# From Foundations to Futures: A State wise Study of Rural Health Infrastructure and Infant Mortality (2014-2020) in India

Seminar paper presented

by

**Vishakha Gupta**

(Indian Economic Service, 2024)

Under the Supervision of

**Prof. Oindrila De**


**INSTITUTE OF ECONOMIC GROWTH**

**Delhi - 110007, India**

February, 2025

**ABSTRACT**

This study examines the relationship between rural health infrastructure and infant mortality rates (IMR) in 21 major Indian states from 2014 to 2020. Using Principal Component Analysis (PCA), we construct a Rural Health Infrastructure Index (RHII) to assess the impact of health centers and personnel density on IMR. Panel regression techniques, including fixed and random effects models, are employed to estimate this relationship. The findings indicate that stronger health infrastructure significantly reduces IMR, with regional disparities persisting. Additionally, the study evaluates the impact of COVID-19 on these trends and examines the role of access to clean drinking water as a moderating factor. The research provides crucial insights for evidence-based policymaking aimed at improving rural healthcare systems.

Keywords: Rural Health Infrastructure, Infant Mortality Rate, Panel Regression, PCA, COVID-19, Clean Drinking Water

JEL Codes: I18, J13, C33

## INTRODUCTION

Infant mortality, often regarded as a key indicator of public health and socio-economic development, reflects the effectiveness of healthcare systems, maternal health services, and broader social determinants of health. Despite substantial progress over the past few decades, India continues to face challenges in reducing infant mortality, particularly in rural areas where healthcare access remains constrained. While the national infant mortality rate (IMR) declined from 39 per 1,000 live births in 2014 to 28 per 1,000 live births in 2020 (SRS, 2021), disparities between states remain stark. Rural IMR continues to be significantly higher than urban IMR, underscoring deep-seated inequities in healthcare provision.

From a theoretical perspective, health plays a central role in economic development. Theories of endogenous growth (Lucas, 1988; Romer, 1990) emphasize that investments in human capital including healthcare enhance productivity, improve labor market outcomes, and contribute to long-term economic growth. The Grossman (1972) model of health capital further posits that better health increases an individual's productivity and earnings, reinforcing the need for strong public health interventions. Amartya Sen's capability approach (1999) adds another dimension, arguing that access to healthcare is not merely an economic variable but a fundamental capability that enables individuals to lead a meaningful life. Without robust healthcare infrastructure, large segments of the rural population remain deprived of essential health services, perpetuating cycles of poverty and ill-health.

Recognizing the importance of healthcare for national development, India has launched several initiatives aimed at strengthening rural health infrastructure. The National Rural Health Mission (NRHM, 2005) later integrated into the National Health Mission (NHM, 2013) sought to expand the reach of primary healthcare services, improve maternal and child health, and enhance the availability of healthcare personnel in rural areas. More recently, the Ayushman Bharat initiative (2018) has sought to transform healthcare delivery through the establishment of Health and Wellness Centres (HWCs) now known as Jan Arogya Mandir and the Pradhan Mantri Jan Arogya Yojana (PMJAY), which provides financial protection for secondary and tertiary healthcare. While these interventions have contributed to improving health indicators, the effectiveness of rural health infrastructure remains uneven across states, with gaps in availability, accessibility, and quality of healthcare services.

At a global level, Sustainable Development Goal 3 (SDG-3) aims to reduce preventable child deaths, with a target of lowering neonatal mortality to at least 12 per 1,000 live births and under-five mortality to 25 per 1,000 live births by 2030. Countries that have successfully achieved low IMR—such as Thailand, Cuba, and Sri Lanka—have done so by prioritizing primary healthcare expansion, strengthening rural health workforce, and investing in maternal and child health

programs. For instance, Thailand's Universal Coverage Scheme (UCS), implemented in the early 2000s, led to a significant decline in infant and child mortality by ensuring access to healthcare at the village level. Cuba's family doctor program, which integrates preventive and curative services at the community level, has been another model of success. Lessons from these countries highlight the need for India to move beyond infrastructure expansion and focus on systemic improvements in healthcare quality, governance, and financing mechanisms.

## LITERATURE REVIEW

The relationship between health infrastructure and infant mortality rates (IMR) has been extensively studied in India, with research focusing on various determinants, methodologies, and regional disparities. Existing studies indicate that despite sustained policy interventions, infant mortality remains a critical challenge, particularly in rural areas. This review synthesizes insights from key contributions to the field and identifies gaps that inform the direction of the present study.

Suriyakala et al. (2016) conducted a regression-based analysis to understand factors influencing IMR in India. Their findings suggest that higher female literacy, increased government health expenditure, and better availability of doctors and government hospitals significantly reduce IMR. The study highlights regional disparities in IMR reduction, emphasizing that policy measures have yielded uneven benefits across states. However, the study does not extend beyond 2011, limiting its ability to capture recent health sector transformations, including those prompted by the COVID-19 pandemic. Similarly, Lakshmi and Sahoo (2013) analyzed health infrastructure in Andhra Pradesh using double-log regression and PCA. Their findings reinforce the significance of hospitals, dispensaries, and doctors in improving key health indicators. However, the study does not differentiate between rural and urban settings, leaving a gap in understanding how rural-specific health infrastructure influences IMR. This is particularly relevant to the present study, which seeks to construct a composite index for rural health infrastructure and analyze its impact on IMR.

Arun et al. (2018) employed an econometric approach (OLS and K-means clustering) to explore the role of health service infrastructure in reducing IMR across Indian states. Their study indicates that the availability of Community Health Centers (CHCs) is significantly associated with lower IMR. This finding aligns with the present study's objective of assessing the impact of rural health infrastructure components, such as CHCs and Sub-Centers (SCs), on infant mortality.

Beyond IMR, some studies examine the role of health infrastructure in broader economic contexts. Pradhan et al. (2011) assessed the relationship between health infrastructure and economic growth through input-output analysis. Their study found a long-term cointegration between health indicators (IMR, life expectancy rate) and economic growth, suggesting that investments in health

infrastructure have significant macroeconomic implications. However, the study lacks a focus on regional variations, making it difficult to discern how rural and urban disparities shape health outcomes. Understanding these regional differences is essential for policy interventions aimed at strengthening rural health services.

The issue of regional disparities is further explored by Lyngdoh (2015), who analyzed interstate variations in rural health infrastructure across North-Eastern states using PCA. The study provides a ranking of states based on infrastructure availability but does not examine its direct impact on IMR. This limitation is addressed in the present study, which seeks to construct an index of rural health infrastructure and statistically link it with IMR across states.

The methodologies employed in the reviewed studies vary significantly. While Suriyakala et al. (2016) used simple regression models, Pradhan et al. (2011) applied advanced time-series techniques such as the Granger causality test and Johansen MLE. PCA has been used to construct health infrastructure indices (Lyngdoh, 2015; Lakshmi & Sahoo, 2013), though these indices have not been directly linked to IMR. Arun et al. (2018) adopted K-means clustering to group states based on health service availability, which provides an alternative approach to identifying disparities. Given these methodological trends, the present study employs PCA to construct the Rural Health Infrastructure Index (RHII), which will be used in panel regression models to examine its impact on IMR. By incorporating pooled OLS and fixed/random effects models, this study seeks to provide a more robust statistical analysis than previous studies.

While existing studies highlight the importance of health infrastructure, they exhibit several limitations that the present study aims to address. First, most studies analyze health infrastructure at a broad state level rather than focusing specifically on rural health facilities, which are crucial for IMR reduction. Second, while PCA has been used to rank infrastructure, prior research does not construct a composite index aggregating multiple health infrastructure indicators to examine their collective impact on IMR. Third, the reliance on cross-sectional or time-series methods does not fully account for state-level heterogeneity, which panel regressions can address. Fourth, many studies use data only until 2011, missing key healthcare policy changes and the potential impact of COVID-19. The present study, covering 2014–2020, provides updated insights. Finally, manpower shortages are often overlooked in the literature, despite their critical role in healthcare delivery. By incorporating both physical infrastructure (hospitals, CHCs, SCs) and manpower indicators (doctors, nurses), the present study offers a more comprehensive analysis.

In conclusion, the literature suggests a strong relationship between health infrastructure and IMR but leaves gaps in rural-specific analysis, composite index construction, and methodological robustness. By constructing an RHII and applying panel data techniques, the present study builds

on previous findings while addressing these limitations. The insights generated will contribute to policy discussions on strengthening rural health services to improve infant survival rates in India.

## RESEARCH OBJECTIVES AND QUESTIONS

This study aims to examine the relationship between rural health infrastructure and infant mortality rates (IMR) across Indian states by constructing a composite Rural Health Infrastructure Index (RHII). It seeks to analyze whether this relationship changed during the COVID-19 pandemic, given the disruptions in healthcare service delivery. Additionally, the study investigates whether the impact of health infrastructure on IMR varies across Empowered Action Group (EAG) states, which historically have weaker health systems compared to non-EAG states. Finally, the research explores the indirect role of clean drinking water access in reducing IMR, particularly assessing whether health infrastructure mediates this effect.

Accordingly, the study addresses the following research questions:

- Is rural health infrastructure significantly correlated with rural IMR?
- Did the relationship between rural health infrastructure and IMR change during the COVID-19 pandemic?
- Does the effect of rural health infrastructure on IMR differ between EAG and non-EAG states?
- What is the indirect impact of clean drinking water access on IMR, and does health infrastructure mediate this relationship?

## METHODOLOGY

### DATASETS & VARIABLES

This study employs a panel dataset (2014–2020) to analyze the relationship between rural health infrastructure and infant mortality rates (IMR) across Indian states. Data is sourced from official reports, including the Rural Health Statistics (RHS), National Family Health Survey (NFHS), Sample Registration System (SRS), and National Sample Survey Organization (NSSO). The variables used in the analysis are categorized as follows:

Dependent Variable:

- Infant Mortality Rate (IMR): The number of infant deaths (children under one year of age) per 1,000 live births. IMR serves as a key indicator of child health and overall healthcare effectiveness in rural areas. Data is obtained from SRS and NFHS.

Independent Variables:

- Rural Health Infrastructure Index (RHII): A composite index constructed using Principal Component Analysis (PCA) to measure the availability and quality of rural healthcare infrastructure. The index includes:
    - Sub-Centers (SCs) per lakh rural population: SCs are the first point of contact in the rural healthcare system, providing basic maternal and child healthcare, immunization, and disease prevention services. Each SC is meant to serve 5,000 people in plains and 3,000 in hilly/tribal areas as per government norms.
    - Primary Health Centers (PHCs) per lakh rural population: PHCs serve as referral units for SCs and provide outpatient care, maternal and child health services, and minor surgical procedures. A PHC is typically meant to serve a population of 30,000 in plains and 20,000 in hilly/tribal areas and should be staffed with at least one medical officer and paramedical personnel.
    - Community Health Centers (CHCs) per lakh rural population: CHCs function as block-level referral centers with specialized services such as surgery, obstetrics, and gynecology. Each CHC is designed to serve 1,20,000 people in plains and 80,000 in hilly/tribal areas and should have at least four specialists (physician, surgeon, gynecologist, pediatrician).
- Health Workforce Availability: The number of doctors, nurses, and auxiliary nurse midwives (ANMs) per 100,000 rural population, representing the accessibility of healthcare personnel in rural areas. Data is extracted from RHS and NSSO.

Control Variable:

- Access to Improved Drinking Water: The percentage of rural households with access to safe and improved drinking water sources, which plays a crucial role in reducing waterborne diseases and improving child survival rates. Data is obtained from NFHS and NSSO.

## CONSTRUCTION OF RURAL HEALTH INFRASTRUCTURE INDEX (RHII) USING PCA AND IHS TRANSFORMATION

The Rural Health Infrastructure Index (RHII) is constructed using Principal Component Analysis (PCA) to create a composite measure of healthcare availability in rural India. It incorporates

the above-mentioned indicators representing healthcare facility density and workforce availability. To ensure comparability, all variables are standardized using Z-scores, and PCA is applied to extract the principal components (PCs). Only those with eigenvalues greater than 1 are retained, as they explain the most significant variance. The final index is computed as a weighted sum of selected PCs, with weights assigned based on their variance contributions:

$$RHII = w_1PC_1 + w_2PC_2 + \ldots\ldots + w_nPC_n$$

The resulting RHII is normalized to a 0–1 scale, where higher values indicate better rural health infrastructure. This approach ensures a statistically robust, data-driven measure for analyzing the impact of healthcare infrastructure on infant mortality rates.

Since some variables in the analysis may have skewed distributions or contain zero values, the Inverse Hyperbolic Sine (IHS) transformation is applied to improve normality and reduce heteroscedasticity. The IHS transformation is defined as:

$$\text{IHS (X)} = \ln\left( X + \sqrt{X2 + 1} \right)$$

This transformation is particularly useful when dealing with variables that include zero or small values, where the traditional log transformation is not applicable. Unlike log transformation, IHS allows for the inclusion of zero values without requiring arbitrary adjustments. By stabilizing variance and reducing skewness, the IHS transformation ensures more reliable regression estimates, making it suitable for econometric analysis involving health infrastructure and infant mortality.

**MODEL SPECIFICATION & TECHNIQUES**

The baseline model specification is as follows:

$\text{Rural\_IMR}_{it} = \alpha_i + \beta_1 \text{RHII}_{it} + \text{DrinkingWater}_{it} + \epsilon_{it}$

where, $\text{DrinkingWater}_{it}$ = Percentage of Rural Households having access to improved sources of drinking water

Rural_IMR = No. of deaths per 1000 live births

$\alpha_i$ = State fixed effects (to control for time-invariant state-specific factors)

$\varepsilon_{it}$ = Error term

We started our analysis with Pooled OLS, which treats all observations as a single cross-section, ignoring state-specific heterogeneity. While useful for initial estimation, it assumes no unobserved differences across states, leading to potential omitted variable bias. To address this, panel regression

techniques both Fixed Effects (FE) and Random Effects (RE) models are employed. The FE model controls for time-invariant state-specific factors, ensuring more reliable estimates, while the RE model assumes these factors are uncorrelated with regressors. The Hausman test rejects the RE specification, confirming that the FE model is the appropriate choice for addressing unobserved heterogeneity and providing consistent estimates. To ensure robustness, heteroskedasticity-consistent standard errors, placebo tests, and alternative model specifications are incorporated in subsequent sections. The analysis further examines regional disparities, particularly between Empowered Action Group (EAG) and non-EAG states, while assessing the moderating role of improved drinking water access.

## RESULTS

This section represents the descriptive statistics, regression results and robustness checks of our study.

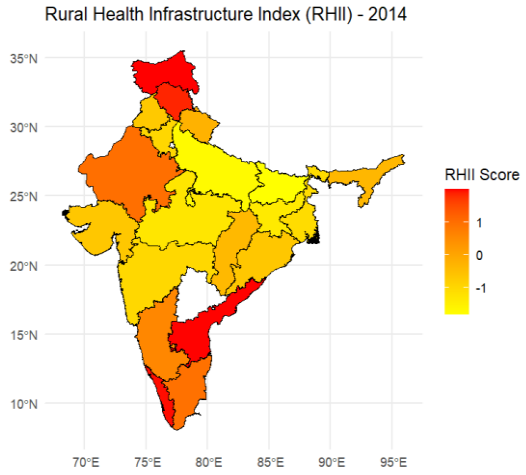
### DESCRIPTIVE STATISTICS

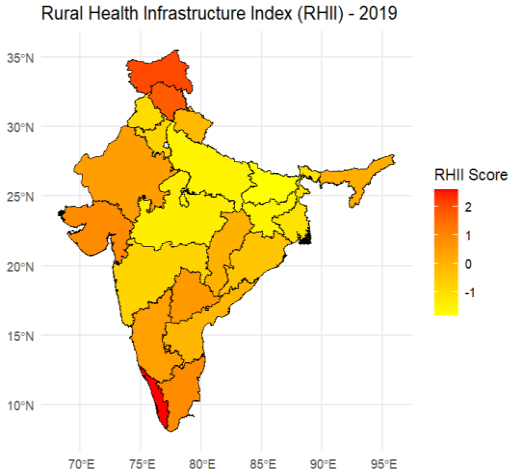
Summary statistics (Table 1) reveal significant regional disparities in rural health infrastructure and infant mortality across 21 Indian states. The average rural IMR stands at 32.38 per 1,000 live births, with values ranging from 4 to 57, highlighting substantial differences in child health outcomes. The availability of healthcare personnel varies widely, with doctors density averaging 4.15 per 100,000 rural population, while nurses and ANM densities stand at 26.56 and 9.71, respectively. The distribution of healthcare facilities also exhibits disparities, with sub-center (SC) density averaging 21.35, primary health center (PHC) density at 3.65, and community health center (CHC) density at just 0.77 per 100,000 population, indicating limited access to secondary healthcare. Despite an overall high level of improved drinking water access (90.28%), some states still fall behind, with coverage as low as 66.33%. These patterns suggest that states with weaker health infrastructure and lower medical personnel density tend to experience higher infant mortality rates, emphasizing the need for targeted interventions to strengthen rural healthcare systems.

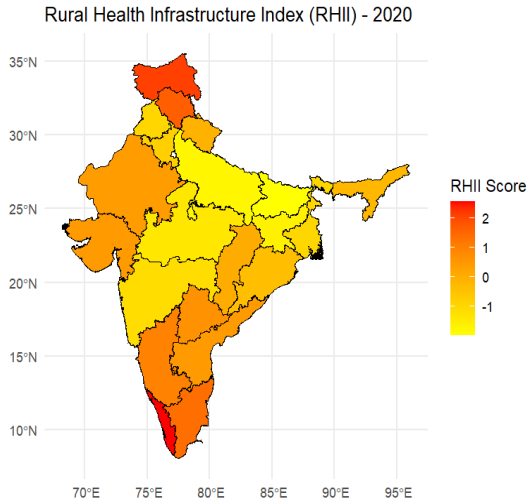
**Table 1: Summary Statistics**

| Variables | Mean | Std.Dev. | Min | Max |
|---|---|---|---|---|
| Rural IMR | 32.377 | 11.084 | 4 | 57 |
| Doctors density | 4.145 | 2.635 | 0.787 | 13.489 |
| Nurses density | 26.558 | 8.281 | 12.344 | 60.01 |
| ANM density | 9.714 | 5.816 | 1.117 | 31.762 |
| SC density | 21.349 | 7.243 | 8.54 | 49.093 |
| PHC density | 3.65 | 1.896 | 1.032 | 9.878 |
| CHC density | 0.768 | 0.343 | 0.053 | 1.931 |
| Drinking water access | 90.28 | 6.875 | 66.333 | 99.167 |

**Infrastructure spread is almost similar across different states**



## REGIONAL COMPARISONS

To visualise the inequalities and disparities, these 21 states have been divided into 5 broad geographical regions as follows:

**Table 2: Regional Classification**

| Region | States |
|---|---|
| Central | Madhya Pradesh, Chattisgarh |
| East & North East | Bihar, Jharkhand, Odisha, West Bengal, Assam |
| North | Jammu & Kashmir, Himachal Pradesh, Punjab, Haryana, Uttarakhand, Uttar Pradesh |
| South | Andhra Pradesh, Karnataka, Kerala, Tamil Nadu, Telangana |
| West | Rajasthan, Gujarat, Maharashtra |

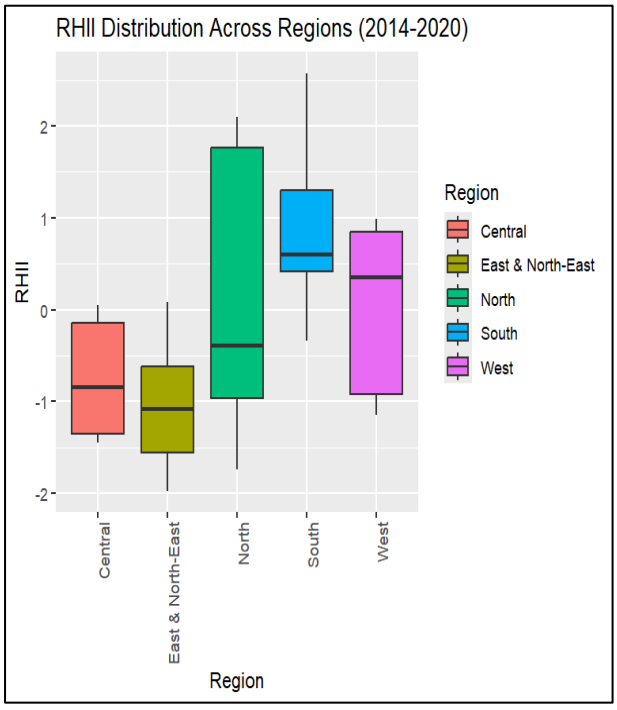



**fig.1**

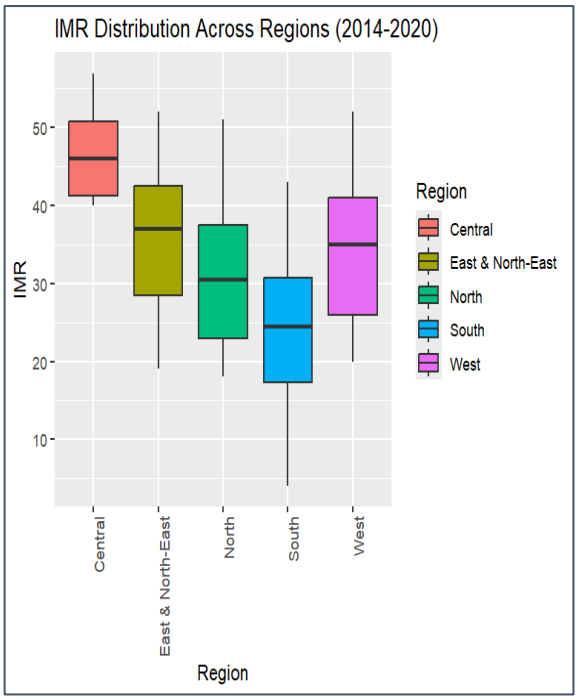



**fig.2**

The first plot(fig.1) illustrates the distribution of the Rural Health Infrastructure Index (RHII) across different regions from 2014 to 2020. The Northern and Southern regions exhibit the highest RHII values, indicating relatively stronger rural health infrastructure. In contrast, Central and East & North-East regions have negative median RHII values, suggesting weaker infrastructure and fewer healthcare resources. The wider interquartile range (IQR) in the North suggests greater variability in health infrastructure across states within the region, while the South shows a more concentrated distribution, indicating relatively uniform healthcare facilities.

The second box plot(fig.2) presents the distribution of Infant Mortality Rate (IMR) across regions. The Central region reports the highest IMR, followed by the East & North-East, both of which also exhibited lower RHII values in the first plot. Conversely, the South has the lowest IMR, aligning with its stronger RHII values. The negative correlation between RHII and IMR is evident, as regions with better rural health infrastructure tend to have lower IMR. The wider range of IMR values in some regions, particularly the West, suggests intra-regional disparities, indicating that certain states within these regions perform significantly better than others. Further scatter plots can be referred in the appendix.

These findings reinforce the hypothesis that strengthening rural health infrastructure is crucial for reducing infant mortality, with particular policy focus needed for Central and East & North-East regions, where both infrastructure and child health outcomes remain weak.

## REGRESSION RESULTS

The results (Table 3) from the pooled OLS, fixed effects (FE), and random effects (RE) models indicate a strong negative relationship between rural health infrastructure (RHII) and infant mortality rate (IMR). Across all specifications, the RHII coefficient is negative and statistically significant, confirming that improvements in rural health infrastructure lead to lower IMR. Specifically, in the FE model (preferred based on the Hausman test), a one-unit increase in RHII is associated with a 2.44 percentage point reduction in IMR, although the effect is slightly larger in the RE (-3.12) and OLS (-4.35) models.

Access to clean drinking water is also significantly associated with lower IMR across all models. The coefficient suggests that a one percentage point increase in access to clean drinking water reduces IMR by approximately 0.63 to 0.90 percentage points, highlighting the importance of basic sanitation and hygiene in reducing infant mortality.

The R-squared values indicate that the FE model explains 50.1% of the variation in IMR, making it a better fit than the OLS model (37.6%). The highly significant constant term suggests that even after accounting for RHII and drinking water access, other unobserved factors contribute to IMR variations across states.

**Table 3: Basic Specification**

| | (OLS) | (FE) | (RE) |
|---|---|---|---|
| VARIABLES | Rural_IMR | Rural_IMR | Rural_IMR |
| RHII | -4.352*** | -2.444** | -3.120*** |
| | (0.608) | (1.173) | (0.951) |
| drinkingwater | -0.632*** | -0.896*** | -0.876*** |
| | (0.107) | (0.0814) | (0.0789) |
| Constant | 89.03*** | 113.0*** | 111.1*** |
| | (9.645) | (7.352) | (7.370) |
| Observations | 146 | 146 | 146 |
| R-squared | 0.376 | 0.501 | |
| Number of States | | 21 | 21 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Another regression result (Table 4) is also obtained by running a model with all explanatory variables used to construct RHII. It suggests that the **RHII-based models (OLS, FE, RE)** consistently indicate a **negative and significant association between rural health infrastructure and IMR**, affirming that better infrastructure correlates with lower infant mortality. However, these models provide an **aggregated effect**, potentially obscuring the nuanced contributions of specific components of health infrastructure.

**Table 4: With all variables**

| | OLS | OLS_RHII | FE | FE_RHII | RE | RE_RHII |
|---|---|---|---|---|---|---|
| VARIABLES | Rural_IMR | Rural_IMR | Rural_IMR | Rural_IMR | Rural_IMR | Rural_IMR |
| Doctors_den | -2.308*** | | 0.731** | | 0.492 | |
| | (0.603) | | (0.358) | | (0.356) | |
| Nurses_den | -0.235* | | 0.0141 | | -0.0556 | |
| | (0.125) | | (0.104) | | (0.0997) | |
| ANM_MW_den | -0.00953 | | -0.0202 | | -0.0638 | |
| | (0.183) | | (0.137) | | (0.130) | |
| PHC_den | 2.431*** | | -1.794* | | -1.354* | |
| | (0.761) | | (0.916) | | (0.790) | |
| CHC_den | -4.113 | | -10.86* | | -6.919 | |
| | (3.973) | | (5.880) | | (4.877) | |
| SC_den | -0.263 | | -0.157 | | -0.169 | |
| | (0.241) | | (0.196) | | (0.194) | |
| drinkingwater | -0.754*** | -0.632*** | -0.997*** | -0.896*** | -0.935*** | -0.876*** |
| | (0.120) | (0.116) | (0.0856) | (0.0814) | (0.0822) | (0.0789) |
| RHII | | -4.352*** | | -2.444** | | -3.120*** |
| | | (0.662) | | (1.173) | | (0.951) |
| Constant | 116.3*** | 89.03*** | 137.4*** | 113.0*** | 130.7*** | 111.1*** |
| | (11.88) | (10.65) | (9.507) | (7.352) | (8.646) | (7.370) |
| Observations | 146 | 146 | 146 | 146 | 146 | 146 |
| R-squared | 0.498 | 0.376 | 0.556 | 0.501 | | |
| Number of States | | | 21 | 21 | 21 | 21 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In the full specification models, the results reveal more intricate relationships:

- Doctor density, which is negatively associated with IMR in OLS, turns positive in FE/RE, suggesting the presence of state-level unobserved heterogeneity influencing healthcare access and outcomes.

- PHC density exhibits a contrasting effect across specifications, with a positive association in OLS but a negative and significant effect in FE/RE models, possibly indicating an endogenous policy response where PHCs are set up in regions with higher IMR.

- Drinking water access remains consistently negative and highly significant across all models, highlighting its fundamental role in reducing infant mortality.

The divergence in results can be attributed to differences in degrees of freedom. The RHII-based models, by reducing the number of estimated parameters, gain efficiency but risk overlooking key infrastructural determinants. Conversely, the full models, while improving granularity, sacrifice degrees of freedom and may introduce multicollinearity among highly correlated infrastructure variables.

## STRONG RELATION BETWEEN IMR & WATER ACCESS

From previous two regression results, it is evident that drinking water access is one of the crucial factors of determining IMR outcomes. We divided the states as per the criteria laid down by the National Rural Drinking Water Programme (NRDWP) which is now merged with Jal Jeevan Mission (JJM) i.e states with less than 55% water access are categorised as low access, between 55% and 89% as medium access and more than 90% as high access. In table 5, the coefficient of water_category is -8.314, showing a strong negative relationship between access to improved drinking water and IMR. This suggests that regions with better drinking water facilities experience significantly lower infant mortality rates, likely due to reduced waterborne diseases and overall improved maternal and child health.

**Table 5: Water access results**

| VARIABLES | Rural_IMR |
|---|---|
| RHII | -3.959*** |
| | (0.656) |
| 3.water_category | -8.314*** |
| | (1.500) |
| Constant | 36.59*** |
| | (1.195) |
| Observations | 146 |
| R-squared | 0.360 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The R-squared value of 0.360 indicates that 36% of the variation in rural IMR is explained by RHII and drinking water access, suggesting a moderate explanatory power. The significance levels ($p<0.01$) further confirm the robustness of these associations. This result is also supported by a meta-analysis conducted by Kremer et al. (2022), which establishes that improved water access is the most cost-effective way to reduce under-5 IMR.

## ROBUSTNESS CHECKS

### COVID SCENARIO

Generated a covid dummy to see whether relationship changed in covid year i.e 2020

$$Rural\_IMR_{it} = \alpha_i + \beta_1 RHII_{it} + \beta_2 COVID_t + \beta_3 (RHII_{it} \times COVID_t) + Drinkingwater_{it} + \epsilon_{it}$$

Where, $COVID_t$ : Dummy variable (1 for 2020, 0 otherwise)

$RHII_{it} \times COVID_t$ : Interaction term

**Table 6: Covid Scenario**

| | OLS | FE | RE |
|---|---|---|---|
| VARIABLES | Rural_IMR | Rural_IMR | Rural_IMR |
| RHII | -4.335*** | -1.789 | -2.569*** |
| | (0.719) | (1.125) | (0.939) |
| $COVID_t$ | -4.007** | -3.579*** | -3.544*** |
| | (1.999) | (0.821) | (0.808) |
| $RHII_{it} \times COVID_t$ | 0.0608 | 0.0688 | 0.0993 |
| | (1.668) | (0.592) | (0.587) |
| drinkingwater | -0.589*** | -0.729*** | -0.725*** |
| | (0.119) | (0.0856) | (0.0822) |
| Constant | 85.66*** | 98.57*** | 98.02*** |
| | (10.87) | (7.664) | (7.630) |
| Observations | 146 | 146 | 146 |
| R-squared | 0.391 | 0.569 | |
| Number of States | | 21 | 21 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6 results show that RHII reduces IMR, but its effect weakens in the fixed effects model, suggesting state-specific factors influence IMR. COVID-19 significantly reduced IMR (~4 points), likely due to improved healthcare attention during the pandemic. However, the interaction term (COVID-19 × RHII) is insignificant, indicating that RHII's impact on IMR did not significantly change during COVID-19.

## EAG STATES ANALYSIS

EAG states refer to the Empowered Action Group (EAG) states in India. These are eight states that were identified by the Government of India in 2001 as being socio-economically weaker, with poor health indicators and high fertility rates. The EAG states are: **Bihar, Jharkhand, Uttar Pradesh, Uttarakhand, Madhya Pradesh, Chhattisgarh, Rajasthan, and Odisha.**

So, a dummy is generated for whether a state is a part of Empowered action group to study the relationship between RHII & rural IMR.

**Table 7: EAG Relationship**

| | OLS | FE | RE |
|---|---|---|---|
| VARIABLES | Rural_IMR | Rural_IMR | Rural_IMR |
| RHII | -3.462*** | -2.721** | -2.970** |
| | (0.788) | (1.103) | (1.199) |
| EAG | 10.71*** | | 10.33*** |
| | (1.303) | | (3.621) |
| EAG_inf | 3.166** | 2.598 | 2.656 |
| | (1.372) | (4.228) | (2.765) |
| drinkingwater | -0.538*** | -0.886*** | -0.853*** |
| | (0.113) | (0.130) | (0.113) |
| Constant | 77.48*** | 112.9*** | 106.0*** |
| | (10.48) | (11.41) | (10.91) |
| Observations | 146 | 146 | 146 |
| R-squared | 0.516 | 0.503 | |
| Number of States | | 21 | 21 |

Table 7 shows that better rural health infrastructure significantly reduces infant mortality (IMR) across Indian states. EAG states have higher IMR than non-EAG states, even after controlling for infrastructure. Since the interaction term (EAG_inf) is statistically insignificant, this suggests that the relationship between IMR and RHII remains similar for both EAG and non-EAG states. This indicates that EAG states are catching up with non-EAG states owing to their respective government's efforts and commitment.

## PLACEBO TESTS

To assess the robustness of the relationship between rural health infrastructure and infant mortality (IMR), a placebo test was conducted by randomly assigning values to the Rural Health Infrastructure Index (RHII). Table 8 results reveal that the **RHII_random** variable is statistically insignificant across all specifications (OLS, FE, and RE), with coefficients close to zero and large standard errors. This suggests that when infrastructure values are randomly assigned, they do not exhibit a meaningful relationship with IMR.

**Table 8: Fake RHII test**

| | OLS | FE | RE |
|---|---|---|---|
| VARIABLES | Rural_IMR | Rural_IMR | Rural_IMR |
| RHII_random | 0.760 | 0.0512 | 0.0675 |
| | (0.752) | (0.176) | (0.179) |
| drinkingwater | -0.632*** | -0.887*** | -0.870*** |
| | (0.115) | (0.143) | (0.134) |
| Constant | 89.51*** | 112.4*** | 110.9*** |
| | (10.31) | (12.87) | (12.63) |
| Observations | 146 | 146 | 146 |
| R-squared | 0.159 | 0.484 | |
| Number of States | | 21 | 21 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

## DISCUSSION

The findings of this study underscore the critical role of rural health infrastructure in reducing infant mortality rates (IMR) across Indian states between 2014 and 2020. The construction of the Rural Health Infrastructure Index (RHII) through Principal Component Analysis (PCA) provides a comprehensive measure of health infrastructure availability, revealing substantial disparities among states. The regression analyses further demonstrate that states with better-equipped rural health infrastructure exhibit significantly lower IMR, reaffirming the necessity of continuous investment in primary healthcare facilities.

The importance of rural health infrastructure has gained renewed attention, particularly in the aftermath of the COVID-19 pandemic. The pandemic exposed deep-seated vulnerabilities in India's

health system, especially in rural areas where shortages of health personnel, inadequate hospital infrastructure, and supply chain disruptions were pronounced. While our study primarily covers the pre-pandemic period, the findings hold significant relevance in the present context, as recent government initiatives under Ayushman Bharat and PM-ASBY aim to strengthen primary healthcare and augment rural health capacities. These interventions align with our study's findings that improved rural health infrastructure has the potential to significantly impact infant mortality.

Another key finding is the persistent shortage of health manpower, which continues to be a critical challenge in rural India. The data suggests that despite the presence of physical infrastructure, the lack of trained healthcare professionals—doctors, nurses, and community health workers—limits the effectiveness of healthcare delivery. This highlights the need for policies that not only expand infrastructure but also ensure adequate human resources through incentives for rural postings, better training facilities, and technological integration such as telemedicine.

The analysis also suggests that while infrastructure improvements correlate with lower IMR, the effect is not uniform across all states. Factors such as governance, efficiency in resource allocation, and socio-economic determinants like maternal education and nutrition play a crucial role in amplifying or limiting the impact of health infrastructure. Therefore, a more nuanced, region-specific approach is necessary to enhance rural healthcare outcomes effectively.

Lastly, our study hints at the potential long-term implications of the pandemic on rural health outcomes. The disruptions caused during the pandemic may have exacerbated existing inequities, making it imperative to track post-pandemic trends in IMR and rural health service delivery. Further research should explore how pandemic-induced shifts—such as increased digital health adoption and expanded insurance coverage—affect long-term infant mortality trends and rural health system resilience.

## CONCLUSION

In conclusion, while India has made commendable progress in improving rural health infrastructure, there remains a significant gap in equitable access and service delivery. The findings reinforce the need for a holistic approach integrating infrastructure development, workforce strengthening, and targeted policy interventions to achieve sustainable reductions in infant mortality. With the vision of 'Viksit Bharat 2047,' investments in rural health infrastructure must be prioritized to ensure that India's development trajectory is inclusive and health outcomes are equitably distributed across regions. A sustained commitment to bridging infrastructural gaps, enhancing healthcare workforce capacity, and implementing region-specific policies will be crucial in achieving long-term improvements in rural health indicators.

**ACKNOWLEDGEMENT**

I am grateful to Prof. Oindrila De for assisting me in my seminar paper. This study would not have been not possible without her valuable insights and suggestions. I am also thankful to Prof. Vikram Dayal and Prof. Saudamini Das for providing their valuable inputs during presentations. I would also like to thank Dr. Naresh for his constant support.

## REFERENCES


Ministry of Health and Family Welfare, Government of India. (2014–2020). *Rural health statistics*. Registrar General of India. (2014–2020).

Registrar General of India. (2014-2020). *Sample Registration System statistical reports(2014-20)*. Office of the Registrar General & Census Commissioner, India. https://censusindia.gov.in

International Institute for Population Sciences (IIPS) & ICF. (2017). *National Family Health Survey (NFHS-4), 2015-16: India*. Mumbai: IIPS.

International Institute for Population Sciences (IIPS) & ICF. (2021). *National Family Health Survey (NFHS-5), 2019-21: India*. Mumbai: IIPS.

Lakshmi, T. S., & Sahoo, D. (2013). Health infrastructure and health indicators: The case of Andhra Pradesh, India. *IOSR Journal of Humanities and Social Science*, *6*(6), 22–29.

Pradhan, J., Kumar, A., & Sanyal, K. (2011). Health infrastructure in India: The input and output association with economic growth. *Journal of Health Management*, *13*(1), 59–75. https://doi.org/10.1177/097206341001300104

Lyngdoh, M. (2015). Inter-state variations in rural healthcare infrastructure in North-East India. *The NEHU Journal*, *13*(2), 31–48.

Suriyakala, V., Deepika, M. G., Amalendu, J., & Deepa, G. (2016). Factors affecting infant mortality rate in India: An analysis of Indian states. In J. Corchado Rodriguez, S. Mitra, S. Thampi, & E. S. El-Alfy (Eds.), *Intelligent systems technologies and applications 2016* (Vol. 530, pp. 707–719). Springer. https://doi.org/10.1007/978-3-319-47952-1_69

Kremer, M., et al. (2015). Potential reduction in child mortality through expanding access to safe drinking water in India. *University of Chicago*.

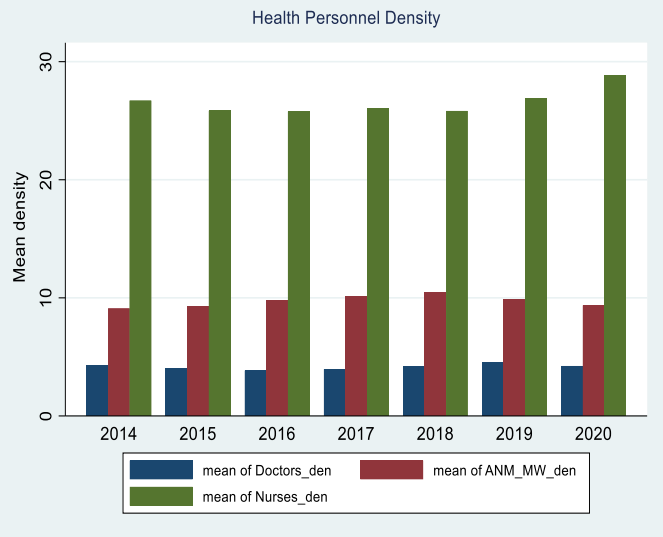
Das, U., & Guha, I. (2022). The puzzle of public health expenditure and healthcare infrastructure in India: An empirical investigation.

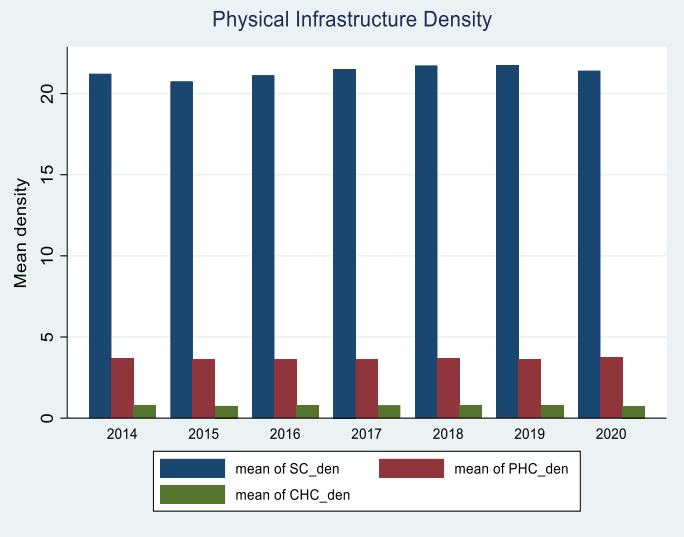
Arun, C., et al. (2018). Influence of health service infrastructure on the infant mortality rate: An econometric analysis of Indian states.

United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*.

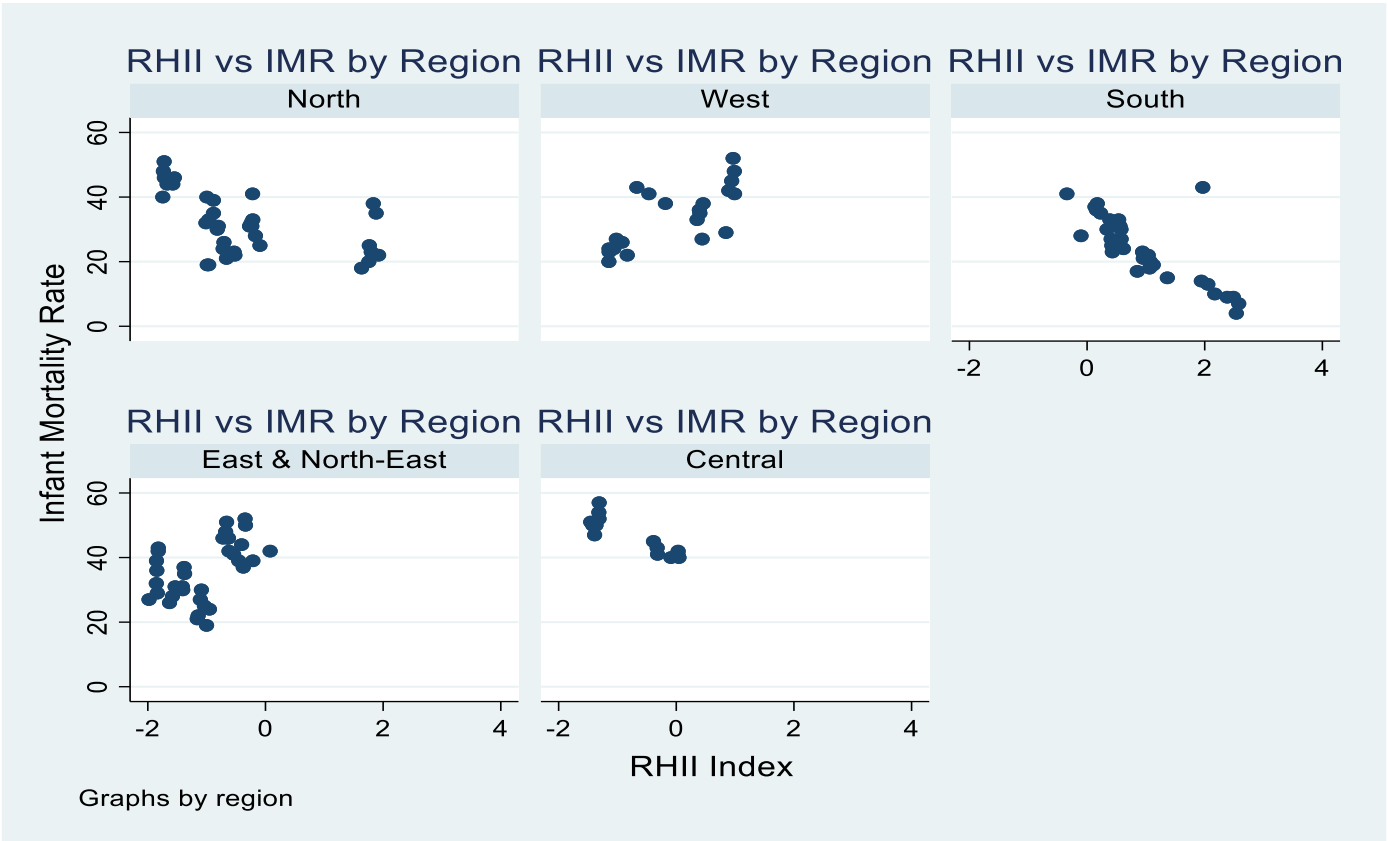
# APPENDIX

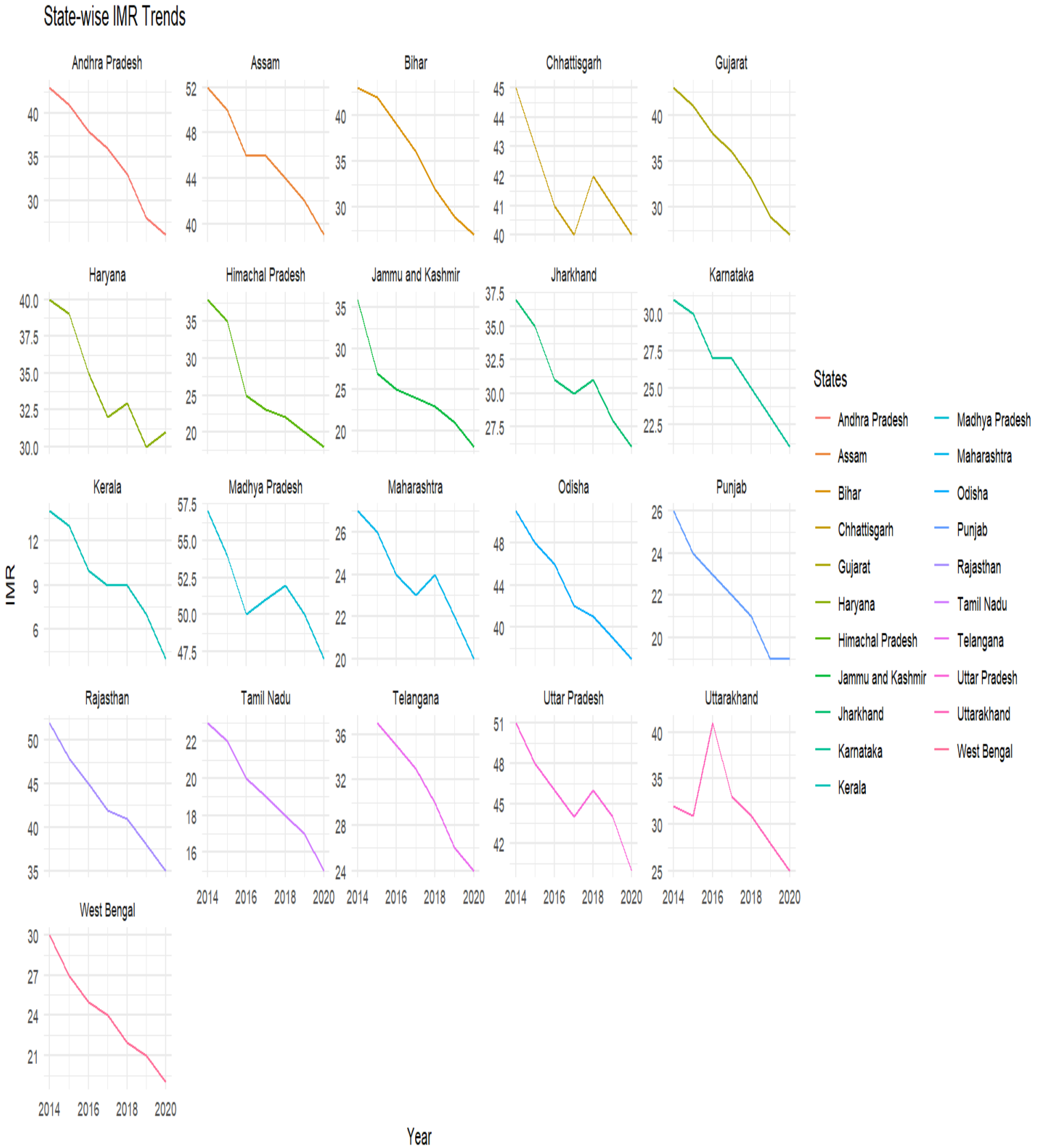
## 1. Manpower v/s Physical Infra




## 2. Southern States shows the clear depiction




State-wise IMR Trends
Andhra Pradesh
Assam
Bihar
Chhattisgarh
Gujarat
Haryana
Himachal Pradesh
Jammu and Kashmir
Jharkhand
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Odisha
Punjab
Rajasthan
Tamil Nadu
Telangana
Uttar Pradesh
Uttarakhand
West Bengal
IMR
Year
States
Andhra Pradesh
Assam
Bihar
Chhattisgarh
Gujarat
Haryana
Himachal Pradesh
Jammu and Kashmir
Jharkhand
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Odisha
Punjab
Rajasthan
Tamil Nadu
Telangana
Uttar Pradesh
Uttarakhand
West Bengal