% of confidence level	5.302579>3.699871	❖ H0 Hypothesis is accepted ❖ H1 hypothesis is rejected
90% level of confidence	5.302579>3.699871	❖ Null Hypothesis is accepted ❖ alternative hypothesis is rejected

Source: Authors Calculations

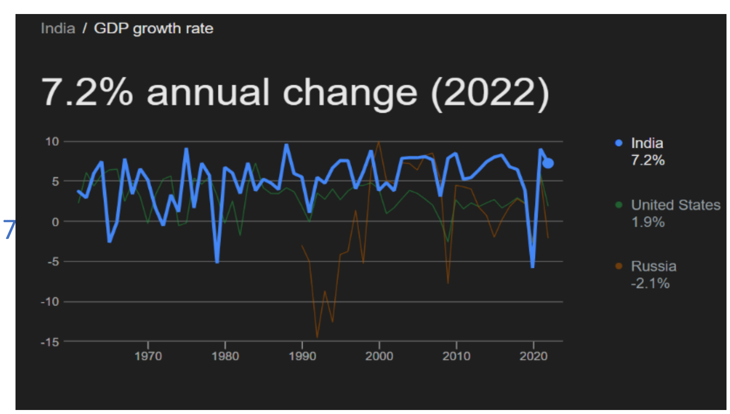


Table: 3
Results of Stationarity analysis (ADF Unit Root Tests) of unemployment rate

H0: Unemployment has a unit root (Non-Stationary of variable)

		t-Statistic	Prob.*
Augmented DF test statistic		-2.956217	0.0526
Test critical values:	1% level	-3.711457	
	5% level	-2.981038	
	10% level	-2.629906	

Source:Researchers Findings

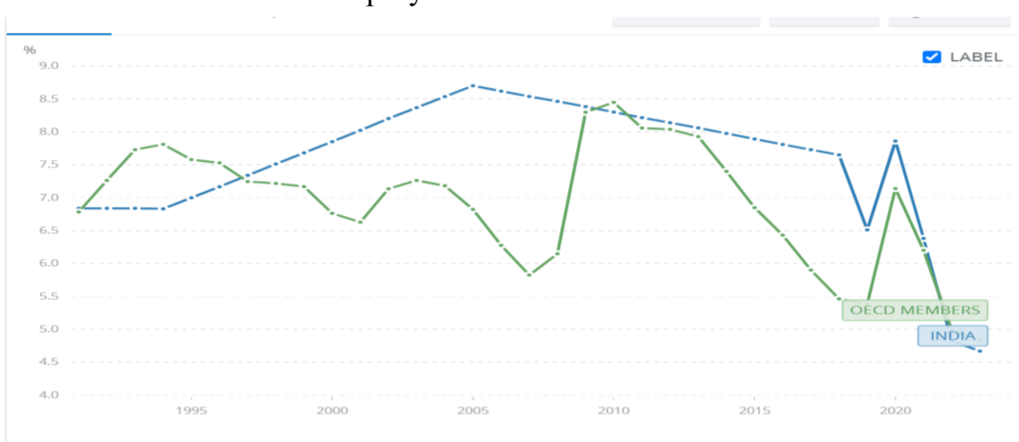
Table 4- Hypothesis analysis table

Confidence levels	Critical values	Hypotheses
99% of confidence level	2.956217<3.711457	❖ H0 Hypothesis is accepted ❖ H1 hypothesis is rejected
95% of confidence level	2.956217<2.981038	❖ Null Hypothesis is accepted ❖ alternative hypothesis is rejected
90% of confidence level	2.956217>2.629906	❖ H0 (Null Hypothesis) is rejected ❖ H1 (alternative hypothesis) is accepted

Source: Authors Calculations

According to Table 1 and 2, the Augmented DF test results show the non-Stationary of all variables at 99% confidence level and 90% confidence level but at 90% confidence level all variables are Stationary. And on the other hand, we chose our confidence level to analyze the relationships between variables at 90 %, so we aren't faced with the non-Stationary of variables. It means that our H0 (GDP has a unit root) at 90% confidence level is rejected, and our alternative hypothesis is confirmed, meaning that unemployment is a Stationary variable at 90% confidence level.

unemployment



The graph-1 relates to GDP growth over the years 1991-2018. As we can see from this graph, it is clear that the GDP growth variable is stable, and our data are distributed around the mean. Which can be deduced to be Stationarity of this variable. The graph-2 shows unemployment during the years 1991-2018. The distribution of the data and as well as the unemployment graph is also clear that unemployment data is stationary too. Which can be deduced as being this Stationarity.

Analysis of the OLS model results:

After performing the Augmented Dickey-Fuller Unit Root Tests we came to the conclusion that at 90% confidence level our variables are Stationary and we do not experience Unit Root problems and now in this section, using the EViews program and the Ordinary least-squares method, we obtained the regression equation, the results of which are shown in table 5 below.

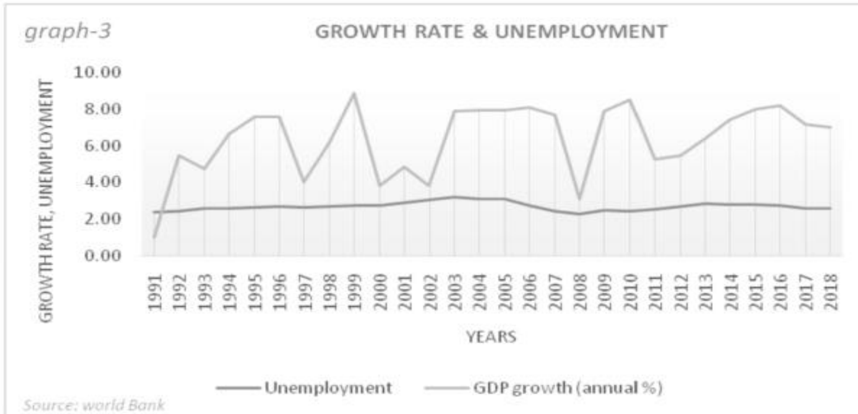
***Table 5- Results of* the OLS model**

Dependent Variable: Unemployment
Method: Least Squares
Sample: 1991-2018
Included observations: 28
Null Hypothesis: There is no significant relationship between variables

Variable	Coefficient	Error	t-Statistic	Prob.
C	2.455890	0.148719	16.51359	0.0000
GDPgrowth	0.035030	0.022386	1.564771	0.1297
R^2	0.086068	Mean dependent var		2.678929
Adjusted R^2	0.050917	S.D. dependent var		0.230480

Source: Authors Calculations

According to table 5: The results of the OLS (ordinary least square method) model estimation show that the relationship between GDP growth and unemployment is positive (Okun's law violation). In other words, with the increase in GDP growth rate, unemployment will go up. Conversely, unemployment decreases as GDP growth declines (Okun's law violation). On the other, since (Prob. = 0.1297) is greater than (0.05) H_0 hypothesis is accepted and H_1 is rejected. That is, there is no significant relationship of economic growth and unemployment.
The regression equation of GDP and unemployment is (***Unemployment = 2.40.035*GDPGROWTH***). This indicates a direct relationship between GDP and unemployment. The GDP growth coefficient is 0.035, which indicates that when changes one percent GDP growth it causes a 0.035 percent change in unemployment, in other words whenever the 100 percent change in GDP growth causes a 0.035 percent change in unemployment in a positive direction. The R-squared is 0.0861, which represents the total impact of the GDP growth on unemployment. It means 8.61 Changes in unemployment are due to GDP growth. The intercept is 2.46 which represents our unemployment without GDP growth which means if we exclude GDP growth, unemployment will be 6.14.



Reasons for rejecting Okun's law:

Violation of Okun's law in the Indian Economy In addition to the three that Mossa, I. (2008) has mentioned, there are several other reasons to be mentioned below:

a) Labor force transfer among sectors: The growth of the service and industry sectors in India has led to the shift of the labor force from agriculture to services and industry. That is, the labor force that was formerly in the agricultural sector as disguised unemployment and not productive, now with Transition to services and industry have contributed to economic growth.

b) Increasing labor productivity by increasing education and expertise has increased the share of each unit of labor in national production.

c) Increasing the labor force in India is another reason that the Okun's law is being violated. That is, every year a large volume of new labor force is joining the body of the Indian economy, which is a major challenge for the economy and the economy must create job for them. That is, no matter how high India's economic growth is, it is still affected by the growth of the labor force and cannot change unemployment.

d) There is another possibility that the current rate of unemployment in India is the natural Unemployment rate or lower than the natural unemployment rate. In this case too, one can understand the violation of Okun's law.

e) The inaccuracy of data on unemployment rate and as well as economic growth is another reason for not accurately calculating and still violating Okun's law. (It's a presumption.)

The above reasons may be part of the cause of the violating of the Okun's law. While there are many other reasons that need to be researched and each of the reasons mentioned above requires a separate investigation.

Conclusions:

Unemployment is one of the most important problems in the economies of countries and finding solutions to reduce unemployment is one of the uttermost important strategies of decision makers and decision makers. In this regard, attention to economic growth as one of the ways to reduce unemployment has received much attention and has been introduced in legal economic literature as the Okun's law. It explains the negative correlation between unemployment and GDP growth.
In other words, according to Okun findings, 1 percent increase in economic growth results in a 3 percent decrease in the unemployment. In the last decade, many studies have been done on Okun law and Okun factor has been evaluated based on different components. The results on the one hand indicate that the Okun coefficient has varied over time from country to country and from region to region and it has been violated at some point in time that some of the world's economic literature has focused on this issue.
This paper estimates Okun's coefficients using least squares method and based on latest World Bank data on Indian unemployment and economic growth. The results of this article show that the Okun's law has been violated in India. That is, there is no negative relationship between unemployment and GDP growth, on the contrary there is a positive relationship between them. The reasons for the violation of Okun's law can be summarized as Labor force transfer among sectors, increasing labor productivity, increasing the labor force in India, current unemployment rate in India is the natural unemployment rate or lower than the natural unemployment rate and inaccuracy of data on unemployment and economic growth. These reasons may be part of the cause of the violation of Okun's law. While there are many other reasons that need to be researched.

References:

Abu, N. (2017). Does Okun's Law Exist in Nigeria? Evidence from the ARDL Bounds Testing Approach. *International License*., 11(2), 131-144. *Retrieved from www.ce.vizja.pl.*
Acemoglu, H. (2018). Is there a Trade-off between Output and Unemployment? Evidence from Okun's Law for G-20 Countries. *Eskisehir Osman Gazi Universitas İİBF DeGise*, 13(2), 147-162. Retrieved from https://or-cid.org/0000-0001-6757-2140.
Amin-Naseri, M. R., & Rasouli, N. (2017). Investigation of the Relationship Between Economic Growth and Effective Factors by Multiple Linear Regression. *Research Gate*. Retrieved from https://www.researchgate.net/publication/323116281.
Chand, K., & Tiwari, R. (2018). Economic Growth and Unemployment Rate: An Empirical Study of Indian. PRAGATI: *Journal of Indian Economy*, 4(2), 130-137. doi:10.17492/pagati. v4i02.11468

David Lancaster, P. T. (2015). Okun's Law and Potential Output. *Research*, ISSN 1448-5109. Retrieved from http://www.rba.gov.au.
Girvaan, S. (2013). Okun's Law: An Empirical Investigation into Eurozone Growth and Unemployment. Economic *Research,* XXXI.

https://data.worldbank.org/country/india
https://www.imf.org/en/Countries/IND
https://labour.gov.in/annual-reports
http://www.mospi.gov.in/3-labour-and-employment

Research Project
Introductory Econometrics

—---

Topic:
A Study of Gender Inequality in the Indian Nation.

—---

By,

Ishan Sardana 12731

Saatwik Batra 12736

Atishay Jain 12713

B.A. (Hons) Economics

DATE OF SUBMISSION: 25th April 2024

Index

Introduction

Gender inequality remains a critical issue in India, impacting the lives of millions of women and girls across the country. Despite significant progress in recent decades, gender disparities persist in various aspects of life, including education, employment, healthcare, and political representation. This study seeks to delve deeper into the state-level dynamics of gender inequality in India, aiming to uncover the multifaceted nature of these disparities and their implications for development and social cohesion.

India's diverse and complex social fabric presents a unique context for examining gender inequality. While some states have made significant strides in promoting gender equity, others continue to grapple with deep-rooted patriarchal norms and practices that perpetuate discrimination against women and girls. Understanding these state-level variations is crucial for designing targeted interventions and policies that can effectively address the underlying causes of gender inequality.

Income, as measured by Gross National Income (GNI) and Net National Product (NNP), is a fundamental metric for evaluating a country's economic health and development. It reflects the total monetary value of goods and services produced within a nation's borders, including income earned from abroad. Gender plays a crucial role in income distribution, with women often facing lower earning potential and limited access to economic opportunities compared to men. This disparity is evident in various sectors, including employment, where women are frequently concentrated in lower-paying jobs and face barriers to advancement. As a result, GNI and NNP figures can highlight gender inequalities in income and economic participation, underscoring the need for policies and initiatives that promote gender equality in the workforce. Efforts to address these disparities can lead to more inclusive economic growth and sustainable development, benefiting society as a whole.

Employment is another critical area where gender inequality is pronounced. Women in India continue to face limited access to formal employment opportunities, particularly in certain states where traditional gender roles are more entrenched. This study will measure the overall participation in both the formal and informal sectors and perhaps provide insight into increasing participation in the formal sector especially by women.

GDP serves as a vital economic indicator. It reflects the overall economic output of each state, providing a measure of its economic strength and development. Understanding GDP variations across states can shed light on disparities in economic opportunities and resources, particularly concerning gender. Analyzing GDP alongside gender-specific data can reveal important insights into the economic dimensions of gender inequality within India.

Healthcare is also a significant concern in the context of gender inequality. Women in many states face barriers to accessing healthcare services, including limited availability of healthcare facilities and cultural taboos that inhibit women from seeking medical treatment. The IMR, CBR, and CDR, will provide this study with a baseline for the extent of the disparities found in this nation.

In this study, the **Gender Inequality Index (GII)** will serve as the focal point for measuring and understanding gender disparities. The GII is a composite measure encompassing reproductive health, empowerment, and economic activity disparities between genders, and offers a comprehensive lens through which to view the impact of gender inequality on human development. While the GII provides a global perspective on gender disparities, this research aims to apply it at the state level in India. By utilizing a range of indicators related to education, employment, healthcare, and other factors, this study seeks to provide a nuanced understanding of gender inequality in the Indian context

In conclusion, this study seeks to provide a comprehensive analysis of gender inequality in the Indian states, highlighting the challenges and opportunities for promoting gender equity. By understanding the state-level dynamics of gender inequality, policymakers can develop targeted interventions that address the root causes of discrimination and empower women and girls to fully participate in and contribute towards India's development.

*Please note that this study will primarily use data from 2000-2001 and 2010-2011.

A Review Of Existing Literature…

"Gender Inequality in India: Persistence, Resistance, and Change" edited by Anju Bakhshi and Jyoti Verma.

"Gender Inequality in India: Persistence, Resistance, and Change," curated by Anju Bakhshi and Jyoti Verma, thoroughly explores the intricate issue of gender disparity in India. Through a series of scholarly essays, the book probes into the diverse facets of gender inequalities and inspects the factors contributing to their endurance, and highlights endeavors toward resistance and transformation.

Commencing with a historical retrospective, the volume traces the origins of gender inequality in India to deeply entrenched social norms, cultural traditions, and patriarchal frameworks. Contributors dissect these entrenched power dynamics, illuminating how they shape gender dynamics and curtail opportunities for women across various domains of life.

A notable strength of the book lies in its intersectional lens towards dissecting gender inequality. By acknowledging the overlapping identities of caste, class, religion, and geographical region, the editors and contributors portray a nuanced understanding of the obstacles encountered by women from marginalized backgrounds. They emphasize the significance of addressing these intersecting forms of discrimination to foster substantial strides towards gender parity.

Additionally, the volume scrutinizes the roles played by diverse stakeholders in either perpetuating or challenging gender inequality. Ranging from policymakers and activists to grassroots organizations and individuals, each chapter explores diverse avenues for catalyzing change. Through case studies and empirical data, successful interventions and initiatives aimed at empowering women and deconstructing oppressive structures are elucidated.

Nevertheless, the book candidly recognizes the persistent hurdles and resistance faced in the journey towards advancing women's rights and achieving gender equality. Deep-rooted attitudes, institutionalized biases, and systemic prejudices continue to impede progress, necessitating sustained endeavors and collective action to surmount.

In conclusion, "Gender Inequality in India: Persistence, Resistance, and Change" offers an extensive and insightful exploration of an issue deeply embedded within Indian society. Through its diverse perspectives and analytical depth, the book not only unveils the complexities of gender inequality but also provides valuable insights into pathways towards substantial change. It is a crucial resource for scholars, policymakers, activists, and individuals committed to fostering gender justice and equality in India and beyond.

<u>"Gender Inequality and Women Empowerment: A Study of Indian Women" by Sanghmitra S. Acharya.</u>

In Sanghmitra S. Acharya's article, "Gender Inequality and Women Empowerment: A Study of Indian Women," the author delves into the intricate relationship between gender inequality and women's empowerment within the Indian context. Through an in-depth empirical investigation, Acharya sheds light on the diverse challenges confronting Indian women and examines potential avenues for meaningful empowerment.

Acharya initiates the discussion by presenting a comprehensive overview of the prevailing gender disparities in India. By drawing on a combination of statistical data and qualitative analysis, the author elucidates the manifold dimensions of inequality experienced by women across varying socio-economic backgrounds. The disparities encompass aspects such as education, employment, healthcare, and political representation, highlighting the pervasive nature of gender discrimination deeply entrenched within Indian society.

At the core of Acharya's examination lies an exploration of women's empowerment initiatives and their efficacy in tackling gender inequality. The author critically assesses governmental policies, interventions by non-governmental organizations, and community-driven efforts aimed at empowering women. Through this evaluation, Acharya endeavors to discern the extent to which these initiatives have succeeded in enhancing women's agency, access to resources, and decision-making capabilities.

A noteworthy aspect of the article is its emphasis on the intersectional nature of gender inequality. Acharya acknowledges that women's experiences are shaped not only by their gender but also by intersecting factors such as caste, class, religion, and geographical location. By adopting an intersectional perspective, the author unveils the intricacies of women's lived realities and underscores the imperative of addressing multiple layers of oppression to achieve holistic empowerment.

Moreover, Acharya delves into the pivotal role of education in fostering women's empowerment and challenging entrenched gender norms. Through the presentation of case studies and empirical evidence, the author illustrates how education can serve as a catalyst for societal transformation, enabling women to defy traditional roles, pursue economic opportunities, and advocate for their rights.

Nevertheless, amidst the acknowledgement of progress in advancing women's empowerment, Acharya also highlights the persistent obstacles and entrenched barriers hindering further advancements. Deep-seated patriarchal attitudes, systemic discrimination, and socio-cultural norms continue to impede women's full participation and inclusion across various spheres of life.

In summary, "Gender Inequality and Women Empowerment: A Study of Indian Women" by Sanghmitra S. Acharya provides valuable insights into the intricate dynamics of gender inequality and women's empowerment within India. Through its meticulous analysis and nuanced approach, the article enriches our understanding of the challenges and prospects for promoting gender equality and fostering women's empowerment in the country. It serves as an invaluable resource for policymakers, scholars, and activists committed to advancing women's rights and achieving gender justice in India and beyond.

<u>"Gender Inequality in India: A Multidimensional Perspective" edited by K. Narayana and R. Srinivasan.</u>

"Gender Inequality in India: A Multidimensional Perspective," edited by K. Narayana and R. Srinivasan, presents a thorough examination of the pervasive issue of gender disparity in India. Through a compilation of essays by esteemed scholars, the book delves into the complex array of factors contributing to gender inequality and explores potential avenues for addressing it.

The volume commences by providing a foundational understanding of the historical and cultural context of gender inequality in India. Contributors elucidate how entrenched patriarchal norms, societal expectations, and institutionalized discrimination have perpetuated gender disparities over time. This historical lens offers crucial insights into the intricate dynamics of gender relations in India.

A notable strength of the book is its multidimensional approach to analyzing gender inequality. Instead of focusing solely on one aspect, the editors and contributors examine gender disparities through various lenses, including education, employment, health, politics, and law. This holistic perspective enables a nuanced understanding of the multifaceted challenges faced by women in Indian society.

Moreover, the volume delves into the intersectionality of gender with other social identities such as caste, class, religion, and ethnicity. By acknowledging the overlapping forms of discrimination faced by marginalized groups, the book underscores the importance of addressing these intersecting axes of oppression to achieve meaningful gender equality.

The editors and contributors critically evaluate numerous policy interventions and initiatives aimed at addressing gender inequality in India. Through empirical research and case studies, they assess the effectiveness of government programs, legal reforms, and grassroots movements in promoting women's rights and empowerment. This analytical approach provides valuable insights for policymakers, activists, and practitioners striving to implement gender-sensitive policies and programs.

While acknowledging progress in certain areas, the book also highlights persistent challenges and barriers hindering gender equality in India. Deep-seated patriarchal attitudes, systemic biases, and socio-cultural norms continue to pose significant obstacles to women's empowerment and full participation in society.

In conclusion, "Gender Inequality in India: A Multidimensional Perspective" offers a thorough and insightful exploration of a pressing issue in Indian society. Through its diverse perspectives and analyses, the book contributes to a deeper understanding of the complexities of gender inequality and provides valuable insights into strategies for promoting gender justice and equality. It is an essential resource for scholars, policymakers, and activists dedicated to advancing women's rights and gender equality in India and beyond.

Data & Methodology

This analysis will look into gender inequality across the states of India, hopefully providing insight into the state of equality (of all genders), and giving the community, policymakers and other officials the ability to look into a period of definitive change in the country.

This section outlines the research design, data collection methods, and data analysis techniques, demonstrating the rigour and validity of the research methods employed. Moreover, any specifications required will be provided in this section.

Symbol	Variable	Definition
$\hat{Y}$	Gender Inequality Index	GII is a composite metric of gender inequality using three dimensions: reproductive health, empowerment and the labour market.
x2	GDP	It measures the monetary value of final goods and services—that is, those that are bought by the final user—produced in a country in a given period of time
x3	Employment	The No Of People Employed
x4	Infant Mortality Rate	Infant mortality rate (or IMR) is the number of infant deaths per 1,000 live births during the year.
x5	Gross national income	Gross national income, abbreviated as GNI, is the sum of the incomes of residents of an economy in a given period.
x6	Net National Income	NNP at factor cost, also known as Net National Income, represents the total income generated by an economy's factors of production, such as labour and capital
x7	Crude Birth Rate(CBR)	The number of live births per 1,000 midyear population.
x8	Crude Death Rate (CDR)	Number of deaths over a given period divided by the person-years lived by the population over that period.
x9	% of Women in Parliament	Women's representation in the Parliament of India, expressed as a Percentage.
ε	Random Error Term	The error in estimation.

After careful thought and discussion, we decided that the most suitable functional form for this analysis would be the **Multiple Linear Regression Model.**

The Model itself would look like the following:

$$\hat{Y} = \beta 1 + \beta 2 \,.\, x2 + \beta 3 \,.\, x3 + \beta 4 \,.\, x4 + \beta 5 \,.\, x5 + \beta 6 \,.\, x6 + \beta 7 \,.\, x7 + \beta 8 \,.\, x8 + \beta 9 \,.\, x9 + \varepsilon$$

Where,

- $\hat{Y}$ is the dependent variable (Gender Inequality Index).

- $\beta 1$ is the intercept coefficient.

- The βi*'s* are the coefficients for the independent variables (Population, Literacy rate, Public and private sector employment, Infant mortality rate, Sex ratio, GDP).

- The Xi's are the independent variables.

- ε is the error term.

Models & Results

Here we will look at 3 different models,

1. A model measuring the coefficients from 1990-2001,
2. A model measuring the coefficients from 2001-2011, &,
3. A model measuring the coefficients from 1990-2011

Beginning with the 1990-2011 model, it will showcase the entire picture and offer a comprehensive overview of the interaction of all the variables over the entire period as a whole, allowing us to identify any long-term relationships between the variables.

R code for the model:

[1]reg1<- lm(GII ~ GDP + Employment + IMR + GNI + NNP + CBR + CDR + Women_participation_in_politics, data = data2011)

[2]summary(reg1)

Console results:

```
Call:

lm(formula = GII ~ GDP + Employment + IMR + GNI + NNP + CBR +

    CDR + Women_participation_in_politics, data = data2011)



Residuals:

       Min         1Q     Median         3Q        Max

-0.0052437 -0.0006670  0.0009071  0.0016597  0.0028057



Coefficients:

                                  Estimate Std. Error t value Pr(>|t|)

(Intercept)                      2.301e-01  2.334e-01   0.986   0.3623
```

```
GDP                                8.561e-05  1.154e-03   0.074   0.9433
Employment                         6.285e-04  8.032e-04   0.783   0.4637
IMR                               -4.448e-03  8.077e-03  -0.551   0.6017
GNI                               -1.742e-07  2.648e-07  -0.658   0.5351
NNP                                1.876e-07  2.885e-07   0.650   0.5397
CBR                                8.453e-03  1.062e-02   0.796   0.4564
CDR                                4.674e-02  1.139e-01   0.410   0.6957
Women_participation_in_politics -9.222e-03  4.285e-03  -2.152   0.0749 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.003769 on 6 degrees of freedom
  (7 observations deleted due to missingness)
Multiple R-squared:  0.9931,   Adjusted R-squared:  0.9838
F-statistic: 107.4 on 8 and 6 DF,  p-value: 6.562e-06
```

<u>Interpretation Of Results:</u>

Coefficient Results are as follows,

GDP is 8.561e-05, which means that for every unit increase in GDP, the GII is estimated to increase by 8.561e-05 units, holding all other variables constant.

Employment is 6.285e-04, which means that for every unit increase in Employment, the GII is estimated to increase by 6.285e-04 units, holding all other variables constant.

IMR is (-4.448e-03), which means that for every unit increase in IMR, the GII is estimated to Decrease by (-4.448e-03) units, holding all other variables constant.

GNI is (-1.742e-07), which means that for every unit increase in GNI, the GII is estimated to Decrease by (-1.742e-07) units, holding all other variables constant.

NNP is 1.876e-07, which means that for every unit increase in NNP, the GII is estimated to increase by 1.876e-07 units, holding all other variables constant.

CBR is 8.453e-03, which means that for every unit increase in CBR, the GII is estimated to increase by 8.453e-03 units, holding all other variables constant.

CDR is 4.674e-02, which means that for every unit increase in CDR, the GII is estimated to increase by 4.674e-02 units, holding all other variables constant.

Women's participation in Politics is (-9.222e-03), which means that for every unit % increase in participation, the GII is estimated to decrease by (-9.222e-03)units, holding all other variables constant.

Hence we can say that the Gender Inequality Index (GII) becomes, Worse, with an increase in GDP, Employment, NNP, CBR, & CDR. And becomes Better with an increase in IMR, GNI, &, Women's participation in Politics.

The fitness of the model is explained as follows,
The **Multiple R-squared** value is 0.9931 which indicates that the proportion of variance in GII is explained by the independent variables in the model.

The Adjusted R-squared value is 0.9838. This means that approximately 98.38% of the variability in the Gender Inequality Index (GII) for the years 1990- 2011 is explained by the independent variables GDP, Employment, IMR, GNI, NNP, CBR, CDR, and Women's participation in politics.

F-statistic is 107.4 with a p-value of 6.562e-06, indicating that the model as a whole is statistically significant.

Now Checking And assessing the significance of coefficients,

GDP: p-value = 0.9433
Employment: p-value = 0.4637
IMR (Infant Mortality Rate): p-value = 0.6017
GNI (Gross National Income): p-value = 0.5351
NNP (Net National Product): p-value = 0.5397
CBR (Crude Birth Rate): p-value = 0.4564

CDR (Crude Death Rate): p-value = 0.6957
Women_participation_in_politics: p-value = 0.0749

None of the variables has p-values less than 0.05, which is the typical threshold for statistical significance. This suggests that, in this model and with this dataset, there is insufficient evidence to conclude that any of these variables have a significant effect on the Gender Inequality Index (GII).

However, the coefficient for the variable **Women's participation in politics** has a p-value of 0.0749. This p-value is greater than the typical significance threshold of 0.05, which means that the coefficient is not statistically significant at the conventional level. However, it is close to 0.05, so it is marginally significant.

Now for the ***1990-2001 model*,**

R code for the model:

[3]data_2001 <- subset(data2011, Years >= 1990 & Years <= 2001)

[4]model_2001 <- lm(GII ~ GDP + Employment + IMR + GNI + NNP + CBR + CDR ,
data = data_2001)

[5]summary(model_2001)

Console results:

```
Call:

lm(formula = GII ~ GDP + Employment + IMR + GNI + NNP + CBR +

    CDR, data = data_2001)



Residuals:

         1          2          3          4          5          6          7

 0.0012016 -0.0006421 -0.0005628 -0.0036169  0.0041328  0.0013464 -0.0039346

         8          9         10         11         12

 0.0030072 -0.0008979  0.0016065 -0.0027339  0.0010937



Coefficients:

              Estimate Std. Error t value Pr(>|t|)

(Intercept)  6.217e-01  8.110e-01   0.767    0.486

GDP          1.661e-03  1.167e-03   1.424    0.228

Employment   3.455e-04  1.267e-03   0.273    0.799
```

```
IMR          1.622e-02  2.089e-02   0.777    0.481

GNI          2.194e-06  1.504e-06   1.459    0.218

NNP         -2.409e-06  1.601e-06  -1.504    0.207

CBR         -3.226e-02  3.557e-02  -0.907    0.416

CDR         -3.988e-02  1.419e-01  -0.281    0.793


Residual standard error: 0.004205 on 4 degrees of freedom

Multiple R-squared:  0.9894,   Adjusted R-squared:  0.9709

F-statistic: 53.49 on 7 and 4 DF,  p-value: 0.0008644
```

<u>Interpretation Of Results:</u>

Coefficient Results are as follows,
GDP is 1.661e-03, which means that for every unit increase in GDP, the GII is estimated to increase by 1.661e-03 units, holding all other variables constant.

Employment is 3.455e-04, which means that for every unit increase in Employment, the GII is estimated to increase by 3.455e-04 units, holding all other variables constant.

IMR is 1.622e-02, which means that for every unit increase in IMR, the GII is estimated to increase by 1.622e-02 units, holding all other variables constant.

GNI is 2.194e-06, which means that for every unit increase in GNI, the GII is estimated to increase by 2.194e-06 units, holding all other variables constant.

NNP is (-2.409e-06), which means that for every unit increase in NNP, the GII is estimated to decrease by (-2.409e-06)units, holding all other variables constant.

CBR is (-3.226e-02), which means that for every unit increase in CBR, the GII is estimated to decrease by (-3.226e-02) units, holding all other variables constant.

CDR is (-3.988e-02), which means that for every unit increase in CDR, the GII is estimated to decrease by (-3.988e-02) units, holding all other variables constant.

Women's participation in Politics is NA, this is mainly due to insufficient data about the subject pre-1997 which is required for the analysis.

Hence we can say that the Gender Inequality Index (GII) becomes, Worse, with an increase in GDP, Employment, GNI, and IMR, And becomes Better with an increase in NNP, CBR, & CDR.

The fitness of the model is explained as follows,
The **Multiple R-squared** value is 0.9894 which indicates that the proportion of variance in GII is explained by the independent variables in the model.

The Adjusted R-squared value is 0.9709. This means that approximately 97.09% of the variability in the Gender Inequality Index (GII) for the years 1990- 2001 is explained by the independent variables GDP, Employment, IMR, GNI, NNP, CBR, And CDR.

F-statistic is 53.49 with a p-value of 0.0008644, indicating that the model as a whole is statistically significant.

Now Checking And assessing the significance of coefficients,

A summary of the t-values and p-values for each coefficient in the model using data from 2001:

GDP: t-value = 1.424, p-value = 0.228
Employment: t-value = 0.273, p-value = 0.799
IMR: t-value = 0.777, p-value = 0.481
GNI: t-value = 1.459, p-value = 0.218
NNP: t-value = -1.504, p-value = 0.207
CBR: t-value = -0.907, p-value = 0.416
CDR: t-value = -0.281, p-value = 0.793

Based on these results, none of the coefficients are statistically significant at the conventional level of 0.05 nor at 0.10. This suggests that, in this model and with this dataset, there is insufficient evidence to conclude that any of these variables have a significant effect on the Gender Inequality Index (GII) for the year 2001.

And finally, we will look at the *2001-2011 model,*

R code for the model:

[7]data_2011 <- subset(data2011, Years >= 2001 & Years <= 2011)

[8]model_2011 <- lm(GII ~ GDP + Employment + IMR + GNI + NNP + CBR + CDR + Women_participation_in_politics, data = data_2011)

[9]summary(model_2011)

Console results:

```
Call:

lm(formula = GII ~ GDP + Employment + IMR + GNI + NNP + CBR +

    CDR + Women_participation_in_politics, data = data_2011)


Residuals:

        12          13          14          15          16          17          18

 5.476e-05  2.325e-04 -1.634e-03  2.155e-03 -1.920e-03  2.673e-03 -1.105e-03

        19          20          21          22

-1.691e-03  1.204e-03  5.196e-04 -4.890e-04


Coefficients:

                                  Estimate Std. Error t value Pr(>|t|)

(Intercept)                      1.793e-01  3.343e-01   0.536    0.646

GDP                             -1.987e-04  1.372e-03  -0.145    0.898

Employment                      -8.130e-04  1.364e-03  -0.596    0.612

IMR                              7.853e-04  1.070e-02   0.073    0.948
```

```
GNI                               -2.362e-07  2.546e-07  -0.928    0.451
NNP                                2.656e-07  2.777e-07   0.957    0.440
CBR                               -2.936e-02  3.313e-02  -0.886    0.469
CDR                                1.679e-01  1.757e-01   0.955    0.440
Women_participation_in_politics -1.007e-03  7.140e-03  -0.141    0.901

Residual standard error: 0.003482 on 2 degrees of freedom
Multiple R-squared:  0.9952,  Adjusted R-squared:  0.9761
F-statistic: 52.06 on 8 and 2 DF,  p-value: 0.01898
```

Interpretation Of Results:

Coefficient Results are as follows,
GDP is **(**-1.987e-04), which means that for every unit increase in GDP, the GII is estimated to decrease by (-1.987e-04) units, holding all other variables constant.

Employment is **(**-8.130e-04), which means that for every unit increase in Employment, the GII is estimated to decrease by (-8.130e-04) units, holding all other variables constant.

IMR is 7.853e-04, which means that for every unit increase in IMR, the GII is estimated to increase by 7.853e-04 units, holding all other variables constant.

GNI is **(**-2.362e-07), which means that for every unit increase in GNI, the GII is estimated to Decrease by (-2.362e-07) units, holding all other variables constant.

NNP is 2.656e-07, which means that for every unit increase in NNP, the GII is estimated to increase by 2.656e-07 units, holding all other variables constant.

CBR is (-2.936e-02), which means that for every unit increase in CBR, the GII is estimated to decrease by (-2.936e-02) units, holding all other variables constant.

CDR is 1.679e-01, which means that for every unit increase in CDR, the GII is estimated to increase by 1.679e-01 units, holding all other variables constant.

Women’s participation in Politics is (-1.007e-03), which means that for every unit % increase in participation, the GII is estimated to decrease by (-1.007e-03) units, holding all other variables constant.

Hence we can say that the Gender Inequality Index (GII) becomes, Worse, with an increase in GDP, Employment, NNP, CBR, & CDR. And becomes Better with an increase in GDP, Employment, GNI, CBR, &, Women’s participation in Politics.

The fitness of the model is explained as follows,
The **Multiple R-squared** value is 0.9952 which indicates that the proportion of variance in GII is explained by the independent variables in the model.

The Adjusted R-squared value is 0.9761. This means that approximately 97.61% of the variability in the Gender Inequality Index (GII) for the years 2001 - 2011 is explained by the independent variables GDP, Employment, IMR, GNI, NNP, CBR, CDR, and Women’s participation in politics.

F-statistic is 52.06 with a p-value of 0.01898, indicating that the model as a whole is statistically significant.

Now Checking And assessing the significance of coefficients,

A summary of the coefficients with their t-values and p-values:
GDP: t-value = -0.145, p-value = 0.898
Employment: t-value = -0.596, p-value = 0.612
IMR: t-value = 0.073, p-value = 0.948
GNI: t-value = -0.928, p-value = 0.451
NNP: t-value = 0.957, p-value = 0.440
CBR: t-value = -0.886, p-value = 0.469
CDR: t-value = 0.955, p-value = 0.440
Women_participation_in_politics: t-value = -0.141, p-value = 0.901

These results suggest that, in this model and with this dataset, there is insufficient evidence to conclude that any of these variables have a significant effect on the Gender Inequality Index (GII) for the year 2011 at the 10 per cent or 5 per cent levels of significance.

Overall Result:

Overall, based on the regression analysis of the data from 2001 and 2011, it can be concluded that, except for Women's Participation in Politics, which has a marginal effect on the Gender Inequality Index (GII), the other indicators (GDP, Employment, IMR, GNI, NNP, CBR, CDR) have insufficient evidence of their effect on the GII. This suggests that these indicators may not be strong predictors of gender inequality in the context of the selected time periods.

Discussion and tests

Comparison between 2001 and 2011 models,

1. **Differences in R-squared values**
   The R-squared value Increased from 2001**(0.9894)** to 2011**(0.9952)**, suggesting that the independent variables in the 2001 model are less effective at explaining the variation in the dependent variable (GII) compared to the 2011 model. There exists a possibility this happened due to Missing data in the 2001 model, Namely, Women's participation in politics.

2. **Differences in adjusted R-squared values**
   The adjusted R-squared value increased from 0.9709 in the 2001 model to 0.9761 in the 2011 model. This indicates that the model using data from 2011 explains about (0.9761−0.9709=0.0052) 0.52% more variability in the GII.

3. **Changes in what worsened and what bettered the GII**
   Since only women's participation in politics was found to have a marginal effect on the GII, changes in this variable can be considered the main driver of any observed changes in the GII. However, due to insufficient data sources regarding this subject, we were unable to consider this variable in the 2001 model.

Comparison with the 1990-2011 model,

4. **Differences in R-squared values**
   In the 1990-2011 model, the R-square value was 0.9931 which is higher than the 2001 Model **(0.9894)** and lower than the 2011 Model **(0.9952)** which implies that the 1990-2011 model was more effective at explaining the variation in the dependent variable (GII) compared to the 2001 model but worse than the 2001-2011 model.

5. **Differences in adjusted R-squared values**
   In the 1990-2011 model, the R-square value was 0.9838. This means that approximately 98.38% of the variability can be explained by this model, which is better than both models.

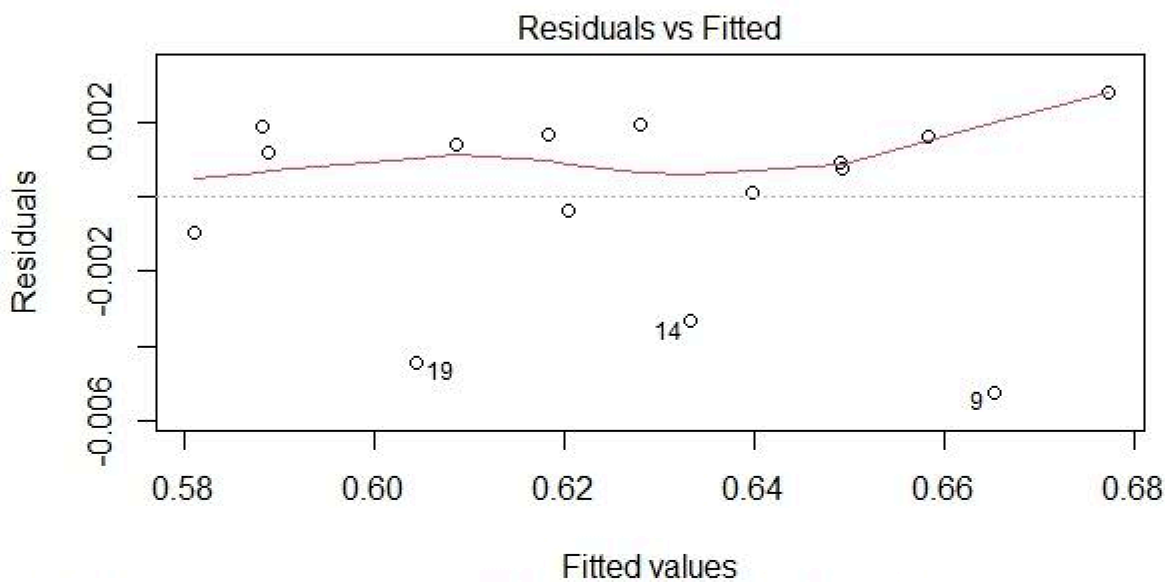
6. **Changes in what worsened and what bettered the GII**

   In the 1990-2011 model, the coefficient for Women's participation in Politics is negative (-9.222e-03), indicating that an increase in women's participation in politics is associated with a lower GII or improved gender equality, although it is only marginally significant (p-value: 0.0749), Moreover, the lack of data pre 1997 led to its removal in the 1990-2001 model and in the 2001-2011 model it has a high p-value indicating that it is not statistically significant.

   Hence, there is no clear trend in the coefficients of these variables across the three models. The coefficients are generally small in magnitude and not statistically significant, suggesting that these variables are not strong predictors of gender inequality in the selected time periods.

Tests of Residuals, Collinearity & Multicollinearity, Autocorrelation and Heteroscedasticity.

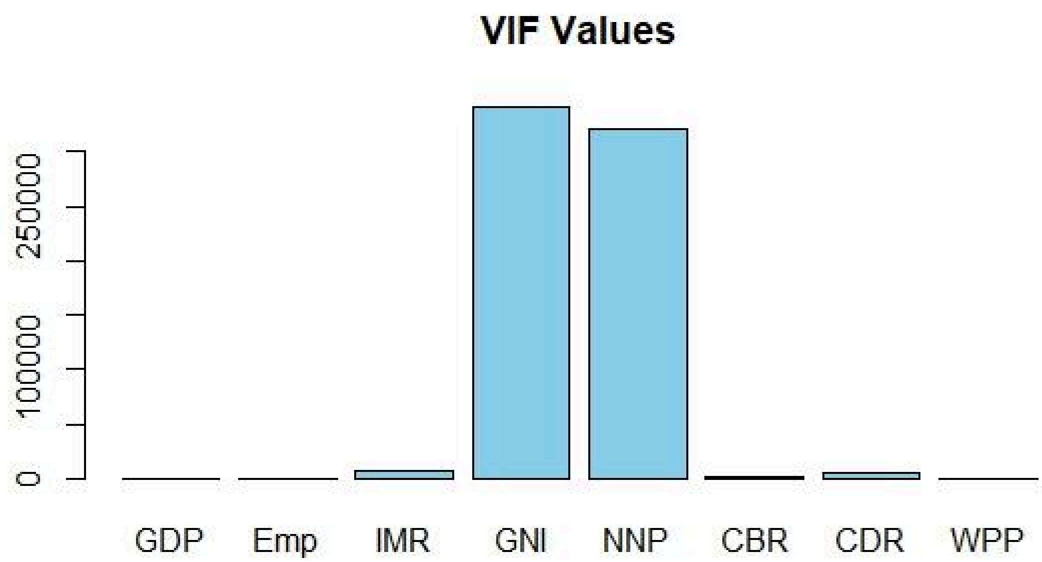
Residuals Analysis: Checking the residuals for normality, homoscedasticity, and independence, Plotting the residuals against the fitted values and checking for patterns or trends



Results:
The points are scattered around the red line, which suggests that the linear regression model is a good fit for the data. The random scatter indicates that the model captures the relationship between the independent and dependent variables reasonably, suggesting that the **linearity assumption is reasonable.**

The red line represents the expected values of the residuals if they were normally distributed. When the points are randomly scattered around this line, it indicates that the assumption of **homoscedasticity is reasonable**. This means that the variance of the residuals is roughly constant across all values of the independent variable(s).
A consistent spread of the residuals suggests that the model's errors have a stable variance, which is essential for the reliability of the regression analysis.

Collinearity Diagnostics: Calculating variance inflation factors (VIFs) to detect multicollinearity among independent variables, where, High VIF values will indicate collinearity issues.



Results:

```
              GDP                          Employment
     4.656801e+00                        3.739192e+01
              IMR                                 GNI
     6.093085e+03                        3.408055e+05
              NNP                                 CBR
     3.215899e+05                        6.791399e+02
              CDR Women_participation_in_politics (WPP)
     5.190823e+03                        2.185791e+01
```

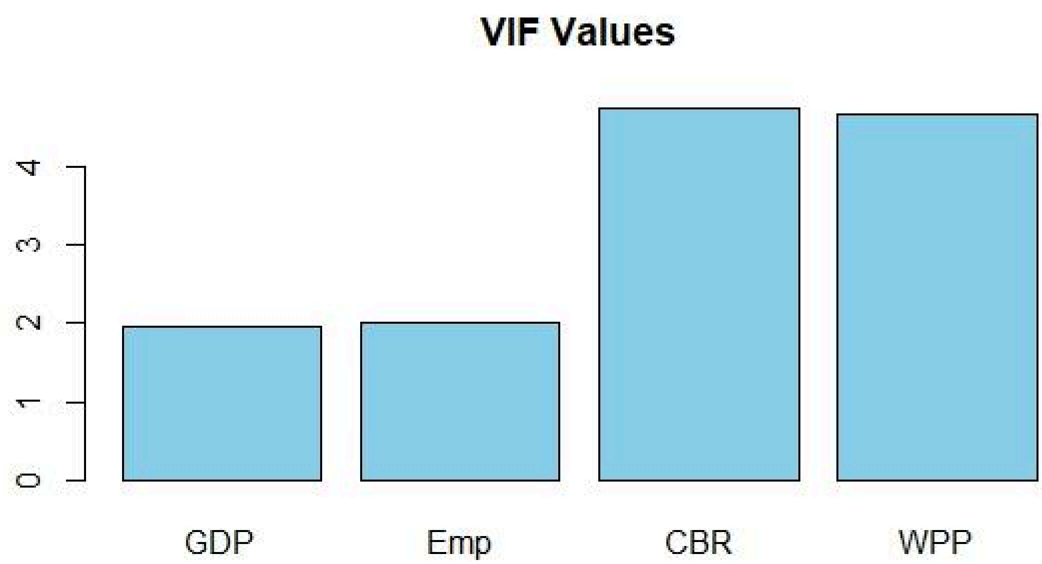
Interpretation:

1. GDP: VIF = 4.66 - Moderate multicollinearity.
2. Employment: VIF = 37.39 - High multicollinearity.

3. IMR: VIF = 6093.08 - Very high multicollinearity.
4. GNI: VIF = 340805.5 - Extremely high multicollinearity.
5. NNP: VIF = 321589.9 - Extremely high multicollinearity.
6. CBR: VIF = 679.14 - Very high multicollinearity.
7. CDR: VIF = 5190.82 - Extremely high multicollinearity.
8. Women_participation_in_politics,(WPP): VIF = 21.86 - Moderate multicollinearity.

High VIF values indicate that there is multicollinearity among the independent variables in the model. This is problematic because multicollinearity can inflate the standard errors of the coefficients, making some variables appear to be less statistically significant than they actually are.

To address this issue, we will rerun the model and drop the variables with extremely high multicollinearity, such as IMR, GNI, NNP, and CDR.



Results:

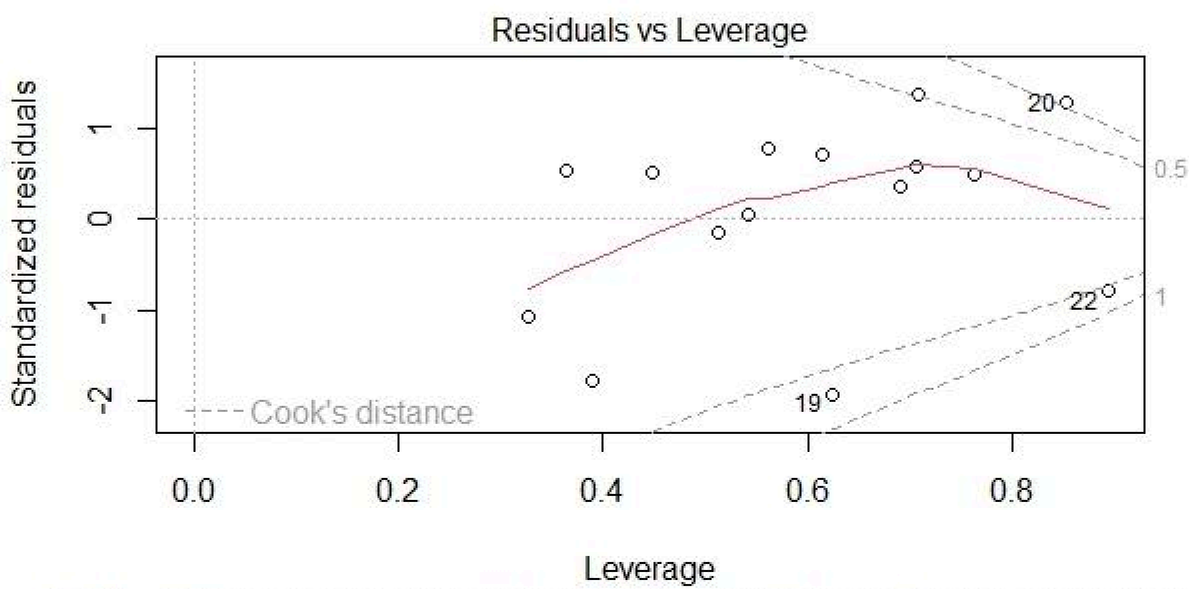
```
             GDP                    Employment
        1.956292                      2.005099
             CBR Women_participation_in_politics
        4.738849                      4.672783
```

The correlation coefficients you provided suggest moderate to high correlations between the following pairs of variables:

1. GDP and Employment: Correlation coefficient = 1.956292. This indicates a moderate positive correlation between GDP and Employment.
2. CBR and Women_participation_in_politics: Correlation coefficient = 4.738849. This indicates a high positive correlation between CBR (Crude Birth Rate) and Women_participation_in_politics.

Further analysis is required to understand the relationships between variables, but due to Our lack of knowledge and expertise, we are unable to do so at this point .

<u>Durbin-Watson Test:</u> Checking for Autocorrelation in the residuals and any values near 2 indicate No Autocorrelation.



H(0): No Autocorrection
H(a): Presence of Autocorrection

Performing the Test,

R code: durbinWatsonTest(reg1)

```
lag Autocorrelation D-W Statistic p-value

   1      -0.4792597      2.854956   0.616

 Alternative hypothesis: rho != 0
```

Interpretation:
A value around 2 suggests no autocorrelation, Furthermore, A p-value above the significance level (0.05) indicates that we failed to reject the null hypothesis, suggesting **no significant autocorrelation**.

Limitations

1. **Skewness in Our Data:** Due to only covering the 1990-2011 time period our data is Skewed (heavy-tailed) which has made proper interpretation difficult.

2. **Using Multiple Data Sources**: While collecting our data we had problems finding one entire data set comprising the entire time period as such we had to use multiple data sources and sets, combining them to complete our data set.

3. **Conflicting Data Sets**: As stated above, we had to use multiple data sources to complete our data set as the collection of the sources were from different time periods and organizations we had conflicting data points so we do the best of our knowledge and ability to choose the data points.

4. **Measuring Only Two Time Periods:** We are aware that the last major country-wide survey occurred in 2020-2021. However, due to the outbreak of the Covid-19 Pandemic, we feel that the data may be skewed or biased by various factors that occurred during the pandemic, predominantly, The Lockdowns, crippled the informal sector of India and had a severe impact on the formal sector. Moreover, it would be foolish of us to overlook the major economic and social ramifications that have occurred due to the pandemic, such as the widespread unemployment and the rampant inflation that has followed this period.

5. **Limited Historical Data on Women's Participation in Parliament:** While this variable had a marginally significant impact on the GII, due to poor data collection we were unable to find data pre-1997 which has led to a loss of critical insight into this crucial variable.

In Conclusion…

Based on the comprehensive regression analysis conducted across different time periods (1990-2011, 1990-2001, and 2001-2011) to assess the relationship between various socio-economic indicators and the Gender Inequality Index (GII), several key findings emerge:

Over the entire period of 1990-2011, the regression analysis suggests that several factors, including GDP, Employment, NNP, CBR, and CDR, are associated with an increase in the Gender Inequality Index (GII). Conversely, improvements in IMR, GNI, and Women's Participation in Politics are linked to a decrease in GII (improved gender equality), indicating a complex interplay of economic, social, and political variables influencing gender inequality.

While this suggests that these factors affect the GII only, women's participation in politics had a marginal significant impact on the GII while, we have inadequate evidence to prove that the other factors have any significant effect on the GII.

When analyzing specific time intervals (1990-2001 and 2001-2011), the relationships between these indicators and GII show variations. For instance, while some variables like GDP, Employment, and GNI consistently show a positive association with GII across all periods, others like NNP, CBR, and CDR exhibit inconsistent relationships depending on the time frame under consideration.

These findings have important implications for policymakers and researchers. While certain factors such as economic growth and employment are often assumed to lead to improvements in gender equality, the analysis suggests a more nuanced understanding is needed. Factors like maternal and child health (IMR), socio-economic development (GNI), and political empowerment (Women's Participation in Politics) play significant roles and should be prioritized in policy interventions aimed at reducing gender inequality.

It's crucial to acknowledge the limitations of this analysis, including data availability and the complexity of the GII as an aggregate measure. Future research could explore additional variables, consider regional or country-specific analyses, and employ more sophisticated methodologies to capture the multifaceted nature of gender inequality.

In conclusion, while some indicators show associations with gender inequality, the analysis underscores the need for a comprehensive, context-specific approach to addressing this complex issue. Policies targeting economic development, health outcomes, and political empowerment of women are essential for fostering greater gender equality and social justice. This research paper underscores the imperative of addressing gender inequality as a fundamental issue of social justice and human rights in India. By understanding the complex dynamics of gender disparities and adopting intersectional approaches, policymakers, scholars, and activists can develop more effective strategies for promoting gender equity and empowering women and girls across the country.

References

1. India Database
2. Previous Economic Surveys
3. The proportion of seats held by women in national parliaments (%) | Data
4. The decline in Rates of Maternal and Infant Mortality
5. 3.3 Vital Statistics
6. TABLE NO. 2.05: INFANT MORTALITY RATES IN INDIA 1911 - 2000
7. Essentials of Econometrics by Gujarati, Damodar N; Porter, Dawn C 4th Edition
8. Introductory Econometrics: A Morden Approach.7th Edition by Wooldridge, J. M. (2019)

The Impact of Foreign Direct Investment on Economic Growth

Gaurang, Divyansh, Aditya

(12722) (12733) (12732)

Student, B.A.(Honours) Economics, P.G.D.A.V. College, University of Delhi, India

ABSTRACT

A key indicator of an economy's health is its GDP (gross domestic product). It shows how big the economy is. It also aids in figuring out and contrasting the economy's growth with that of other economies as well as during various time periods. The primary driver of the GDP's growth is investment. An increase in investment level has a multiplier effect that raises income and employment levels. In open economies, foreign investment plays a major role. Both direct and indirect investments are possible. The purpose of the current study is to determine how the growth rate of foreign direct investment (FDI) in India and the growth rate of GDP in India are related to each other. Analysis of correlation was done in the study, and also, the results of regression analysis indicated the no significant impact of growth rate of FDI on growth rate of GDP.

Keywords: FDI, GDP, Economic Growth, Indian Economy

JEL Classifications: E01, F21, F43

INTRODUCTION

Gross Domestic Product (GDP) is one of the main economic indicators that gives snapshot of the economic position of a country. It is used to measure the size and growth of economy. One of the important factor which contributes to growth of GDP is the investment. According to Keynes investment plays key role in economic growth. Increase in level of investment will lead to increase in employment and income, which will lead to increase in demand in an economy. In short with increase in investment income and employment in an economy will increase multiple times depending upon propensity to consume in that economy.

Developing countries typically struggle with low investment since their citizens have inadequate incomes and have little capacity to take on risk. It will be expected that foreign direct investment (FDI) would have a major impact on rising investments in the Indian economy's development rate. Using the economy of Kazakhstan as an example, Rakhmatullayeva et al. (2020) analyzed the effect of FDI inflows on the economic growth of the host nation. The study attempted to assess the impact of foreign direct investment (FDI) using a multiple regression model for the years 2000–2017. The analysis's findings concluded that FDI had no detrimental effects on economic growth. The authors did note, nevertheless, that monitoring the expansion of the national economy does not require the presence of a positive link. Sokang (2018) explored how foreign direct investment affected Cambodia's economic development. The time series data spanning the years 2006 to 2016 was used by the author. Multiple regression analysis and correlation analysis were used to assess the data that had been gathered. The study discovered that FDI has an advantageous impact on the chosen nation's economic expansion. In order to attract more foreign direct investment within the nation, the report advised the government to start modernizing the domestic markets. Balasubramanyam et al. (1996) used a novel growth theory framework to investigate the function of foreign direct investment in the growth process in developing nations with varying trade policy regimes. Additionally, Balasubramanyam et al. (1999) examined how foreign direct investment (FDI) contributes to economic growth and found that the relationship between FDI and human capital has an enormous impact on growth performance. It will be expected that FDI will boost the host country's economic growth (Zhang, 2001; Hermes and Lensink, 2003). GDP has been influenced by FDI through the transfer of technology and know-how (Hansen and Rand, 2006; Berthélemy and Démurger, 2000). The association with FDI and growth is less clear than the relationship between FDI and growth, as noted by Choe (2003). Studies by Basu et al. (2003) and Basu and Guariglia (2007) looked at how liberalization affected the dynamics of the connection between GDP and FDI.

Using panel data analysis for a sample of 18 Latin American countries from 1970 to 1999, Bengoa and Sanchez-Robles (2003) investigated the relationship between economic freedom, foreign direct investment (FDI), and economic growth. They observed that FDI is continually significantly and positively correlated with economic growth.

n 2015, Rahaman and Chakraborty investigated if there was an a causal connection between GDP (gross domestic product) and foreign direct investment (FDI). The study's primary emphasis was Bangladesh. To verify the existence of a long-run equilibrium unity, the researchers employed the cointegration test approach.

To confirm that there is a unidirectional causal relationship between FDI and GDP, the Granger causality test was further used. The findings showed that FDI inflow was much lower than in other nearby countries. The nation requires sophisticated facilities and infrastructure, skilled labor, an abundance of power generation, a macroeconomic framework which encourages investments, and political stability in order to draw overseas capital. Johnson (2006) employed panel and cross-sectional data sources. It chose 90 nations and the data was collected between 1980 and 2002. The author observed that FDI inflows had a favorable effect on the economic growth of the chosen nations. The feasibility of a causal link between FDI and economic development as economic factors was examined by Chowdhury and Mavrotas (2003). An new econometric methodology was utilized in the study to evaluate the causal relationship between the variables under analysis. The Toda-Yamamoto test for cause and effect served as the foundation for the study's approach.

The idea that more foreign direct investment (FDI) would encourage economic growth is supported theoretically. nevertheless studies have indicated that a rise in FDI and an increase in GDP can have either a good or negative impact, or no effect at all (Türkcan et al., 2008).

The expansion of the Indian economy following independence may be divided into two periods: pre-liberalization, or 1947–1990, and post-liberalization, or after 1990. In the years before the 1990s, India's industrial mechanism possessed both financial and technological strength.

The New Economic Policy of the 1990s relaxed FDI regulations while concurrently reorganizing domestic businesses. Many of the obstacles that were before placed in the way of foreign direct and indirect investment were removed. The formalities needed to get technical and financial clearances for collaboration underwent a thorough revision. The nation's central bank was granted the authority to consequently approve a number of certain businesses. This accelerated the expansion of the time-series data covering the years 1969 to 2000 and strengthened competition within the industry. The nations that were chosen were all developing nations: Malaysia, Thailand, and Chile. The study's conclusions unambiguously suggested that, in the instance of Chile, foreign direct investment is caused by its gross domestic product, or GDP. On the other hand, there was abundant proof of the existence of a reciprocal causal connection between the variables under investigation for two of the other nations included in the study: Malaysia and Thailand.

RESEARCH OBJECTIVES AND RESEARCH METHODOLOGY

The study's goal is to quantify the influence and link between foreign direct investment and India's GDP growth rate. Data on foreign direct investment and India's economic growth rate from 2001 to 2019 will be examined as part of this study. Data from the World Bank website (http://api.worldbank.org/v2/en/country/IND?downloadformat=excel) was used throughout the study for the variables. As part of the study, the World Bank's data on FDI (net inflows) was used for calculating the growth rate of FDI. Regression and correlation are the statistical techniques employed in the present investigation.

RESEARCH HYPOTHESIS

The hypothesis devised to appraise the significance of association between growth rate of FDI net inflows and growth rate of GDP of India and significance of impact of growth rate FDI (net inflow) on growth rate of GDP of India have been stated as:

H1: There is significant relationship between growth rate FDI (net inflows) and growth rate of GDP of India

H2: There is significant impact of growth rate FDI (net inflows) on growth rate of GDP of India

Table 1: Growth rate of GDP and FDI (net inflows) in India

Year	Growth rate of GDP (annual %)	Foreign direct investment, net inflows (Annual %)
2001	4.823966	43.07429
2002	3.803975	1.577068
2003	7.860381	−29.3145
2004	7.922937	47.45447
2005	7.923431	33.89337
2006	8.060733	175.5262
2007	7.660815	25.95532
2008	3.086698	72.05773
2009	7.861889	−18.0271
2010	8.497585	−23.0022
2011	5.241345	33.22191
2012	5.456359	−34.256
2013	6.386106	17.32539
2014	7.410228	22.81677
2015	7.996253	27.28098
2016	8.256306	1.020415
2017	7.043821	−10.1049
2018	6.119587	5.382962
2019	5.023873	20.15287

Source: http://api.worldbank.org/v2/en/country/IND?downloadformat=excel

DATA ANALYSIS

Fig The following Table 1 exhibits the growth rate of GDP and FDI (net inflows) from the year 2001 to 2019.

Figure 1 , The GDP and FDI (net flows) growth rates in India are presented graphically in Figure 1. The Indian economy has expanded 6.65% between 2001 and 2019, as can be shown. During this time, 2008 had the lowest yearly growth rate in the nation, growing at a mere 3.09%. India's net inflows of foreign direct investment (FDI) increased at an average pace of 21.69% between 2001 and 2019. During the research period, India registered the greatest annual growth rate of foreign direct investment (175.5262%) in 2006. It is also apparent that the nation as a whole experienced negative growth rates in 2003, 2009, 2010, 2012, and 2016.

Figure 1: Growth rate of GDP and FDI (net inflows) in India

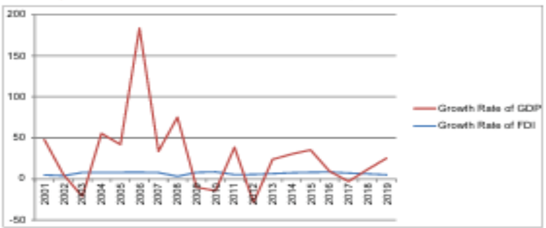


Table 2: Model summary

Model	R	R square	Adjusted R square	Std. error of the estimate
1	0.043[a]	0.002	−0.057	47.9100949

a. Predictors: (Constant), Growth Rate of FDI

Table 3: ANOVA

Model		Sum of squares	df	Mean square	F	Sig.
1	Regression	71.045	1	71.045	0.031	0.862[a]
	Residual	39021.412	17	2295.377		
	Total	39092.458	18			

a. Predictors: (Constant), growth rate of FDI
b. Dependent variable: Growth rate of GDP

Table 4: Coefficients

Coefficients[a]						
Model		Unstandardized coefficients		Standardized coefficients	t	Sig.
		B	Std. error	Beta		
1	(Constant)	29.760	47.190		0.631	0.537
	Growth rate of FDI	−1.213	6.896	−0.043	−0.176	0.862

a. Dependent variable: Growth rate of GDP

Impact of Growth Rate FDI Inflow and Growth Rate of GDP of India

The study used regression analysis to calculate the association between the growth rates of India's GDP and net foreign direct investment inflows. Table 2 shows the summary of the model. The results reveal that the dependent variable (growth rate of GDP) can be determined by the independent variables in the model (growth rate of FDI) by 0.2%. This is confirmed by the R2 value of 0.002.

Table 3, the ANOVA findings, which show if the overall regression model fits the data well, are shown in Table 3. The findings demonstrate that the statistically significant growth rate of GDP is not predicted by the growth rate of FDI (net inflows). Table 4 displays the coefficients of regression

The findings suggested that the growth rate of FDI (net inflows) had no discernible effect on the growth rate of India's GDP. Therefore, the hypothesis H2—that is, that the growth rate of FDI (net inflows) has a substantial impact on the growth rate of India's GDP—is rejected.

In theory, foreign direct investment may boost growth in the economy. This is due to the fact that foreign direct investment often moves from capital-rich countries to capital-poor countries. It also gives a boost to elevated production and the invention of new technologies. However, because it may cause a disturbance in the market, foreign direct investment might be detrimental to growth. Furthermore, it could help the transferring community while inhibiting the development of the nation receiving.

CONCLUSION

One major measure of the size of the economy is the gross domestic product. The GDP growth rate is useful for calculating and contrasting the economy's growth with that of other economies as well as throughout various time periods. Every economy can benefit from investment. It is the primary factor in any nation's progress. An increase in investment level has a multiplier effect that raises income and employment levels. Foreign investment is one of the major sources of funding, particularly for emerging economies. It is possible to invest overseas directly or indirectly. The goal of the current study was to determine the link between the influence of India's GDP growth rate and the growth rate of foreign direct investment in India. It was discovered that the growth rate of FDI and the growth rate of the economy had no discernible link. Regression analysis findings also showed that the growth rate of FDI had no discernible effect on the growth rate of GDP.

REFERENCES

Alfaro, L., Chanda, A., Kalemli-Ozcan, S., Sayek, S. (2004), FDI and economic growth: The role of local financial markets. Journal of International Economics, 64(1), 89-112.

Balasubramanyam, V.N., Salisu, M., Sapsford, D. (1996), Foreign direct investment and growth in EP and is countries. The Economic Journal, 106(434), 92-105.

Balasubramanyam, V.N., Salisu, M., Sapsford, D. (1999), Foreign direct investment as an engine of growth. The Journal of International Trade and Economic Development, 8(1), 27-40.

Basu, P., Chakraborty, C., Reagle, D. (2003), Liberalization, FDI, and growth in developing countries: A panel cointegration approach. Economic Inquiry, 41(3), 510-516.

Basu, P., Guariglia, A. (2007), Foreign direct investment, inequality and growth. Journal of Macroeconomics, 29, 824-839.

Bengoa, M., Sanchez-Robles, B. (2003), Foreign direct investment, economic freedom and growth: New evidence from Latin America. European Journal of Political Economy, 19, 529-545.

Berthélemy, J.C., Démurger, S. (2000), Foreign direct investment and economic growth: Theory and application to China. Review of

Development Economics, 4(2), 140-155.

Carkovic, M., Levine, R. (2002), Does Foreign Direct Investment Accelerate Economic Growth? Available from: https://www.piie.com/publications/chapters_preview/3810/08iie3810.pdf.

Choe, J.I. (2003), Do foreign direct investment and gross domestic investment promote economic growth? Review of Development Economics, 7(1), 44-57.

Chowdhury, A., Mavrotas, G. (2003), FDI and Growth: What Causes What? Paper presented at the WIDER Conference on Sharing Global Prosperity, WIDER. Helsinki: World Bank Group.

De Mello, L.R. (1999), Foreign direct investment-led growth: Evidence from time series and panel data. Oxford Economic Paper, 51, 133-151.

Durham, J.B. (2004), Absorptive capacity and the effects of foreign direct investment and equity foreign portfolio investment on economic growth. European Economic Review, 48, 285-306.

Hansen, H., Rand, J. (2006), On the Causal Links Between FDI and Growth in Developing Countries. Copenhagen: University of Copenhagen and Development Economics Research Group.

Hermes, N., Lensink, R. (2003), Foreign direct investment, financial development and economic growth. Journal of Development Studies, 40(1), 142-163.

Johnson, A. (2006), The Effects of FDI Inflows on Host Country Economic Growth, No. 58, Working Paper Series in Economics and Institutions of Innovation. Tamil Nadu: Royal Institute of Technology, CESIS

Centre of Excellence for Science and Innovation Studies.
Rahaman, A., Chakraborty, S. (2015), Effects of foreign direct investment on GDP: Empirical evidence from developing country. Advances in Economics and Business, 3(12), 587-592.
Rakhmatullayeva, D., Kuliyev, I., Beisenbaiyev, Z., Tabeyev, T. (2020), Assessment of the influence of FDI on the economic growth of the host country: Evidence from Kazakhstan. E3S Web of Conferences, 159, 06007.

What is Regression?

Regression analysis is a statistical technique used to model and analyze the relationship between a dependent variable and one or more independent variables. Its purpose is to understand how changes in the independent variables are associated with changes in the dependent variable. Regression analysis aims to find the best-fitting mathematical model that represents this relationship. It's a versatile tool employed in various fields, from predicting sales trends in digital marketing to understanding the impact of advertising campaigns on consumer behavior. There are different types of regression models, such as linear regression, logistic regression, polynomial regression, and more, each suited to different types of data and relationships.

Applications of Regression Analysis in Real-Life Scenarios

Business and marketing:

- Predicting sales based on factors like advertising expenditure, product price, and consumer demographics.
- Analyzing customer behavior and purchase patterns to optimize marketing strategies.
- Forecasting demand for products or services based on historical data and market trends.

Finance and economics:

- Modeling stock prices or commodity prices based on economic indicators and market factors.
- Analyzing the impact of interest rates, inflation, and other variables on economic growth or household income.
- Estimating the effect of policy changes (e.g., tax reforms) on economic variables like employment or consumer spending.

Healthcare and medicine:

- Studying the relationship between risk factors (e.g., age, lifestyle, genetics) and the prevalence of diseases.
- Analyzing the effectiveness of medical treatments or drugs based on patient characteristics and other variables.
- Predicting the length of hospital stays or the likelihood of readmission based on patient data.

Environmental studies:

- Modeling the impact of factors like industrial emissions, deforestation, or climate change on air quality or biodiversity.
- Predicting the spread of wildfires or the occurrence of natural disasters based on weather conditions and environmental factors.

Engineering and manufacturing:

- Optimizing product design or manufacturing processes based on the relationship between input variables and output quality.
- Predicting the lifespan or failure rate of components or systems based on usage patterns and environmental conditions.

Social sciences:

- Analyzing the relationship between socioeconomic factors (e.g., income, education, family structure) and various outcomes like academic performance or crime rates.
- Studying the impact of policy interventions or social programs on variables like unemployment or poverty levels.
- These are just a few examples, and regression analysis has numerous other applications in various fields. By identifying and quantifying the relationships between variables, regression helps in making informed decisions, optimizing processes, forecasting outcomes, and gaining insights into complex real-world phenomena.

Introduction

Understanding the relationship between a country's defense spending and its economic output (GDP) is crucial for policymakers and economists. This project aims to develop a linear regression model to analyze the relationship between India's defense budget and its GDP over the last 10 years, from 2014 to 2023.

Data

The data used in this project is sourced from the CSV file "newfile - Sheet1.csv," which contains yearly records of India's GDP (in USD trillions) and defense budget (in USD billions) from 2014

Year	Y GDP (in USD trillions)	X Defense Budget (in USD billions)
2014	2.04	37.4
2015	2.11	39.1
2016	2.3	45.2
2017	2.65	52.5
2018	2.72	55.9
2019	2.87	61.9
2020	2.87	65.9
2021	3.01	71.1
2022	3.17	76.7
2023	3.34	81.2

to 2023. The CSV file has the following structure:

Year	Y GDP (in USD trillions)	X Defense Budget (in USD billi	XiYi	Xi^2	Yi^2	xi=Xi-Xbar	yi=Yi-Ybar	xi^2	yi^2	xiyi	ei	ei^2	Yi hat
2014	2.04	37.4	76.296	1398.76	4.1616	-21.29	-0.668	453.2641	0.446224	14.22172	-0.06583639334	0.004334430688	2.105
2015	2.11	39.1	82.501	1528.81	4.4521	-19.59	-0.598	383.7681	0.357604	11.71482	-0.04391897349	0.001928876232	2.153
2016	2.3	45.2	103.96	2043.04	5.29	-13.49	-0.408	181.9801	0.166464	5.50392	-0.0264505846	0.0006996334257	2.326
2017	2.65	52.5	139.125	2756.25	7.0225	-6.19	-0.058	38.3161	0.003364	0.35902	0.1170771595	0.01370706127	2.532
2018	2.72	55.9	152.048	3124.81	7.3984	-2.79	0.012	7.7841	0.000144	-0.03348	0.09091199918	0.008264991595	2.629
2019	2.87	61.9	177.653	3831.61	8.2369	3.21	0.162	10.3041	0.026244	0.52002	0.07120877513	0.005070689656	2.798
2020	2.87	65.9	189.133	4342.81	8.2369	7.21	0.162	51.9841	0.026244	1.16802	-0.04192670756	0.001757848807	2.911
2021	3.01	71.1	214.011	5055.21	9.0601	12.41	0.302	154.0081	0.091204	3.74782	-0.04900283507	0.002401277845	3.059
2022	3.17	76.7	243.139	5882.89	10.0489	18.01	0.462	324.3601	0.213444	8.32062	-0.04739251085	0.002246050085	3.217
2023	3.34	81.2	271.208	6593.44	11.1556	22.51	0.632	506.7001	0.399424	14.22632	-0.004669928884	0.00002180823578	3.344
SUMMATION (∑) :	**27.08**	**586.9**	**1649.074**	**36557.63**	**75.063**	**0**	**0**	**2112.469**	**1.73036**	**59.7488**	**0**	**0.04043266784**	
AVERAGE:	**2.708**	**58.69**	**164.9074**										

SRF=	
b1= Ybar-b2*Xbar	**1.04801963**
b2= ∑xiyi/∑xi^2	**0.02828387067**
Yi hat=b1+b2*Xi	

eiXi	eiYi hat	eixi
-2.462281111	-0.1386406731	1.401656814
-1.717231863	-0.09459791029	0.8603726906
-1.195566424	-0.06153597801	0.3568183863
6.146550872	0.2965474113	-0.7247076172
5.081980754	0.2390156462	-0.2536444777
4.407823181	0.199298495	0.2285801682
-2.762970028	-0.1220874995	-0.3022915615
-3.484101574	-0.1498998114	-0.6081251832
-3.635005582	-0.1524803095	-0.8535391204
-0.3791982254	-0.01561937071	-0.1051200992
0	**0**	**0**

Methodology

The analysis was performed using the R programming language. The steps followed to build the linear regression model are:

1.Import the CSV file into R using the read.csv() function.

2.Preprocess the data by renaming columns to remove special characters.

3.Identify the column indices for GDP (dependent variable) and defense budget (independent variable).

4.Fit the linear regression model using the lm() function, with GDP as the response variable and defense budget as the predictor variable.

Results

The summary of the fitted linear regression model is as follows:

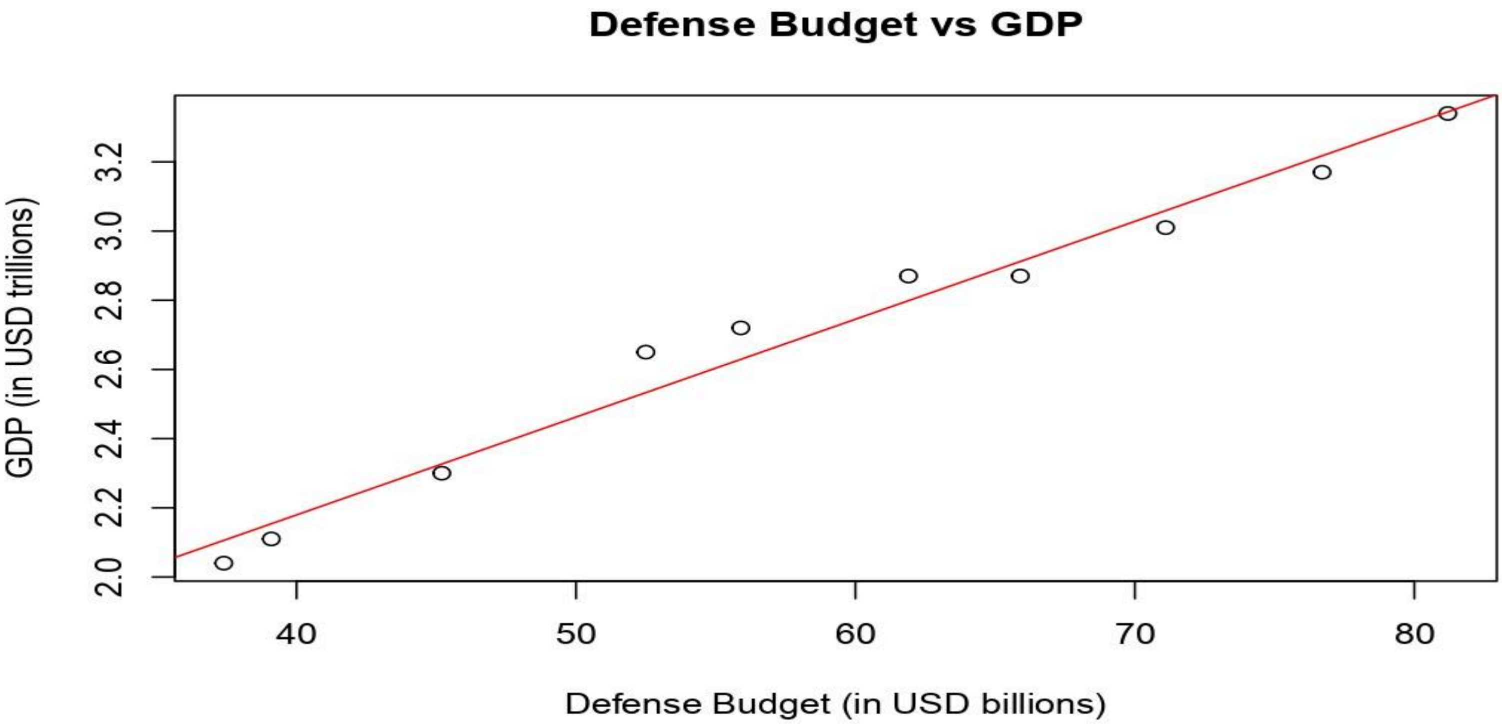
```
Call:
lm(formula = data[, y_col] ~ data[, x_col], data = data)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.06584 -0.04652 -0.03419 0.05224 0.11708 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  1.048020   0.093522  11.21  3.61e-06 ***
data[, x_col] 0.028284   0.001547  18.29  8.23e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.07109 on 8 degrees of freedom
Multiple R-squared:  0.9766,    Adjusted R-squared:  0.9737 
F-statistic: 334.4 on 1 and 8 DF,  p-value: 8.229e-08
```

Scatterplot

The scatter plot visualizing the data points and the fitted regression line is shown below:

```
plot(data[, x_col], data[, y_col],
     xlab = "Defense Budget (in USD billions)",
     ylab = "GDP (in USD trillions)",
     main = "Defense Budget vs GDP")
abline(model, col = "red")
```

Scatter plot of defense budget vs GDP with fitted regression line



Interpretation

The linear regression model reveals a positive and statistically significant relationship between India's defense budget and its GDP. The estimated slope coefficient of 0.028284 indicates that for every one billion USD increase in the defense budget, the GDP is expected to increase by approximately 0.028284 trillion USD, holding other factors constant.

The high values of R-squared (0.9766) and adjusted R-squared (0.9737) suggest that the linear model fits the data well and explains a significant portion of the variability in GDP based on the defense budget.

Predictions

Using the fitted linear regression model, predictions of GDP were made for hypothetical defense budget values of 60, 70, and 80 billion USD:

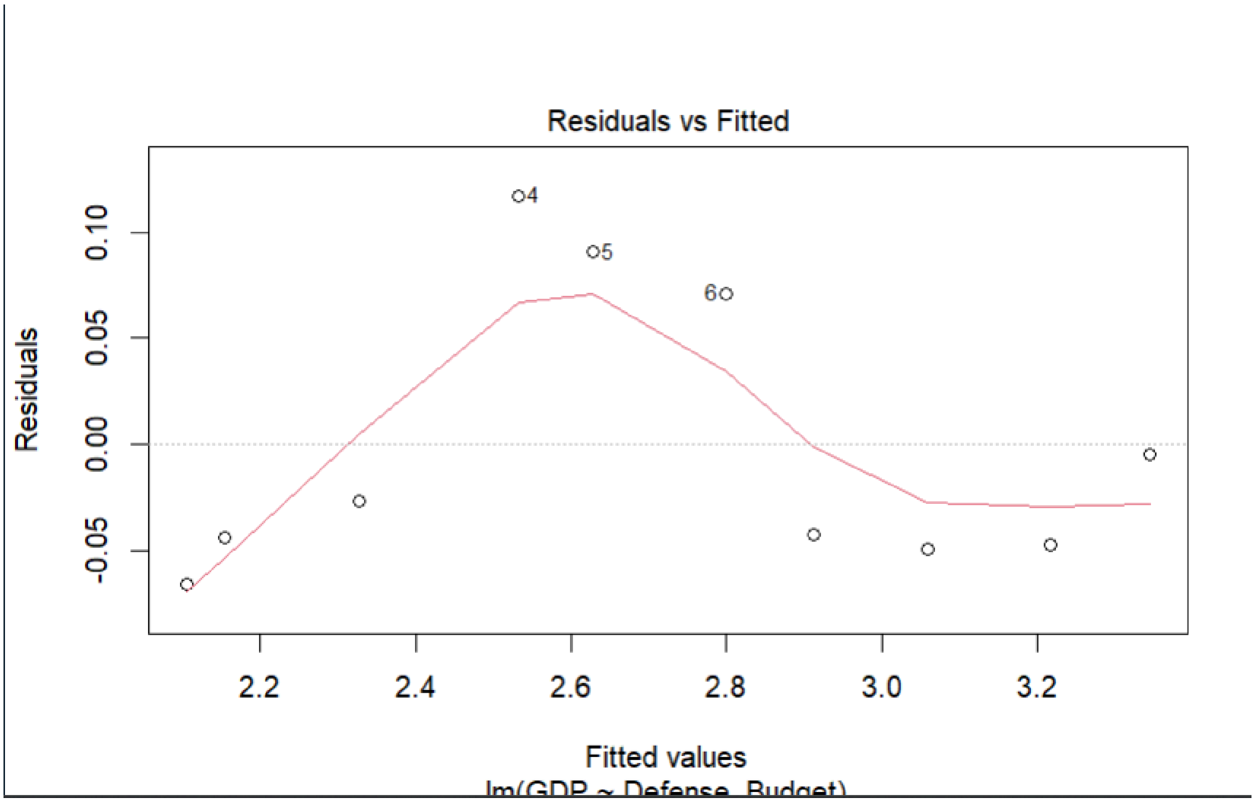
```
> # Make predictions for new defense budget values
> new_data <- data.frame(setNames(c(60, 70, 80), colnames[x_col]))
> predictions <- predict(model, newdata = new_data)
Warning message:
'newdata' had 3 rows but variables found have 10 rows
> print(predictions)
       1        2        3        4        5        6        7        8        9       10
2.105836 2.153919 2.326451 2.532923 2.629088 2.798791 2.911927 3.059003 3.217393 3.344670
```

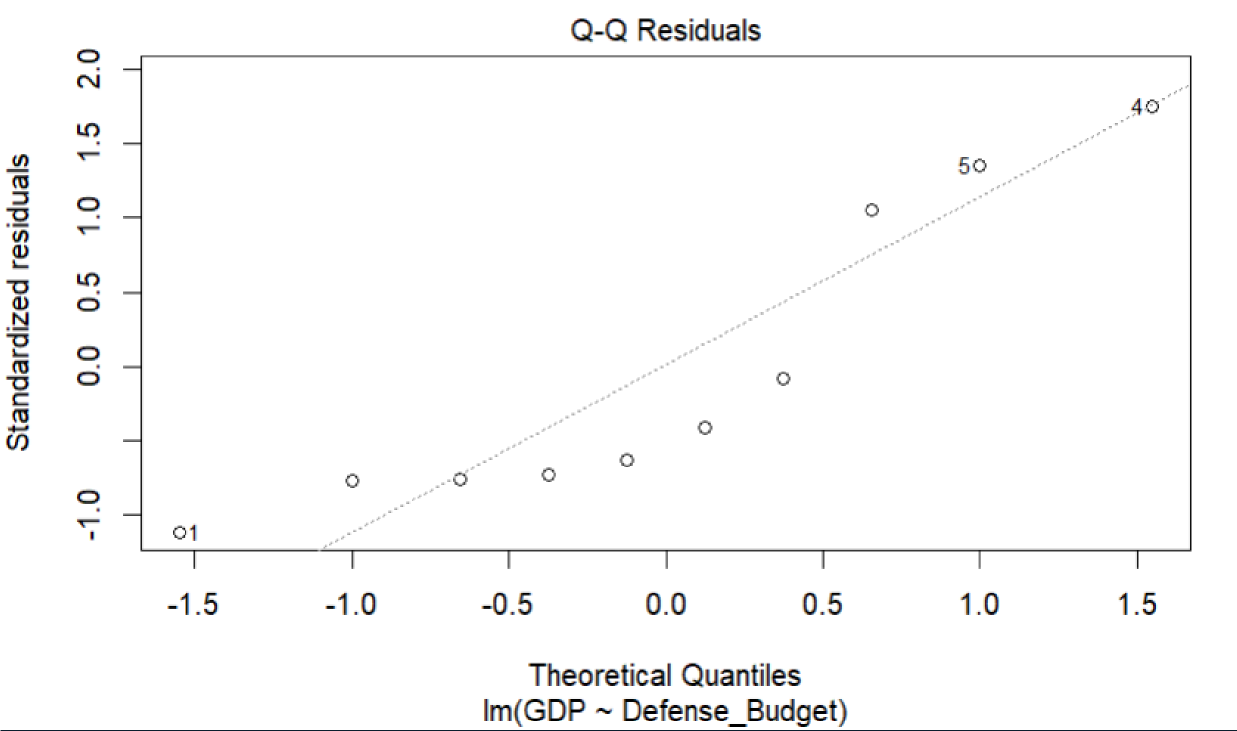
These predictions provide insights into how India's GDP might change in response to different levels of defense spending.

Residual Plots:

The residual plots generated by plot(model, which = c(1, 2)) are used to assess the assumptions of linearity and homoscedasticity (constant variance of residuals) in the linear regression model.

Residuals vs. Fitted Values Plot: This plot displays the residuals (the difference between the observed and predicted values) on the y-axis and the fitted values (predicted values from the regression model) on the x-axis. Ideally, the residuals should be randomly scattered around the horizontal line at 0 with no discernible pattern. If there is a curved or funnel-shaped pattern, it may indicate a violation of the linearity assumption or the presence of heteroscedasticity.
Normal Q-Q Plot: The Normal Q-Q (Quantile-Quantile) plot compares the distribution of the residuals to a normal distribution. If the residuals follow a normal distribution, the points should roughly align along the straight diagonal line. Significant deviations from this line may indicate a violation of the normality assumption for the residuals.
Assumption Tests:


Residuals vs Fitted
Residuals
0.10
0.05
0.00
-0.05
4
5
6
2.2
2.4
2.6
2.8
3.0
3.2
Fitted values
lm(GDP ~ Defense_Budget)


Q-Q Residuals
Standardized residuals
2.0
1.5
1.0
0.5
0.0
-1.0
4
5
1
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
Theoretical Quantiles
lm(GDP ~ Defense_Budget)

Model Summary:

The summary(model) output provides a comprehensive summary of the fitted linear regression model, including:

```
Call:
lm(formula = GDP ~ Defense_Budget, data = data)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.06584 -0.04652 -0.03419  0.05224  0.11708 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
(Intercept)    1.048020   0.093522   11.21 3.61e-06 ***
Defense_Budget 0.028284   0.001547   18.29 8.23e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.07109 on 8 degrees of freedom
Multiple R-squared:  0.9766,	Adjusted R-squared:  0.9737 
F-statistic: 334.4 on 1 and 8 DF,  p-value: 8.229e-08
```

```
> # Confidence intervals and hypothesis testing
> confint(model)  # Confidence intervals for coefficients
                    2.5 %     97.5 %
(Intercept)    0.83235688 1.26368238
Defense_Budget 0.02471701 0.03185073
> summary(model)$coefficients  # t-tests for coefficients
                 Estimate  Std. Error  t value     Pr(>|t|)
(Intercept)    1.04801963 0.093522272 11.20610 3.605579e-06
Defense_Budget 0.02828387 0.001546771 18.28575 8.228902e-08
```

Coefficient estimates (intercept and slope)
Standard errors of the coefficient estimates
t-values and p-values for testing the significance of the coefficients
Residual standard error
Multiple R-squared (coefficient of determination)
Adjusted R-squared
F-statistic and p-value for the overall model significance
This information helps in interpreting the strength and statistical significance of the relationship between the predictor variable (defense budget) and the response variable (GDP), as well as assessing the overall fit of the linear regression model.

Conclusion

This linear regression analysis establishes a strong positive relationship between India's defense budget and its GDP over the last 10 years. The model suggests that an increase in defense spending is associated with a corresponding increase in economic output. However, it's important to note that this analysis considers only the direct relationship between these two variables and does not account for other potential factors influencing GDP.

The positive association between defense spending and GDP can be attributed to various factors. Increased defense expenditure can stimulate economic activity through government procurement of goods and services from domestic industries, creating employment opportunities and driving production. Additionally, investments in defense research and development may lead to technological advancements and innovation, which can have spillover effects on other sectors of the economy.

However, it is crucial to understand that excessive defense spending may divert resources away from other productive sectors, such as education, healthcare, and infrastructure development, which are essential for long-term economic growth. Policymakers must strike a balance between maintaining national security and allocating resources effectively to promote sustainable economic development.

Moreover, the relationship between defense spending and GDP may not be linear across all levels of expenditure, and there could be diminishing returns beyond a certain threshold. Future research could explore the potential nonlinear or time-lagged effects of defense spending on GDP to gain a more nuanced understanding of this relationship.

It is also important to consider the broader geopolitical context and the evolving security landscape when analyzing defense expenditure. Changes in regional power dynamics, emerging threats, and international treaties or agreements may influence a country's defense spending priorities, which could impact the relationship with economic output.

Future research could explore the inclusion of additional variables, such as government spending in other sectors, foreign trade, demographic factors, and geopolitical indicators, to develop a more comprehensive model for predicting India's economic growth. Incorporating these factors could enhance the explanatory power of the model and provide policymakers with a more holistic understanding of the drivers of economic growth.

Furthermore, conducting similar analyses for other countries or regions could shed light on potential variations in the relationship between defense spending and GDP across different economic and political contexts. Cross-country comparisons could yield valuable insights and inform policy decisions related to defense budgeting and economic planning.

R CODE-

```
# Import data
data <- read.csv(text = "Year,Y GDP (in USD trillions),X Defense
Budget (in USD billions)
2014,2.04,37.4
2015,2.11,39.1
2016,2.3,45.2
2017,2.65,52.5
2018,2.72,55.9
2019,2.87,61.9
2020,2.87,65.9
2021,3.01,71.1
2022,3.17,76.7
2023,3.34,81.2", header = TRUE)

# Rename columns
colnames(data) <- c("Year", "GDP", "Defense_Budget")

# Data exploration and preprocessing
summary(data)
str(data)
cor(data$GDP, data$Defense_Budget)  # Correlation

# Build linear regression model
model <- lm(GDP ~ Defense_Budget, data = data)
```

```
summary(model)

# Assumptions and diagnostics
plot(model, which = c(1, 2))  # Residual plots
bptest(model)  # Heteroscedasticity test
dwtest(model)  # Autocorrelation test

# Outlier and influence analysis
influenceIndexPlot(model)  # Influence plot
outlierTest(model)  # Outlier test

# Confidence intervals and hypothesis testing
confint(model)  # Confidence intervals for coefficients
summary(model)$coefficients  # t-tests for coefficients

# Model validation and generalization
data_train <- data[1:8, ]  # Training data
data_test <- data[9:10, ]  # Test data
model_train <- lm(GDP ~ Defense_Budget, data = data_train)
predictions <- predict(model_train, newdata = data_test)
mse <- mean((data_test$GDP - predictions)^2)  # Mean squared
error
```

GROUP 8- YASH SAINI(12719)

MIHIT JAIN(12711)

ADITYA RATHORE(12744)

DATA GATHERED FROM-

1.Ministry of Finance, Government of India:
Website: https://www.indiabudget.gov.in/

2.Reserve Bank of India (RBI):
Website: https://www.rbi.org.in/

3.Ministry of Defence, Government of India:
Website: https://www.mod.gov.in/

4.Stockholm International Peace Research Institute (SIPRI):
Website: https://www.sipri.org/databases

SOFTWARE USED- R-STUDIO, CLAUDE AI, GOOGLE SHEETS.

Group-6

Submitted to: Dr. Geeta Aggarwal

Submitted by:
Sayesha Ali Khan, 21099
Khalid Hanief, 21444

CINEMASCOPE
A CINEMATIC EXPERIENCE UNLEASHED

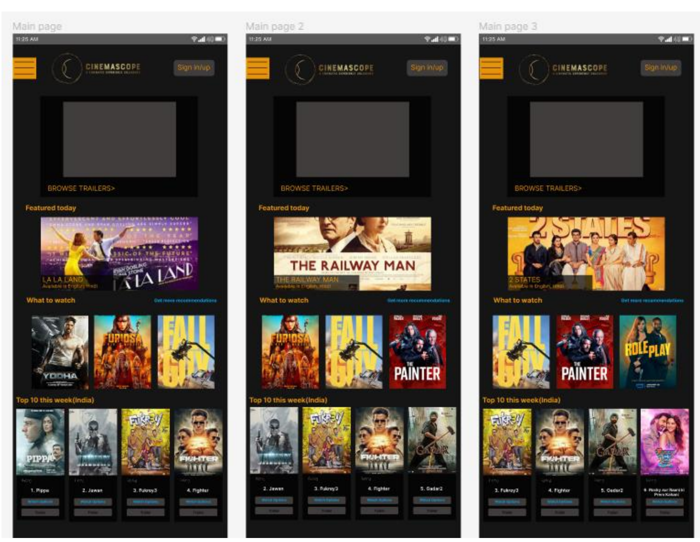
REPORT FOR THE MONTH OF FEBRUARY

CinemaScope is a groundbreaking movie information and recommendation app that seamlessly integrates the visual appeal of Figma with the powerful functionality of Python and MySQL for a comprehensive and immersive cinematic experience. This unique fusion of design and technology aims to redefine the way users explore, discover, and engage with movies.

Week 1:

In Figma,

Introduction: The design process of the Cinemascope app aimed to create an intuitive and visually appealing user interface for a movie-centric application. The project included the creation of a Homepage, Login frames for both Users and Admins, and the incorporation of animations to enhance user engagement.

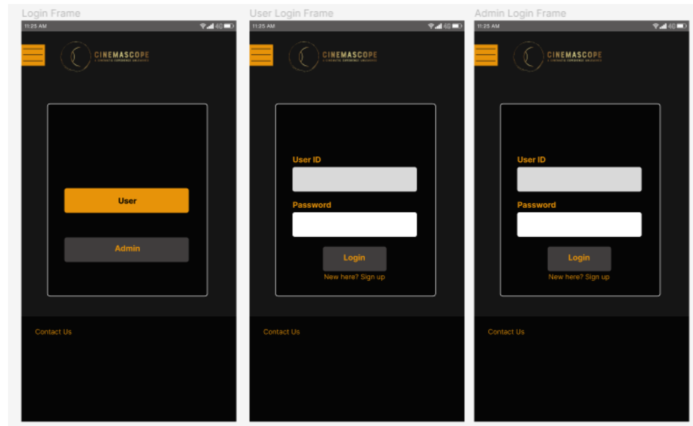


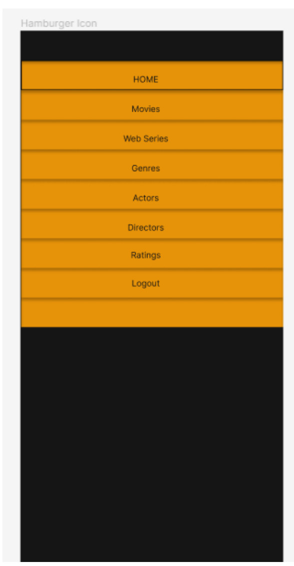
Research and Planning: Before diving into design, extensive research was conducted to understand user preferences, industry trends, and the specific needs of a cinema-related app.

Homepage Design:
The homepage was designed to be visually striking and user-friendly.
Key elements included movie banners, a search bar, and featured categories for easy navigation. The color scheme was chosen to evoke a cinematic atmosphere, utilizing dark backgrounds and vibrant accents.

Login Frames:

- **User Login:**
  A simple and secure login frame was created for users.
- **Admin Login:**
  The admin login frame featured additional security measures and authentication processes. A clean and professional design was maintained to reflect the administrative nature of the account.





Hamburger icon:
In order to create the menu, hamburger icon was introduced that consisted of the options of Homepage, Movies page, Web series page, Genres, Actors page, Directors page and Ratings page.

Animation Integration in Figma:
Figma was utilized to implement animations seamlessly into the design.
Micro-interactions were added to buttons, providing visual feedback upon user interaction.
Page transitions and loading animations were designed to enhance the overall user experience.

In MySQL,

Ensured establishing proper relationship between tables, setting primary keys, foreign keys and considered normalization for efficient data storage.

Creation of Database:

Code-

Output-

Creation of Tables:

Code-

Output-

Insertion of Data in Tables:

Code-

Output-

<u>Goals for next month</u>*:*

1)Personalized Recommendations:

Utilize Python algorithms to analyze user preferences, viewing history, and ratings to generate personalized movie recommendations. Boosting user experience by suggesting films tailored to individual tastes.

3)Creation of other sections:

Use Figma to create further pages such as Movies, Genres, Actors, Directors etc.

2)Advanced Search and Filtering:

Implement advanced search functionalities allowing users to find movies based on various criteria such as genre, release year, director, and actors. Incorporate smart filtering options for a more refined and tailored search experience.

4)Integration of MySQL and Figma API

Integrate the MySQL database with the Figma interface for synchronization of data.

Project Report

FITNESS-FORGE

Nishant Jain | Group 5 | 26 February 2024 | Computer Applications

Fitness-Forge Monthly Report - February 2024

1. Executive Summary:

- During February, significant progress was made in the development of Fitness-Forge. The Home Page and user authentication components were successfully designed, showcasing a visually appealing and user-friendly interface.

2. Web Design Overview:

HOME PAGE (INDEX.HTML):

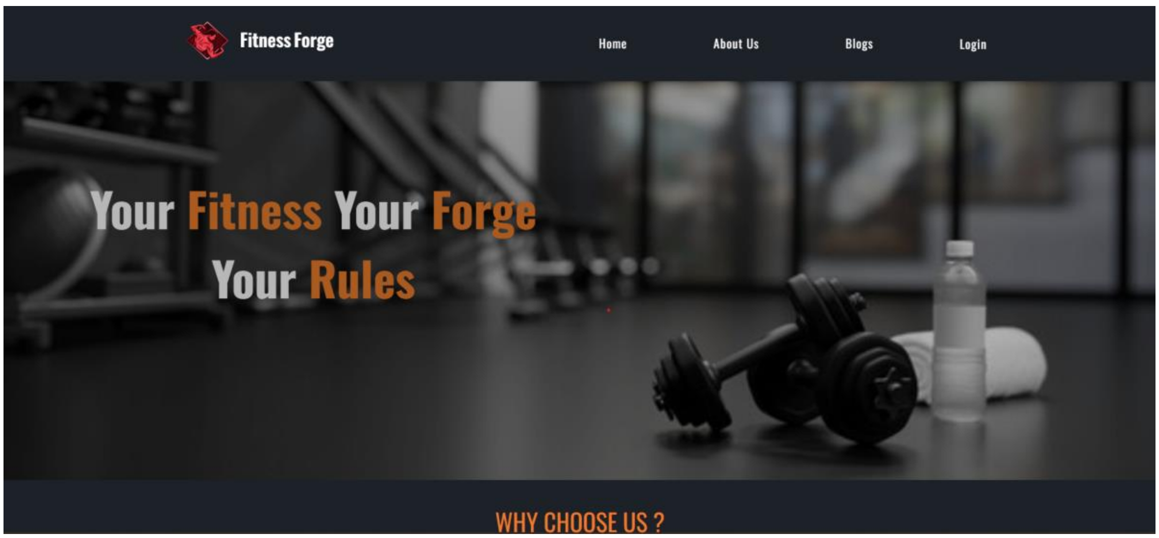
- The Home Page serves as the gateway to FitnessForge, offering an intuitive and engaging interface. Key features include dynamic buttons for user registration and login, crafted using HTML and CSS.

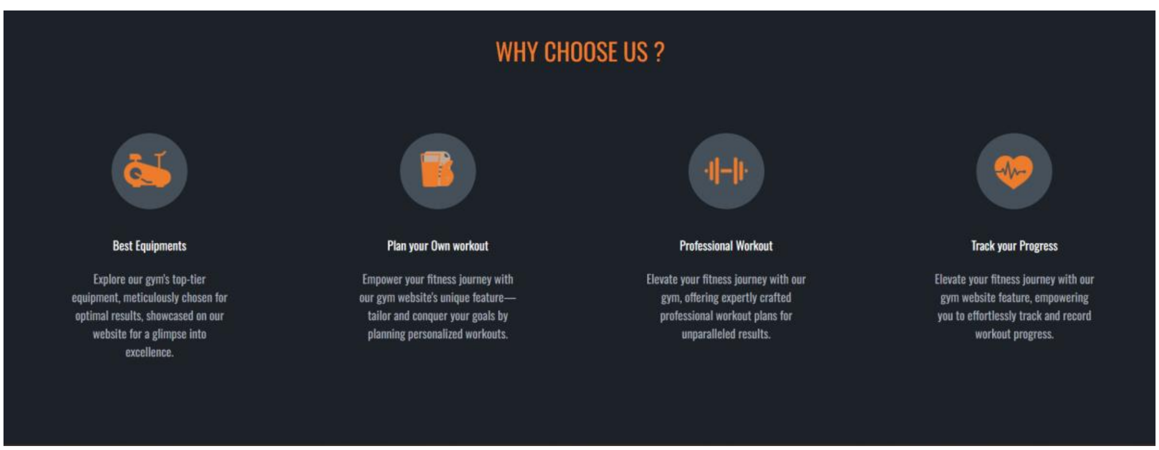
LOGIN/REGISTRATION FORM:

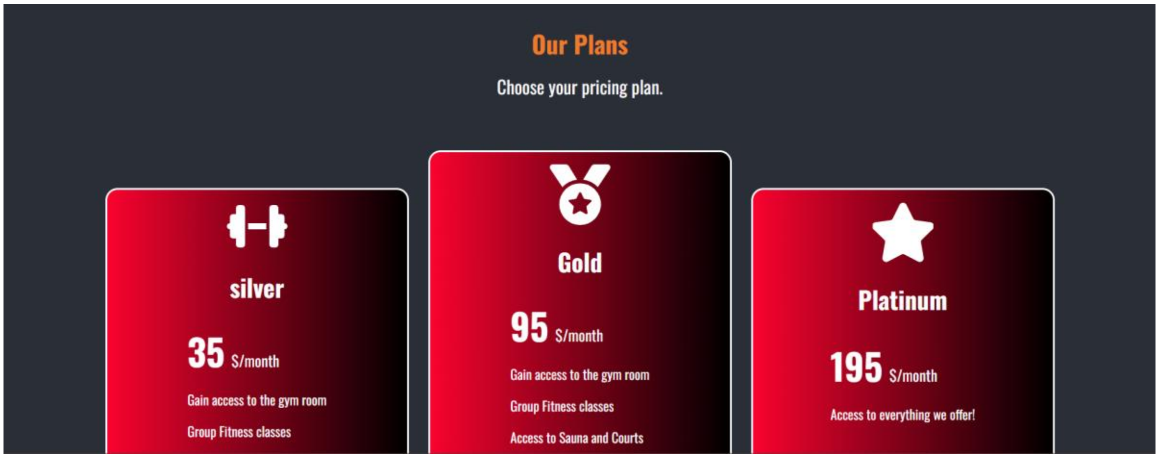
- A seamless and responsive login/registration form was created, enhancing user experience. HTML and CSS were leveraged to ensure a clean design, with Font Awesome icons incorporated for visual appeal.

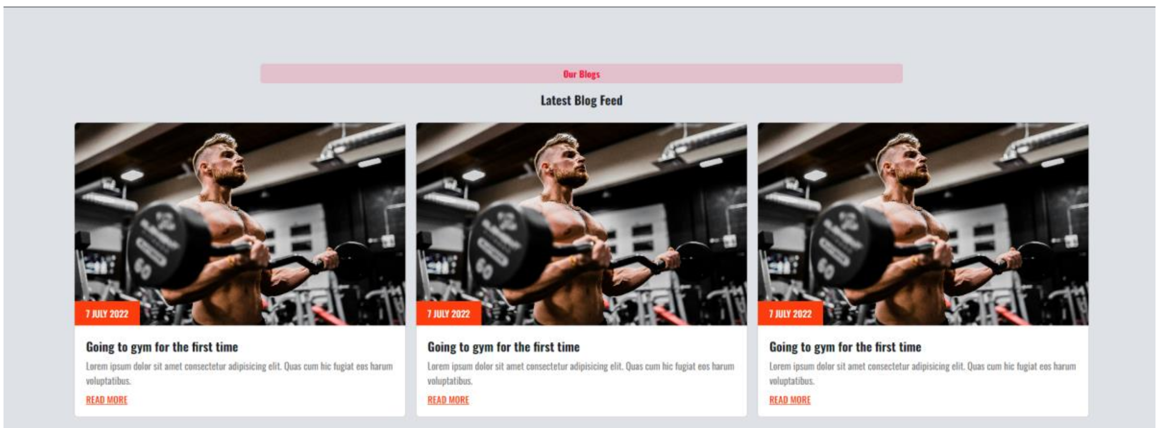
FONT AWESOME ICONS:

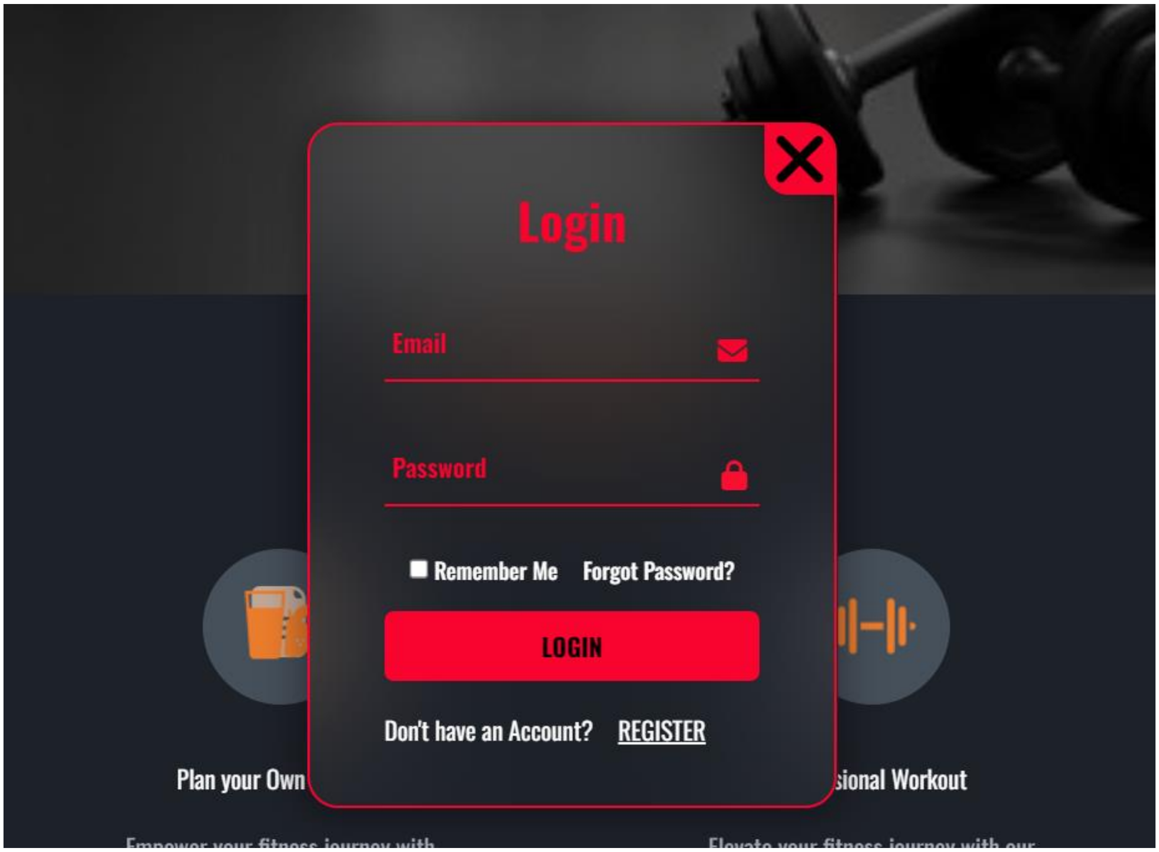
- Font Awesome icons were strategically integrated to enhance the aesthetic appeal of the website, contributing to a modern and visually appealing design.

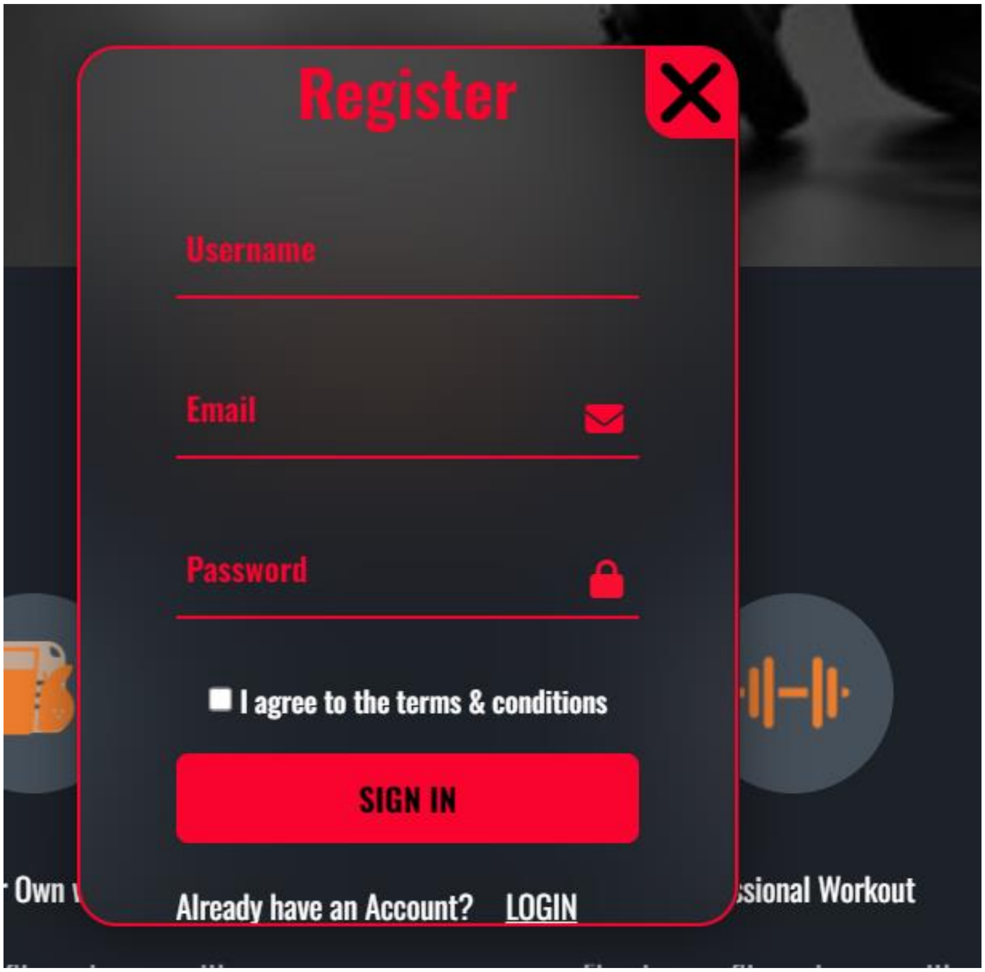





Our Plans
Choose your pricing plan.
silver
35 $/month
Gain access to the gym room
Group Fitness classes
Gold
95 $/month
Gain access to the gym room
Group Fitness classes
Access to Sauna and Courts
Platinum
195 $/month
Access to everything we offer!


Our Blogs
Latest Blog Feed
7 JULY 2022
Going to gym for the first time
Lorem ipsum dolor sit amet consectetur adipisicing elit. Quas cum hic fugiat eos harum voluptatibus.
READ MORE
7 JULY 2022
Going to gym for the first time
Lorem ipsum dolor sit amet consectetur adipisicing elit. Quas cum hic fugiat eos harum voluptatibus.
READ MORE
7 JULY 2022
Going to gym for the first time
Lorem ipsum dolor sit amet consectetur adipisicing elit. Quas cum hic fugiat eos harum voluptatibus.
READ MORE


Login
Email
Password
Remember Me
Forgot Password?
LOGIN
Don't have an Account?
REGISTER
Plan your Own
sional Workout



3. Development Roadmap:

UPCOMING PAGES:

➢ The development roadmap includes the creation of additional pages such as the Physical Details Form, Workout Plans Section, Payment Gateway, and User Dashboard. These pages will be crafted with a focus on responsive design and optimal user experience.

JAVASCRIPT FUNCTIONALITY:

➢ The next phase involves implementing JavaScript functionality to enhance the interactivity of Fitness-Forge. This includes form validation, dynamic content updates, and seamless user interactions.

5. Next Steps:

COMPLETION OF WEB PAGES:

- Proceeding with the development of the remaining web pages, ensuring consistency in design and functionality.

JAVASCRIPT INTEGRATION:

- Initiating the integration of JavaScript functionality to create a dynamic and responsive user interface.

6. Conclusion:

- The progress made in February lays a strong foundation for the Fitness-Forge project. The combination of meticulous design, user-friendly interfaces, and planned JavaScript integration sets the stage for a robust and engaging fitness platform.