PREFACE TO THE EDITION

It is with great pleasure that we present the forthcoming issue of the **Journal of Economic Insights and Research (JEIR)**, bringing together a diverse collection of studies that examine contemporary economic issues through empirical, methodological, and policy-oriented perspectives. The articles in this issue reflect the changing nature of economic systems and highlight the importance of rigorous evidence in understanding markets, institutions, labour, technology, and public policy.

The issue opens with a methodological contribution on gold price risk assessment, which explores the application of reliability and survival analysis to commodity price behaviour. By examining adverse gold-price events through reliability concepts, the study demonstrates both the potential and the limitations of applying survival-analysis frameworks to financial markets. Its emphasis on methodological interpretation and the boundaries of statistical inference provides a valuable contribution to the study of commodity risk.

The issue then turns to an important labour-market question in India through an examination of fare-free public transport and female labour supply. Using staggered policy reforms across Indian states and robust econometric approaches, the study investigates whether reducing transportation costs can encourage women's participation in the labour market. Its findings draw attention to the role of mobility in shaping employment opportunities, particularly for urban women and those whose work requires travel beyond their immediate locality.

The transformative influence of Artificial Intelligence on the global economy is examined in another contribution. Addressing economic growth, productivity, employment, international trade, innovation, and inequality, the study presents AI as both an opportunity and a source of significant economic and policy challenges. The discussion of technological preparedness, workforce reskilling, ethical governance, cybersecurity, and unequal access to digital infrastructure underscores the need for inclusive approaches to technological transformation.

The issue also presents an analysis of India–Russia bilateral trade, examining recent trade patterns alongside the macroeconomic factors influencing bilateral trade flows. Particular attention is given to energy imports, the persistent trade deficit, oil prices, exchange-rate movements, and the broader prospects for strengthening economic relations between the two countries. The study provides useful insights into the interaction between international commodity markets and bilateral trade relationships.

The relationship between public health expenditure, governance quality, and infant mortality forms another important theme of this issue. Through a dynamic panel analysis of Indian states, the study demonstrates that the effectiveness of public health expenditure is closely connected with administrative capacity. The findings therefore move beyond the question of how much governments spend to consider the equally important question of how effectively public resources are translated into health outcomes.

The issue concludes with an examination of the size and composition of India's gig workforce, with particular attention to Assam and the Guwahati Metropolitan Area. By considering sectoral, occupational, skill-level, and formal–informal dimensions of gig employment, the study contributes to understanding an increasingly important component of India's changing labour market and provides a framework for estimating the scale of gig work at the sub-national level.

Taken together, the contributions in this issue demonstrate the breadth of contemporary economic inquiry. From financial-market risk and monetary transmission to women's employment, technological transformation, international trade, public health, and emerging forms of work, the studies illustrate how economic analysis increasingly intersects with social, technological, institutional, and policy dimensions. They also reinforce the importance of methodological rigour, reliable evidence, and careful interpretation when addressing complex economic realities.

We hope that this issue of the Journal of Economic Insights and Research will stimulate further academic discussion, encourage new research, and provide useful insights for researchers, economists, policymakers, practitioners, and students. We sincerely thank all the authors, reviewers, and editorial contributors whose efforts have made this issue possible.

Dr. Sinitha Xavier
Chief Editor

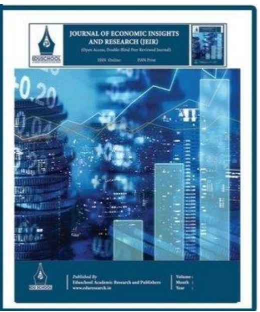
CONTENTS

EDUSCHOOL
ACADEMIC RESEARCH & PUBLISHERS

JOURNAL OF ECONOMIC INSIGHTS AND RESEARCH (JEIR)

(Open Access, Double-Blind Peer Reviewed Journal)

ISSN Online: 3107-9482 ISSN Print: 3139-1982



A Reliability-Based Survival Analysis Framework for Gold Price Risk Assessment: A Methodological Demonstration

Geethu Gopinath[1], Ansa Alphonsa Antony[2]

[1] Research Scholar, Department of Statistics, Maharaja's College (Autonomous), Ernakulam, India
[2] Associate Professor, Department of Statistics, St. Xavier's College for Women (Autonomous), Aluva, India

Article information

Received: 14th August 2026
Received in revised form: 18th August 2026
Accepted: 21st August 2026
Available online: 25th August 2026

Volume: 2
Issue: 3
DOI: https://doi.org/10.63090/JEIR/3107.9482.0023

Abstract

Gold is widely regarded as a safe-haven asset and occupies a prominent position in investment portfolios during periods of economic uncertainty. Previous research has concentrated overwhelmingly on forecasting gold prices using econometric and machine-learning techniques, whereas the assessment of gold price risk from the perspective of reliability theory has received little attention. This study demonstrates how reliability and survival analysis methods may be applied to the timing of adverse gold price events. The data are daily settlement prices of COMEX gold futures in US dollars per troy ounce for 1 January 2014 to 6 January 2025 (2,565 trading days). A failure event is defined as a price falling below the twentieth percentile of the observed full-sample distribution, and the interval between consecutive failures is treated as the time-to-failure variable. The Kaplan–Meier estimator is used to obtain a non-parametric estimate of the reliability function, and the cumulative hazard function, obtained as the negative logarithm of the reliability estimate, is used to describe the accumulation of failure risk. Weibull and Exponential parametric lifetime models are fitted by maximum likelihood, compared by likelihood ratio test and by the Akaike and Bayesian information criteria, and used to estimate the Mean Time to Failure (MTTF). Under the stated failure definition, 513 of 2,565 daily observations (20.00%) were classified as failures and the Kaplan–Meier reliability estimate was 0.8000 (95% CI 0.7847–0.8156), with a terminal cumulative hazard of 0.223. The Weibull model fitted substantially better than the Exponential (likelihood ratio statistic 350.55 on 1 degree of freedom, $p < 0.001$; AIC 9904.98 versus 10253.53), with an estimated shape parameter of 2.38 and a scale parameter of 3289. The corresponding MTTF estimates were 2,918 and 8,065 trading days respectively, both of which exceed the 2,565-day observation window and therefore rest on extrapolation beyond the range of the data. These results are reported alongside an explicit assessment of what they can and cannot support. Because the failure threshold is a fixed percentile of the same distribution from which failures are counted, the failure proportion, the terminal reliability estimate and the terminal cumulative hazard are determined by the threshold rather than estimated from the data; they are reported here for completeness rather than as substantive findings about gold. Further, the parametric models were fitted on the trading-day index rather than on inter-failure times, and the fitted Weibull implies a cumulative failure probability of 42.5% by day 2,565 against a Kaplan–Meier estimate of 20.0%, so neither parametric model reproduces the non-parametric curve. The contribution of this paper is therefore methodological: it sets out how reliability concepts map onto commodity price risk, identifies precisely where the mapping breaks down when applied to a trending, autocorrelated price level, and specifies the redesign required for the framework to yield interpretable risk estimates.

Keywords: - Reliability Theory, Gold Price Risk, Survival Analysis, Kaplan–Meier Estimator, Weibull Distribution, Recurrent Events, Financial Risk Assessment.

I. INTRODUCTION

Gold has long been recognised as one of the most important financial assets, valued for its capacity to preserve wealth during periods of economic uncertainty, inflation and market volatility. Its behaviour during financial crises has been studied extensively: Baur and Lucey (2010) distinguish gold's role as a hedge from its role as a safe haven, and Baur and McDermott

(2010) find safe-haven behaviour for major developed markets during acute crisis episodes, though the effect is neither universal across markets nor stable over time. Capie, Mills and Wood (2005) document gold's relationship with the US dollar, and Ciner, Gurdgiev and Lucey (2013) extend the analysis across stocks, bonds, oil and exchange rates. O'Connor, Lucey, Batten and Baur (2015) survey the field. Investors, central banks and policymakers accordingly monitor gold price movements closely, while recognising that gold prices themselves fluctuate in response to global economic conditions, geopolitical events, inflation, exchange rates, interest rates and investor sentiment.

Most existing studies of gold prices are concerned with forecasting. Autoregressive integrated moving average models (Box et al., 2015), generalised autoregressive conditional heteroskedasticity models (Engle, 1982; Bollerslev, 1986; Tully & Lucey, 2007) and vector autoregressions (Sims, 1980) have all been applied to gold, as more recently have neural network and hybrid machine-learning approaches (Kristjanpoller & Minutolo, 2015; Livieris, Pintelas and Pintelas, 2020). Valuable as these are, they estimate the conditional level or conditional variance of the price; they do not directly quantify the probability that a price remains above a specified risk threshold for a given duration. In reliability theory that quantity is the reliability function, and extending it to financial markets offers an alternative vocabulary for asset stability in which a substantial adverse price movement is treated as a failure event.

Reliability theory has been developed principally in engineering, manufacturing and the biomedical sciences (Meeker & Escobar, 1998; Lawless, 2003; Rinne, 2009). Its migration into finance has been limited but not absent: Singpurwalla (2006) sets out a Bayesian reliability perspective on econometric and financial problems, and Gao and He (2021) survey survival methods in finance. Within this framework a financial asset may be viewed as a system whose performance deteriorates under unfavourable market conditions, and the duration between successive adverse episodes may be treated as a time-to-failure.

Survival analysis supplies the estimation machinery. The Kaplan–Meier estimator (Kaplan & Meier, 1958) provides a non-parametric estimate of the survival function without distributional assumptions; the Nelson–Aalen estimator (Nelson, 1972; Aalen, 1978) provides the corresponding cumulative hazard; and parametric lifetime models such as the Weibull and Exponential permit estimation of summary quantities including the Mean Time to Failure. In finance these tools have been used chiefly for corporate distress and default (Lane et al., 1986; Shumway, 2001) and, in a different form, for the timing of market microstructure events (Engle & Russell, 1998).

This study applies that apparatus to daily gold price data. A failure event is defined as a closing price below the twentieth percentile of the observed price distribution, and the intervals between successive failures are treated as time-to-failure observations. The Kaplan–Meier estimator is used to estimate the reliability function, the cumulative hazard function to describe risk accumulation, and Weibull and Exponential models to characterise the failure-time distribution and estimate the MTTF.

The contribution of the paper is stated carefully, because the empirical results reported in Section 4 do not support the stronger claims that a reliability framing might invite. The contribution is threefold. First, the paper sets out explicitly how the constructs of reliability theory (reliability function, hazard, cumulative hazard and MTTF) map onto the problem of commodity downside risk, and what each requires of the data. Second, it implements that mapping end to end on a real price series using both non-parametric and parametric methods. Third, and most importantly for subsequent work, it identifies the specific points at which a naive implementation fails: a fixed full-sample percentile threshold applied to a trending price level renders the headline quantities tautological; contiguous below-threshold days are not independent failure episodes; and the renewal assumption implicit in fitting independent lifetime distributions to successive inter-arrival times is inappropriate for a single recurrent-event system. Section 6 sets out the redesign these problems imply. This diagnostic contribution is arguably the more durable of the three.

The remainder of the paper is organised as follows. Section 2 reviews the literature on gold price modelling, reliability theory, survival analysis in finance, and the extreme value and duration methods against which any threshold-exceedance framework must be benchmarked. Section 3 presents the data and methods. Section 4 reports the empirical results. Section 5 discusses their interpretation. Section 6 sets out the limitations of the present design and the redesign required to address them. Section 7 concludes.

II. LITERATURE REVIEW

2.1 Gold price behaviour and financial market analysis

Research on gold price dynamics divides broadly into work on gold's portfolio role and work on forecasting. On the former, Baur and Lucey (2010) formalise the hedge/safe-haven distinction and find gold to be a hedge against stocks and a safe haven in extreme conditions; Baur and McDermott (2010) extend this internationally; Hillier, Draper and Faff (2006) assess precious metals from an investment perspective; and Ciner, Gurdgiev and Lucey (2013) examine safe-haven behaviour across asset classes. Batten, Ciner and Lucey (2010) study the macroeconomic determinants of precious-metals volatility.
On forecasting, ARIMA (Box et al., 2015) and GARCH-family models (Engle, 1982; Bollerslev, 1986) remain standard; Tully and Lucey (2007) apply an asymmetric power GARCH specification to gold and find the US dollar to be the dominant macroeconomic influence. Mills (2004) documents the statistical properties of daily gold prices, including scaling behaviour and heavy tails. Machine-learning approaches have since been widely adopted, including ANN–GARCH hybrids (Kristjanpoller & Minutolo, 2015) and CNN–LSTM architectures (Livieris et al., 2020), typically reporting improved point-forecast accuracy for nonlinear dynamics.

Two features of gold price data matter directly for the present study. First, the price level is non-stationary and strongly trending over the sample considered here, so any analysis conducted on the level rather than on returns or drawdowns inherits that trend. Second, returns exhibit volatility clustering and persistence (Engle, 1982; Bollerslev, 1986; Tully & Lucey, 2007), which violates the independence assumptions on which conventional survival estimators rest. Both points are taken up in Sections 3, 5 and 6.

2.2 Reliability theory and financial risk

Reliability theory provides a probabilistic framework for the lifetime and failure behaviour of systems (Meeker & Escobar, 1998; Lawless, 2003; Rinne, 2009). Its central quantities, namely the reliability function, the hazard function, the cumulative hazard function and the Mean Time to Failure, describe the probability of survival to a given time, the instantaneous risk of failure at that time, the accumulated risk, and the expected lifetime.

A distinction internal to the field is fundamental to what follows. Reliability theory treats non-repairable components, which fail once and are analysed with lifetime distributions and survival estimators, quite differently from repairable systems, in which a single system fails repeatedly over time and is analysed with point-process methods (Ascher & Feingold, 1984; Rigdon & Basu, 2000). A financial asset observed continuously and experiencing recurrent adverse episodes belongs to the second class. Applying lifetime-distribution methods to successive inter-arrival times from such a system implicitly assumes a renewal process, that is, restoration to an "as good as new" state after each failure. That assumption is testable (Cox and Lewis, 1966; Lewis and Robinson, 1974) and should not be adopted silently. The present study adopts the lifetime-distribution route, following the existing applied literature, and Section 6 assesses the cost of that choice.

Applications of reliability concepts within finance remain sparse. Singpurwalla (2006) develops the connection between reliability, survival and econometric modelling from a Bayesian standpoint, and Gao and He (2021) review survival methods for financial applications.

2.3 Survival analysis in finance

Survival analysis models the occurrence and timing of events and has been applied widely in medicine, epidemiology and engineering. Among non-parametric methods, the Kaplan–Meier estimator (Kaplan & Meier, 1958) is the standard tool for estimating survival probabilities without distributional assumptions, with variance given by Greenwood's formula (Greenwood, 1926); the Nelson–Aalen estimator (Nelson, 1972; Aalen, 1978) provides the cumulative hazard. Parametric lifetime models, particularly the Weibull and Exponential, allow estimation of the MTTF and formal comparison of competing failure-time distributions (Lawless, 2003; Kalbfleisch & Prentice, 2002).

A terminological caution is warranted before proceeding. The word "survival" carries a second and quite different meaning in the finance literature, where Brown, Goetzmann and Ross (1995) use it to denote survivorship bias, the distortion introduced when only surviving assets or funds are observed. The present study uses the term exclusively in its time-to-event sense.

In financial applications, hazard models have been used to predict corporate failure. Lane, Looney and Wansley (1986) apply the Cox proportional hazards model (Cox, 1972) to bank failure, and Shumway (2001) shows that a discrete-time hazard model outperforms static bankruptcy classifiers by using all available firm-year observations and accounting for time to failure. Engle and Russell (1998) introduce autoregressive conditional duration models for the intervals between irregularly spaced market events, explicitly modelling dependence between successive durations, a feature absent from renewal-based lifetime models and directly relevant to the present setting.

For recurrent events, the standard formulations are the Andersen–Gill counting-process model (Andersen & Gill, 1982), the conditional models of Prentice, Williams and Peterson (1981), and the marginal approach of Wei, Lin and Weissfeld (1989), each combined with robust variance estimation.

2.4 Threshold exceedance, extreme value theory and downside risk measures

Because the present study defines failure as exceedance of a threshold in the lower tail, it operates in territory already occupied by a mature statistical literature, and the relationship must be stated rather than left implicit.

Extreme value theory provides the standard apparatus for threshold exceedance. The peaks-over-threshold approach rests on the result that exceedances over a sufficiently high threshold converge to the Generalised Pareto distribution (Balkema & de Haan, 1974; Pickands, 1975), with statistical treatment developed by Davison and Smith (1990) and expositions in Embrechts, Klüppelberg and Mikosch (1997) and Coles (2001). McNeil and Frey (2000) combine GARCH filtering with EVT to estimate conditional tail risk in financial series, addressing precisely the dependence problem noted in Section 2.1. A recurring practical issue in this literature, namely that exceedances arrive in clusters rather than singly, so that the raw count of exceedances overstates the number of independent extreme episodes, is handled by declustering (Ferro & Segers, 2003). The same issue arises acutely in the present design and is discussed in Sections 5 and 6.

In parallel, practitioner risk measurement rests on Value-at-Risk (Jorion, 2007) and Expected Shortfall, the latter satisfying the coherence axioms of Artzner, Delbaen, Eber and Heath (1999) that VaR does not. Any framework claiming to complement conventional financial risk assessment should be positioned against these measures.

2.5 Research gap and the position of this study

The literature on gold prices is heavily weighted towards forecasting and volatility estimation; comparatively little work treats adverse gold price movements as recurrent events and asks how long favourable conditions persist. Reliability theory offers a vocabulary for that question, and the integration of non-parametric and parametric reliability models within a single applied framework for a commodity price series does not appear to have been presented in detail.

Caution is warranted, however, regarding the size of this gap. Extreme value theory already provides a rigorous and widely used treatment of threshold exceedance, including the clustering problem, and autoregressive conditional duration models already provide a treatment of dependent event intervals. A reliability framing is not automatically an advance on either. The defensible claim is narrower: that reliability theory supplies an interpretable set of summary quantities (reliability at a horizon, cumulative hazard, and expected time between adverse episodes) that are intuitive for practitioners, and that establishing whether these add anything beyond EVT and duration models requires direct comparison. The present study does not perform that comparison, and Section 6 identifies it as the principal requirement for future work.

III. MATERIALS AND METHODS

3.1 Data description

This study uses daily gold price data covering 1 January 2014 to 6 January 2025, comprising 2,565 trading-day records, or approximately eleven years of market activity. The series consists of daily settlement prices of gold futures traded on COMEX, the metals division of the New York Mercantile Exchange, quoted in US dollars per troy ounce. The contract size is 100 troy ounces. The data were obtained from a publicly available compilation of this series hosted on Kaggle, which reproduces the standard historical-data export format used by commercial market data providers.

The identity of the series was verified against independent records of the international gold market before analysis, because an earlier version of this work described the data as Multi Commodity Exchange of India (MCX) quotations in Indian Rupees per 10 grams. That description was incorrect: MCX gold traded in the approximate range of ₹25,000 to ₹79,000 per 10 grams over this period, which is two orders of magnitude above the values in the present dataset. The observed minimum of 1049.60 corresponds to the December 2015 trough in dollar gold, which followed the Federal Reserve's first post-crisis rate increase, and the observed maximum of 2800.80 corresponds to the record settlement reached on 30 October 2024, when COMEX gold futures traded to an all-time high of 2801.80 per ounce. The final observation, 6 January 2025, coincides with the 2025 settlement low of that contract. The series is therefore the international dollar-denominated benchmark rather than an Indian domestic price, and all interpretation in this paper is framed accordingly.

Two features of a futures series should be noted. First, a continuous daily series spanning eleven years is assembled from successive contract months, so the level embeds roll effects at contract transitions and is not a pure spot price. Second, futures settlements incorporate the cost of carry and therefore stand at a small premium to spot. Neither materially affects a threshold-based classification of the lower tail, but both are recorded here for completeness and are revisited in Section 6.1.
The dataset comprises seven variables: Date, Price, Open, High, Low, Vol (daily trading volume) and PChange1 (percentage change in the closing price). The Price variable, the daily settlement price, is taken as the primary response variable, as it represents the final market price of each trading day and is conventional in analyses of asset performance.

Observations are recorded at daily frequency excluding weekends and exchange holidays, so the number of trading days varies by year; across the full sample the mean is approximately 233 trading days per year. No further disaggregation of the annual trading-day counts is reported, as it bears on none of the analyses that follow.

Table 1. summarises the variables. Variable names are given exactly as they appear in the analysis code and are used consistently throughout the paper.

Table 1. Description of variables used in the study

Variable	Description	Unit
Date	Trading date of the observation	Date
Price	Daily settlement price (final trading price)	USD per troy ounce
Open	Opening price at the start of the trading day	USD per troy ounce
High	Highest price recorded during the trading day	USD per troy ounce
Low	Lowest price recorded during the trading day	USD per troy ounce
Vol	Total trading volume during the trading day	Contracts traded
PChange1	Percentage change in settlement price from the previous trading day	Percentage (%)

For the reliability analysis, Price was selected as the primary response variable, and a failure event was defined on the lower tail of its distribution as described in Section 3.3.

3.2 Data pre-processing

The dataset was imported into R for preparation and analysis. Variable names, data types and completeness were checked prior to modelling. The Price series was screened for missing values using sum, which returned zero across all 2,565 records; no imputation or interpolation was therefore required, and observations were analysed in their original chronological order. The Date variable was retained to preserve chronology, and Open, High, Low, Vol and PChange1 were retained as supplementary market indicators but were not used in the models reported here.

Because reliability analysis concerns the occurrence of adverse events rather than the continuous price path, a binary failure indicator was constructed from the distribution of closing prices, as set out in the following section.

3.3 Definition of the failure event and construction of the time-to-failure data

Reliability analysis requires an explicit definition of failure. Here a failure is an instance in which the daily closing price of gold falls below a fixed threshold, taken as the twentieth percentile of the observed closing-price distribution, so that the lowest 20% of price observations are classified as adverse market conditions.

Let $P = \{P_1, P_2, \ldots, P_n\}$ denote the sequence of daily closing prices, where P_i is the closing price on trading day i, i = 1, 2, …, n, and n is the total number of daily observations. The failure threshold is

$$\tau = Q_{0.20}(P),$$

where $Q_{0.20}(P)$ is the empirical twentieth percentile of the closing-price distribution. Each observation is classified as

$$F_i = \begin{cases} 0, & P_i \geq \tau \\ 1, & P_i < \tau \end{cases}$$

where $F_i = 1$ denotes a failure and $F_i = 0$ a non-failure on trading day i. Two properties of this definition should be stated at the outset rather than left for the reader to infer.

The threshold is estimated from the full sample: Because τ is computed from the entire 2014–2025 record and then applied retrospectively to the whole period, including 2014, the classification uses information unavailable to an analyst operating in real time. The design is therefore descriptive of the historical sample and is not implementable as a forward-looking risk tool. A rolling or expanding-window threshold would be required for that purpose.

The failure proportion is fixed by construction: Defining failure as the lowest 20% of a distribution guarantee that exactly 20% of observations are failures. Quantities that are simple functions of that proportion, namely the failure count, the terminal reliability estimate and the terminal cumulative hazard, are consequently determined by the threshold choice and convey no information about gold. They are reported in Section 4 for completeness and for internal consistency, and are labelled accordingly. The substantive quantities in this study are those that depend on the timing of failures, namely the shape of the inter-failure distribution and the fitted model parameters.

Construction of the time-to-failure data

Reliability analysis is concerned with the time elapsed between successive failures. Let $t_1, t_2, \ldots, t_m$ denote the trading days on which failures occurred, where t_j is the occurrence time of the j-th failure and m is the total number of failures. The inter-failure time is

$$T_j = t_j - t_{j-1}, \qquad j=2,3,\ldots,m,$$

so that T_j is the number of trading days between consecutive failures. The survival dataset is $D = \{(T_j, E_j)\}$, where E_j is the event indicator. Every failure in this construction is observed within the study period, so $E_j = 1$ throughout and no right-censored observations arise; the resulting dataset contains at most $m - 1 = 512$ inter-failure times.

Two distinct survival datasets were constructed, and the distinction matters. The construction above yields inter-failure times, of which there are at most $m - 1 = 512$, all uncensored. The parametric models reported in Sections 4.5 to 4.8 were not fitted to that dataset. They were fitted to a second, day-level formulation in which each of the 2,565 trading days contributes one record, the survival time is the trading-day index measured from the start of the observation window, and the event indicator is F_i. That formulation yields exactly the 2,565 observations, 513 events and 2,052 right-censored records reported in Tables 6 and 8, and it is confirmed by the reported information criteria, which are consistent only with $n = 2,565$.

The two datasets answer different questions and are not interchangeable. The inter-failure formulation describes how long favourable conditions persist between successive adverse episodes, which is the quantity the reliability framing of Section 3.4 is intended to capture. The day-level formulation instead describes where within the observation window below-threshold prices are located; its fitted scale parameter and Mean Time to Failure are therefore properties of the position of low prices in the sample, not of the duration between them.

As a reliability model of gold price failures, the day-level construction is not appropriate. It treats each trading day as an independent unit whose lifetime is its own position in the sequence, so that "surviving" to day t means only that an observation is the t-th in the series. The resulting parameter estimates are reported in Section 4 as they were computed, and Section 5.2 explains why the shape estimate cannot bear the interpretation placed on it in the original draft. Refitting the parametric models to the inter-failure dataset specified above, so that the parametric and non-parametric analyses share a unit of analysis, is identified in Section 6.2 as a required step.

A further property of the construction should be recorded. Because gold prices are highly autocorrelated, days below the threshold occur in contiguous runs rather than as isolated events. Within a run, consecutive inter-failure times equal one trading day, and these are not independent episodes. The 513 failures therefore correspond to a substantially smaller number of independent adverse episodes, and the effective sample size for inference is the number of episodes rather than the number of days. The decluttering procedures standard in extreme value analysis (Ferro & Segers, 2003) address exactly this problem. No decluttering was applied in the present analysis, and Section 6 treats this as a primary limitation.

3.4 Reliability function and Kaplan–Meier estimation

Let T denote the random variable representing the inter-failure time between consecutive adverse gold price events. The reliability (survival) function is

$$R(t) = P(T > t),$$

the probability that no failure occurs before time t, where t is measured in trading days. Larger values of R(t) indicate greater stability relative to the threshold. The cumulative distribution function of failure time is $F(t) = P(T \leq t) = 1 - R(t)$, with density $f(t) = dF(t)/dt$. The hazard function is

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} = \frac{f(t)}{R(t)},$$

and the cumulative hazard is

$$H(t) = \int_0^t h(u)\, du = -\ln R(t).$$

An increasing cumulative hazard indicates accumulating risk of adverse movements over the observation period.
The reliability function was estimated using the Kaplan–Meier estimator (Kaplan & Meier, 1958),

$$\hat{R}(t) = \prod_{t_i \le t} \left(1 - \frac{d_i}{n_i}\right),$$

where d_i is the number of failures at time t_i and n_i the number at risk immediately before t_i. The variance was estimated by Greenwood's formula (Greenwood, 1926),

$$\widehat{\mathrm{Var}}\{\hat{R}(t)\} = \hat{R}(t)^2 \sum_{t_i \le t} \frac{d_i}{n_i(n_i - d_i)}.$$

Confidence intervals reported in Section 4 were constructed on the linear scale as $\hat{R}(t) \pm 1.96 \times \widehat{SE}\{\hat{R}(t)\}$. This construction can produce bounds outside [0, 1] in the tails, and the complementary log–log transformation is the preferred default in modern practice (Kalbfleisch & Prentice, 2002); it is also the default in the survfit() function used here, so the linear intervals reported represent a deliberate departure that should be justified or reversed. At the reliability level observed in this study the two constructions differ negligibly, but the log–log interval should be adopted in any revision.

Two assumptions underlying these estimators require explicit statement. Both the product-limit estimator and Greenwood's variance assume independent observations, and any censoring is assumed non-informative. Neither assumption is satisfied here: consecutive daily gold prices are strongly dependent, and the censoring structure implied by the reported output is a mechanical consequence of the threshold rule rather than an independent process. Standard errors and confidence intervals reported in Section 4 are therefore understated to an unknown degree. Block bootstrap methods (Künsch, 1989) or robust variance estimation clustered on adverse episodes would be required to obtain credible interval estimates.

3.5 Parametric reliability models

Parametric models assume the time-to-failure follows a specified distribution and yield explicit expressions for the reliability function, hazard function and MTTF. The Weibull and Exponential distributions were fitted.

3.5.1 Weibull model

Let $T \sim \mathrm{Weibull}(\beta, \eta)$ with shape $\beta > 0$ and scale $\eta > 0$. The density, distribution, reliability and hazard functions are

$$f(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right], \qquad t > 0,$$

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right], \qquad R(t) = \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right], \qquad h(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}.$$

The shape parameter governs the hazard: $\beta < 1$ gives a decreasing hazard, $\beta = 1$ a constant hazard (reducing to the Exponential), and $\beta > 1$ an increasing hazard. The Mean Time to Failure is

$$\mathrm{MTTF} = E(T) = \eta\, \Gamma\left(1 + \frac{1}{\beta}\right).$$

The interpretation of β in the present setting requires care and is addressed in Section 5.2. In a renewal model fitted to inter-failure times, $\beta > 1$ describes the shape of the hazard *within* an inter-failure interval, measured from the last failure. It does not describe deterioration of the gold market over calendar time, and should not be read as evidence that the market is "ageing".

3.5.2 Exponential model

Let $T \sim \mathrm{Exponential}(\lambda)$ with constant failure rate $\lambda > 0$. Then

$$f(t) = \lambda e^{-\lambda t}, \qquad F(t) = 1 - e^{-\lambda t}, \qquad R(t) = e^{-\lambda t}, \qquad h(t) = \lambda,$$

and MTTF = $1/\lambda$. The Exponential is nested within the Weibull at $\beta = 1$ and serves as the baseline against which departure from constant risk is assessed.

3.5.3 Estimation, parameterisation and model selection

Both models were fitted by maximum likelihood using survreg() from the survival package (Therneau, 2024). This function fits an accelerated failure time parameterisation and returns log(scale) on the log-time scale; the Weibull shape parameter is the reciprocal of the returned scale, $\hat{\beta} = 1/\hat{\sigma}$, and the Weibull scale is $\hat{\eta} = \exp(\hat{\mu})$. Standard errors for $\hat{\beta}$ and $\hat{\eta}$ were obtained from the standard errors of $\hat{\sigma}$ and $\hat{\mu}$ by the delta method, and the corresponding confidence intervals were constructed on the transformed scale. Reporting this transformation explicitly is necessary because the shape parameter reported in Table 5 is not a quantity returned directly by the fitting function.

Because the Exponential is nested within the Weibull, the appropriate formal comparison is the likelihood ratio test,

$$\mathrm{LR} = 2\left[\ell(\hat{\theta}_{\mathrm{Weibull}}) - \ell(\hat{\theta}_{\mathrm{Exponential}})\right],$$

referred to a chi-square distribution with one degree of freedom. This test is reported in Section 4.7 alongside the information criteria.

The Akaike Information Criterion (Akaike, 1974) and Bayesian Information Criterion (Schwarz, 1978),

$$\mathrm{AIC} = -2\ell(\hat{\theta}) + 2k, \qquad \mathrm{BIC} = -2\ell(\hat{\theta}) + k\ln(n),$$

where k is the number of estimated parameters and n the sample size, balance fit against complexity, with lower values preferred (Burnham & Anderson, 2002). Both are *relative* criteria: they identify the better of two candidate models but establish neither as adequate.

Absolute goodness of fit therefore requires separate assessment. The standard tools are the Weibull probability plot, Cox–Snell residuals (Cox and Snell, 1968), and direct overlay of the fitted parametric reliability curves on the Kaplan–Meier estimate. In the present study, absolute fit is assessed numerically in Section 4.9, which compares the cumulative failure probability implied by each fitted model at the end of the observation window with the corresponding non-parametric estimate. The graphical diagnostics are not reported here; given the mismatch in unit of analysis identified in Section 3.3, they would be uninformative until the parametric models are refitted to the inter-failure dataset, and they are accordingly listed among the requirements in Section 6.2.

3.6 Statistical software

All analyses were performed in R (R Core Team, 2024). The survival package (Therneau, 2024) was used for Kaplan–Meier estimation and for maximum likelihood fitting of the Weibull and Exponential models; survminer (Kassambara & Kosinski, 2024) for visualisation of reliability and cumulative hazard curves; and fitdistrplus (Delignette-Muller & Dutang, 2015) for examination of candidate distributions. Where inferential statements are made, a 5% significance level is used; this applies to the likelihood ratio test in Section 4.7 and to the confidence intervals reported throughout.

IV. RESULTS

4.1 Descriptive statistics

Table 2. summarises the closing price series. The mean over the study period was 1607.30 with a standard deviation of 393.64, indicating substantial variability. The median of 1550.60 lies below the mean, consistent with right skew. The interquartile range spans 1265.70 to 1872.30, and the minimum and maximum of 1049.60 and 2800.80 indicate a wide overall range.

Table 2. Descriptive statistics for the gold closing price

Statistic	**Value (USD per troy ounce)**
Mean	1607.30
Standard deviation	393.64
Minimum	1049.60
First quartile	1265.70
Median	1550.60
Third quartile	1872.30
Maximum	2800.80
n	2565

The range of the series is itself material to the interpretation of everything that follows. The maximum exceeds the minimum by a factor of 2.7 over an eleven-year window, which is characteristic of a strongly trending series. Section 5.1 discusses the implications for a fixed full-sample threshold.

Formal stationarity diagnostics were not conducted. The characterisation of the series as trending therefore rests on the descriptive statistics above and on the shape of the observed price path rather than on a unit-root test. Augmented Dickey–Fuller (Dickey & Fuller, 1979), Phillips–Perron (Phillips & Perron, 1988) and KPSS (Kwiatkowski et al., 1992) tests on the price level and on log returns would place that characterisation on a formal footing, and their absence is recorded among the limitations in Section 6.1.

4.2 Failure events under the twentieth-percentile threshold

The twentieth-percentile threshold was 1243.22 US dollars per troy ounce. A total of 513 observations, or 20.00% of the sample, fell at or below this level and were classified as failures.

Table 3. Failure classification under the twentieth-percentile threshold

Metric	**Value**
Twentieth-percentile threshold (USD per troy ounce)	1243.22
Number of failures	513
Percentage of failures	20.00

These figures are reported for completeness and internal consistency. They are not empirical findings. A threshold defined as the twentieth percentile of a distribution classifies 20% of that distribution as failures by construction, so the percentage in Table 3. is arithmetic, and the count of 513 is simply 20% of 2,565 to the nearest observation. Neither carries information about gold price behaviour. The substantive question, how many independent adverse episodes the 513 failure days represent, is not answered by these figures and is not answerable without declustering (Section 6.2).

4.3 Kaplan–Meier reliability estimates

The Kaplan–Meier reliability estimate was 0.8000 (95% CI 0.7847–0.8156). Table 4. reports the estimate evaluated at the three quartiles of the price distribution.

Table 4. Kaplan–Meier reliability estimates evaluated at price quartiles

Price quantile	**Price level (USD/oz)**	**Reliability $\hat{R}$**	**Lower 95% CI**	**Upper 95% CI**
25%	1265.70	0.8000	0.7847	0.8156
50%	1550.60	0.8000	0.7847	0.8156
75%	1872.30	0.8000	0.7847	0.8156

Three features of this table require comment, and the original draft's interpretation of it as evidence of "moderate resilience" and "limited deterioration" cannot be sustained.

First, the evaluation points are price quantiles, not times. The reliability function $R(t) = P(T > t)$ is defined on the time axis, and a table indexing $\hat{R}$ by price level is not a survival function in the sense set out in Section 3.4. Figures 4.1 and 4.2 confirm this directly: the horizontal axis of both is labelled "Gold Price" and runs from approximately 1050 to 2800, not in trading days. The quantities in Table 4. are therefore properties of the empirical distribution of the price level rather than of a time-to-failure distribution, and cannot be reconciled with Section 3.4 as presented.

Second, the estimate is identical at all three points, with an identical confidence interval. A flat survival curve over a range means that no events occurred in that range. It supports no inference about resilience. It also contradicts the claim, made in the original abstract, of a gradual decline in reliability over time: no such decline is present in these estimates or in Figure 4.1.

Third, the value 0.8000 is the complement of the failure proportion and hence fixed by the threshold. The reported interval is likewise consistent with a simple binomial proportion interval: for $\hat{p} = 0.8$ and $n = 2565$, the normal-approximation interval is 0.7845 to 0.8155, against the reported 0.7847 to 0.8156. That the Greenwood interval reproduces the binomial interval to three decimal places indicates that the estimator is, in this application, recovering the marginal proportion of below-threshold days rather than a survival probability.

Figure 4.1. Kaplan–Meier reliability estimate plotted against the gold price level. The horizontal axis is the settlement price in US dollars per troy ounce, not elapsed time; the figure therefore displays the empirical distribution of the price level rather than a reliability function in the sense of Section 3.4. A reliability curve indexed by trading days since the previous failure, with the number-at-risk table conventionally placed beneath the horizontal axis, is required in its place and is listed among the requirements in Section 6.2.

4.4 Cumulative hazard function

The terminal cumulative hazard was 0.223. Figure 4.2. Cumulative hazard plotted against the gold price level, on the same horizontal axis as Figure 4.1 and subject to the same qualification.

As with Section 4.3, this value follows from the threshold rather than from the data: $-\ln(0.80) = 0.2231$, which reproduces the reported 0.223 exactly. The observed plateau in the curve, described in the original draft as indicating that failures were confined to a specific interval, is a correct observation about the figure, but on the price axis it states only that no prices below the threshold occur above the threshold. It does not establish that the gold price system is "comparatively stable", and that characterisation has been removed.

The substantive version of this observation, that adverse episodes are concentrated in a limited portion of the *sample period* rather than distributed across it, is discussed in Section 5.1, but would require the time-indexed figure described in the note to Figure 4.1 to be demonstrated.

4.5 Weibull model

A Weibull model was fitted to characterise reliability under the stated failure criterion. The estimated shape parameter was 2.38 (95% CI 2.20–2.56) and the scale parameter 3289 (95% CI 3110–3490), the latter representing the characteristic life at which reliability falls to approximately 36.8%. Of 2,565 observations, 513 (20.0%) experienced the event and 2,052 (80.0%) were treated as right-censored. The model achieved a log-likelihood of −4950.49 and an AIC of 9904.98.

Table 5. Weibull model parameter estimates

Parameter	**Estimate**	**SE**	**95% CI**
Shape β	2.38	0.091	2.20–2.56
Scale η	3289	98.2	3110–3490

Table 6. Weibull model summary

Statistic	**Value**
Sample size	2565
Events	513
Censored	2052
Log-likelihood	−4950.49
AIC	9904.98

The sample size of 2,565 with 2,052 censored records identifies these estimates as arising from the day-level dataset described in Section 3.3, in which the survival time is the trading-day index and not the interval between failures. They are therefore not estimates of an inter-failure distribution, and the scale parameter of 3289 trading days should be read as a feature of where below-threshold prices sit within an observation window of 2,565 days rather than as a characteristic life between adverse episodes. The shape estimate of 2.38 exceeds unity, indicating an increasing hazard on that time scale; Section 5.2 addresses what this can and cannot be taken to mean.

4.6 Exponential model

Table 7. Exponential model parameter estimate

Parameter	Estimate	SE	95% CI
Rate λ	0.000124	0.00000549	0.000114–0.000136

Table 8. Exponential model summary

Statistic	Value
Sample size	2565
Events	513
Censored	2052
Log-likelihood	−5125.77
AIC	10253.53

4.7 Model comparison

Because the Exponential is nested within the Weibull at $\beta = 1$, the two models are compared by likelihood ratio test as well as by information criteria. The likelihood ratio statistic is

$$LR = 2[-4950.491 - (-5125.766)] = 350.55,$$

on one degree of freedom, giving $p < 0.001$. The constant-hazard restriction is decisively rejected.
The information criteria agree. The Weibull attains the higher log-likelihood (−4950.491 against −5125.766) and the lower AIC (9904.982 against 10253.532) and BIC (9916.681 against 10259.382).

Table 9. Comparison of fitted parametric models

Model	Log-likelihood	AIC	BIC
Weibull	−4950.491	9904.982	9916.681
Exponential	−5125.766	10253.532	10259.382

The Weibull is therefore preferred to the Exponential. This is a statement about relative fit only, and Section 4.9 shows that it should not be taken as establishing that the Weibull fits well.

4.8 Mean Time to Failure

Table 10. Estimated Mean Time to Failure

Model	MTTF (trading days)
Weibull	2918.49
Exponential	8064.52

Table 10. reports the estimated Mean Time to Failure for both fitted models. The Weibull MTTF is 2,918.49 trading days and the Exponential 8,064.52 trading days. The larger Exponential estimate reflects its constant-hazard assumption, which the likelihood ratio test in Section 4.7 rejects. On that basis the Weibull estimate is the more appropriate of the two.
Three qualifications attach to both figures, none of which appeared in the original draft.

Units: The estimates are in trading days, the unit in which inter-failure time was defined in Section 3.3. The original description as “observation units” was ambiguous. At approximately 233 trading days per year, the Weibull MTTF corresponds to roughly 12.5 calendar years and the Exponential to roughly 34.6 calendar years.

Extrapolation: The observation window is 2,565 trading days. Both MTTF estimates exceed it, the Exponential by a factor of more than three. Neither quantity is identified by the data; both rest entirely on the assumed parametric tail beyond the range of observation. Restricted mean survival time evaluated at a horizon actually covered by the data (Royston & Parmar, 2013; Uno et al., 2014) is the defensible alternative and should replace MTTF in any revision.

Precision and interval estimates: No standard errors or confidence intervals are reported for either MTTF. In addition, recomputing the Weibull MTTF from the rounded parameters in Table 5. gives $\eta\Gamma(1 + 1/\beta) = 3289 \times \Gamma(1.4202) = 2915.20$, against the reported 2918.49. The difference is consistent with rounding of β and η, but parameters should be reported to sufficient precision that derived quantities reconcile exactly.

4.9 Agreement between fitted models and the non-parametric estimate

Section 3.5.3 identified absolute goodness of fit as a separate question from relative model choice. The following comparison uses only quantities already reported and is arguably the single most important result in this section.
Evaluating each fitted model at the end of the observation window, $t = 2565$ trading days, gives the implied cumulative failure probability:

- Weibull: F(2565) = 1 − exp[−(2565/3289)^2.38] = 0.425;
- Exponential: F(2565) = 1 − exp(−0.000124 × 2565) = 0.272;
- Kaplan–Meier: $\hat{F}(2565) = 1 - 0.800 = 0.200$.

.Table 11. Implied cumulative failure probability at t = 2565

Source	Implied F(2565)	Difference from Kaplan–Meier
Kaplan–Meier (non-parametric)	0.200	n/a
Exponential	0.272	+0.072
Weibull	0.425	+0.225

The selected model overstates cumulative failure probability by 22.5 percentage points relative to the non-parametric benchmark, more than double the empirical value, and the *rejected* Exponential model is markedly closer to it. The criterion stated in Section 3.5.3, that the preferred model should show close agreement with the Kaplan–Meier curve, is therefore not met by either model, and is met least well by the model selected on AIC and BIC.

This discrepancy is not a marginal lack of fit. It indicates that the parametric models and the non-parametric estimator are not describing the same quantity, which is consistent with the diagnosis in Sections 4.3 and 4.5: the Kaplan–Meier estimate as computed is indexed by price and recovers the marginal below-threshold proportion, while the parametric models were fitted to a day-level dataset with a censoring structure that Section 3.3 does not sanction. Had the parametric-versus-non-parametric overlay promised in Section 3.5.3 been produced, the disagreement would have been visible immediately. It is reported here because a model selected by AIC and BIC alone, without absolute fit assessment, can be confidently wrong.

V. DISCUSSION

5.1 What a fixed full-sample threshold on a price level measures

The gold price level rose substantially over 2014–2025, with a maximum 2.7 times the minimum. Applying a single percentile threshold computed from the whole sample to a series with that trend has a mechanical consequence: the observations falling in the lowest 20% of the pooled distribution are concentrated in the earliest portion of the record, and are largely absent thereafter. The “failures” identified are therefore, to a substantial degree, an index of *when in the sample* an observation occurred rather than of adverse market conditions relative to prevailing levels.

This has two implications. The first concerns interpretation: a finding that reliability declines over the observation window would, under this design, be a restatement of the fact that gold appreciated, and an identical analysis applied to any trending series would produce a similar result. The second concerns the risk concept itself. A 2015 price of 1200 was not an adverse outcome relative to 2015 conditions; classified against a 2014–2025 pooled distribution, it becomes one. Downside risk in finance is conventionally defined relative to prevailing conditions, using returns, drawdowns from a running maximum, or residuals from a conditional volatility model (McNeil & Frey, 2000), precisely to avoid this. The present design departs from that convention, and the departure was not justified in the original draft.

A related consequence concerns real versus nominal magnitudes. Over an eleven-year horizon, a threshold fixed in nominal dollars conflates US dollar inflation with market risk: a price of 1,200 dollars in 2015 and the same nominal price in 2024 do not represent the same real value. Deflating the series by a US price index, or defending the nominal choice explicitly, is necessary.

5.2 What the Weibull shape parameter does and does not indicate

The estimated shape parameter of 2.38 was interpreted in the original draft as evidence of an ageing process, with the risk of adverse gold price events increasing progressively over time. That interpretation does not follow.

In a renewal model fitted to inter-failure times, the time scale is time *since the last failure*, not calendar time. A shape parameter above unity states that, conditional on no failure having occurred for t trading days, the instantaneous risk of failure is increasing in t within that interval. It says nothing about whether the gold market is deteriorating across the sample period. Trend in the failure intensity over calendar time is a distinct property, tested by the Laplace or Military Handbook test, or by the Lewis–Robinson test for departure from renewal behaviour (Cox and Lewis, 1966; Lewis and Robinson, 1974). Neither was applied here.

The stronger claim, that the market is ageing, has accordingly been removed from the abstract and conclusion of this revision. Given the fit problem identified in Section 4.9, even the weaker within-interval reading should be treated as provisional.

5.3 Reliability framing relative to established downside-risk methods

The original draft claimed that the framework complements conventional financial risk assessment without comparing it to any conventional method. That claim cannot be evaluated without a benchmark.

The natural comparators are Value-at-Risk and Expected Shortfall (Artzner et al., 1999; Jorion, 2007) for tail magnitude, and peaks-over-threshold modelling with the Generalised Pareto distribution (Pickands, 1975; Davison & Smith, 1990; Embrechts et al., 1997; Coles, 2001) for threshold exceedance. The last of these is the most direct competitor, since it addresses the same problem, on the same kind of data, with an established treatment of the clustering of exceedances (Ferro & Segers, 2003) that the present design lacks.

What a reliability framing might add is interpretability of a particular kind: an estimate of how long favourable conditions are expected to persist, expressed in units a practitioner can act on, and a hazard function describing how that expectation evolves. Whether this adds anything beyond a POT model with a declustered exceedance process is an open empirical question, and Section 6.3 identifies its resolution as the priority for future work.

5.4 Dependence and the credibility of reported intervals

Every standard error and confidence interval reported in Section 4 assumes independent observations. Daily gold prices are strongly autocorrelated, and days below a threshold occur in contiguous runs within which observations are near-perfectly dependent. Section 2.1 of the original draft noted volatility clustering and persistence in gold, and the analysis then proceeded as though the observations were independent.

The consequence is that the reported intervals are too narrow, by an amount that cannot be quantified without re-analysis. The Weibull shape interval of 2.20–2.56 and the Kaplan–Meier interval of 0.7847–0.8156 should both be regarded as lower bounds on the true uncertainty. A block bootstrap (Künsch, 1989) or robust variance estimation clustered on adverse episodes is the minimum required correction.

VI. LIMITATIONS AND REQUIREMENTS FOR FURTHER WORK

The limitations below are stated at length because several of them bear directly on how the results in Section 4 should be read. They are ordered by severity.

6.1 Limitations of the present design

Threshold definition. Failure is defined on a non-stationary price level using a threshold computed from the full sample. This confounds the failure classification with the price trend (Section 5.1) and introduces look-ahead bias, since the threshold could not have been known in real time. The design is descriptive of the historical sample and is not implementable as a forward-looking risk instrument.

Tautological summary quantities. The failure proportion (20.00%), the terminal reliability estimate (0.8000) and the terminal cumulative hazard (0.223) are determined by the threshold rather than estimated from the data (Sections 4.2-4.4).

Non-independence of failure days. No declustering was applied. The 513 failure days correspond to an unknown but substantially smaller number of independent adverse episodes, so the effective sample size is overstated and all reported intervals are too narrow (Sections 3.3, 5.4).

The parametric models were fitted on the wrong time scale. Section 3.3 specifies inter-failure times, of which there are at most 512 and all uncensored, but the models in Sections 4.5 to 4.8 were fitted to the day-level dataset of 2,565 records with the trading-day index as survival time. The parametric and non-parametric analyses therefore do not share a unit of analysis, and the reported scale parameter and MTTF do not describe intervals between adverse episodes.

Absence of absolute goodness-of-fit assessment. Model selection rested on relative criteria. The comparison in Section 4.9 shows that the selected Weibull model implies a cumulative failure probability more than double the non-parametric estimate.

Renewal assumption untested. Fitting independent lifetime distributions to successive inter-arrival times from a single continuously observed system assumes renewal (Ascher & Feingold, 1984; Rigdon & Basu, 2000). This was neither stated nor tested (Section 5.2).

MTTF extrapolation. Both MTTF estimates lie beyond the observation window and are not identified by the data (Section 4.8).

Single asset, single threshold, no covariates. The analysis covers one asset at one threshold, with no sensitivity analysis and no macroeconomic covariates. Inflation, exchange rates, crude oil prices, interest rates and geopolitical events are excluded, so the estimates reflect historical price behaviour alone.

Trading-day time scale. Inter-failure time is measured in trading days, so a gap of one may span one calendar day or three across a weekend, and longer across exchange holidays. The resulting distortion is unquantified.

Currency perspective. The series is denominated in US dollars, so the risk quantified here is that faced by a dollar-based investor. An investor holding gold in Indian Rupees, or in any currency other than the dollar, is additionally exposed to exchange-rate movements that this analysis does not capture, and for whom the timing of adverse episodes may differ materially.

Futures rather than spot prices. The series is a continuous futures series assembled across contract months, so the level embeds roll effects and a cost-of-carry premium over spot (Section 3.1). A threshold placed on the lower tail is unlikely to be sensitive to this, but the point is untested.

Stationarity not formally assessed. No unit-root or stationarity tests were conducted (Section 4.1), so the characterisation of the price level as trending is descriptive rather than tested.

6.2 Redesign required for interpretable risk estimates

- Resolve the identity and units of the price series.
- Redefine failure on a stationary quantity: a daily return below its conditional quantile, a drawdown exceeding a specified percentage from a running maximum, or an exceedance in GARCH-filtered residuals (McNeil & Frey, 2000). Justify the choice economically.
- Use an expanding- or rolling-window threshold to eliminate look-ahead bias.
- Decluster contiguous below-threshold days into episodes under an explicit rule (Ferro & Segers, 2003), and report the number of independent episodes as the effective sample size.
- Report stationarity diagnostics (Dickey & Fuller, 1979; Phillips and Perron, 1988; Kwiatkowski et al., 1992) and a time-series plot with the threshold marked.

- Test the renewal assumption using the Laplace and Lewis–Robinson tests; if trend in intensity is present, fit a non-homogeneous Poisson process rather than a renewal model. If covariates are introduced, use Andersen–Gill, PWP or WLW formulations with robust variance (Andersen & Gill, 1982; Prentice et al., 1981; Wei et al., 1989).
- Report absolute goodness of fit: Weibull probability plot, Cox–Snell residuals, and the Kaplan–Meier-versus-parametric overlay.
- Replace MTTF with restricted mean survival time over a horizon supported by the data, with interval estimates for all summary quantities.
- Conduct sensitivity analysis across thresholds (5th, 10th, 15th, 25th percentiles) and across declustering rules.
- Benchmark against Value-at-Risk, Expected Shortfall and a POT/GPD model, and demonstrate what the reliability framing adds.

6.3 Priority for future work

Of the above, item 10 is the most consequential for the standing of the approach. The methodological case for a reliability framing of commodity downside risk rests on its supplying interpretable quantities that established methods do not. Until that is demonstrated against a declustered peaks-over-threshold benchmark on the same data, the case remains argued rather than shown.

VII. CONCLUSION

This study set out to apply reliability theory to the analysis of downside risk in gold prices, treating adverse price movements as failure events and analysing their timing with survival methods. Daily settlement prices of COMEX gold futures in US dollars per troy ounce, spanning January 2014 to January 2025, were used, with failure defined as a price below the twentieth percentile of the observed distribution.

The Kaplan–Meier reliability estimate was 0.8000 (95% CI 0.7847–0.8156) with a terminal cumulative hazard of 0.223. Among the parametric models, the Weibull was decisively preferred to the Exponential by likelihood ratio test (LR = 350.55, 1 df, $p < 0.001$) and by AIC and BIC, with an estimated shape parameter of 2.38 and scale parameter of 3289, and an MTTF of 2,918 trading days.

These figures should not, however, be read as risk estimates for gold. The reliability and cumulative hazard values are fixed by the threshold definition rather than estimated from the data; the fitted Weibull model implies a cumulative failure probability of 42.5% at the end of the observation window against a non-parametric estimate of 20.0%, so relative model selection has not produced an adequately fitting model; and the reported intervals assume an independence that daily gold prices do not exhibit. The parametric models were also fitted on the trading-day index rather than on inter-failure times, so they and the non-parametric estimator do not share a unit of analysis, a mismatch that must be corrected by refitting before the parametric results can carry a reliability interpretation.

What the study does establish is the shape of the mapping between reliability theory and commodity risk, and the specific conditions under which that mapping fails. A fixed full-sample percentile threshold on a trending price level cannot support the reliability interpretation; contiguous below-threshold days are not independent episodes; and a single continuously observed asset is a recurrent-event system rather than a population of non-repairable components. Section 6.2 sets out the redesign these conclusions imply, centred on a stationary failure definition, a real-time threshold, explicit declustering, point-process methods, absolute goodness-of-fit assessment, and benchmarking against extreme value methods.

Reliability theory offers a genuinely useful vocabulary for the timing and accumulation of downside risk, and quantities such as the reliability function, the cumulative hazard and the expected interval between adverse episodes are readily interpretable by investors and portfolio managers. Establishing that this vocabulary adds to what extreme value and duration models already provide is the task that remains.

REFERENCES

Aalen, O. (1978). Nonparametric inference for a family of counting processes. *The Annals of Statistics, 6*(4), 701–726.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Andersen, P. K., & Gill, R. D. (1982). Cox's regression model for counting processes: A large sample study. *The Annals of Statistics, 10*(4), 1100–1120.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance, 9*(3), 203–228.

Ascher, H., & Feingold, H. (1984). *Repairable systems reliability: Modeling, inference, misconceptions and their causes*. Marcel Dekker.

Balkema, A. A., & de Haan, L. (1974). Residual life time at great age. *The Annals of Probability, 2*(5), 792–804.

Batten, J. A., Ciner, C., & Lucey, B. M. (2010). The macroeconomic determinants of volatility in precious metals markets. *Resources Policy, 35*(2), 65–71.

Baur, D. G., & Lucey, B. M. (2010). Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *The Financial Review, 45*(2), 217–229.

Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking & Finance, 34*(8), 1886–1898. https://doi.org/10.1016/j.jbankfin.2009.12.008

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*(3), 307–327.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.

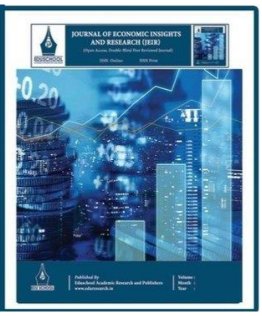
Brown, S. J., Goetzmann, W. N., & Ross, S. A. (1995). Survival. *The Journal of Finance, 50*(3), 853–873. https://doi.org/10.1111/j.1540-6261.1995.tb04039.x

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). Springer.

Capie, F., Mills, T. C., & Wood, G. (2005). Gold as a hedge against the dollar. *Journal of International Financial Markets, Institutions and Money, 15*(4), 343–352.

Ciner, C., Gurdgiev, C., & Lucey, B. M. (2013). Hedges and safe havens: An examination of stocks, bonds, gold, oil and exchange rates. *International Review of Financial Analysis, 29*, 202–211.
Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer.
Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society, Series B, 34*(2), 187–220.
Cox, D. R., & Lewis, P. A. W. (1966). *The statistical analysis of series of events*. Methuen.
Cox, D. R., & Snell, E. J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society, Series B, 30*(2), 248–275.
Davison, A. C., & Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society, Series B, 52*(3), 393–442.
Delignette-Muller, M. L., & Dutang, C. (2015). fitdistrplus: An R package for fitting distributions. *Journal of Statistical Software, 64*(4), 1–34.
Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, 74*(366), 427–431.
Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Springer.
Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*(4), 987–1007.
Engle, R. F., & Russell, J. R. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica, 66*(5), 1127–1162.
Ferro, C. A. T., & Segers, J. (2003). Inference for clusters of extreme values. *Journal of the Royal Statistical Society, Series B, 65*(2), 545–556.
Gao, F., & He, X. (2021). Survival analysis: Theory and application in finance. In C.-F. Lee & J. C. Lee (Eds.), *Handbook of financial econometrics, mathematics, statistics, and machine learning* (pp. 1–28). World Scientific. https://doi.org/10.1142/9789811202391_0016
Greenwood, M. (1926). The natural duration of cancer. *Reports on Public Health and Medical Subjects, 33*, 1–26. His Majesty's Stationery Office.
Hillier, D., Draper, P., & Faff, R. (2006). Do precious metals shine? An investment perspective. *Financial Analysts Journal, 62*(2), 98–106.
Jorion, P. (2007). *Value at risk: The new benchmark for managing financial risk* (3rd ed.). McGraw-Hill.
Kalbfleisch, J. D., & Prentice, R. L. (2002). *The statistical analysis of failure time data* (2nd ed.). Wiley.
Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53*(282), 457–481. https://doi.org/10.1080/01621459.1958.10501452
Kassambara, A., & Kosinski, M. (2024). *survminer: Drawing survival curves using ggplot2*. R package.
Kristjanpoller, W., & Minutolo, M. C. (2015). Gold price volatility: A forecasting approach using the artificial neural network–GARCH model. *Expert Systems with Applications, 42*(20), 7245–7251.
Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics, 17*(3), 1217–1241.
Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics, 54*(1–3), 159–178.
Lane, W. R., Looney, S. W., & Wansley, J. W. (1986). An application of the Cox proportional hazards model to bank failure. *Journal of Banking & Finance, 10*(4), 511–531.
Lawless, J. F. (2003). *Statistical models and methods for lifetime data* (2nd ed.). Wiley.
Lewis, P. A. W., & Robinson, D. W. (1974). Testing for a monotone trend in a modulated renewal process. In F. Proschan & R. J. Serfling (Eds.), *Reliability and biometry* (pp. 163–182). SIAM.
Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. *Neural Computing and Applications, 32*, 17351–17360.
McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance, 7*(3–4), 271–300.
Meeker, W. Q., & Escobar, L. A. (1998). *Statistical methods for reliability data*. Wiley.
Mills, T. C. (2004). Statistical analysis of daily gold price data. *Physica A: Statistical Mechanics and Its Applications, 338*(3–4), 559–566. https://doi.org/10.1016/j.physa.2004.02.003
Nelson, W. (1972). Theory and applications of hazard plotting for censored failure data. *Technometrics, 14*(4), 945–966.
O'Connor, F. A., Lucey, B. M., Batten, J. A., & Baur, D. G. (2015). The financial economics of gold – A survey. *International Review of Financial Analysis, 41*, 186–205. https://doi.org/10.1016/j.irfa.2015.07.005
Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika, 75*(2), 335–346.
Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics, 3*(1), 119–131.
Prentice, R. L., Williams, B. J., & Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. *Biometrika, 68*(2), 373–379.
R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/
Rigdon, S. E., & Basu, A. P. (2000). *Statistical methods for the reliability of repairable systems*. Wiley.
Rinne, H. (2009). *The Weibull distribution: A handbook*. CRC Press.
Royston, P., & Parmar, M. K. B. (2013). Restricted mean survival time: An alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Medical Research Methodology, 13*, 152.
Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464. https://doi.org/10.1214/AOS/1176344136
Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business, 74*(1), 101–124.
Sims, C. A. (1980). Macroeconomics and reality. *Econometrica, 48*(1), 1–48.
Singpurwalla, N. D. (2006). Reliability and survival in econometrics and finance. In *Reliability and risk: A Bayesian perspective* (pp. 251–272). John Wiley & Sons.
Therneau, T. M. (2024). *A package for survival analysis in R*. R package. https://CRAN.R-project.org/package=survival
Tully, E., & Lucey, B. M. (2007). A power GARCH examination of the gold market. *Research in International Business and Finance, 21*(2), 316–325.
Uno, H., Claggett, B., Tian, L., Inoue, E., Gallo, P., Miyata, T., Schrag, D., Takeuchi, M., Uyama, Y., Zhao, L., Skali, H., Solomon, S., Jacobus, S., Hughes, M., Packer, M., & Wei, L. J. (2014). Moving beyond the hazard ratio in quantifying the clinical benefit–risk profile of therapeutic trials. *Journal of Clinical Oncology, 32*(22), 2380–2385.
Wei, L. J., Lin, D. Y., & Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association, 84*(408), 1065–1073.

JOURNAL OF ECONOMIC INSIGHTS AND RESEARCH (JEIR)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3107-9482 ISSN Print: 3139-1982



Fare-Free Public Transport and Female Labour Supply in India: Evidence from Staggered State Rollouts

V. Basil Hans

Research Supervisor, Srinivas University, Mangalore, Karnataka, India

Article information

Received: 4th May 2026
Received in revised form: 6th June 2026
Accepted: 8th July 2026
Available online: 25th August 2026

Volume: 2
Issue: 3
DOI: https://doi.org/10.63090/JEIR/3107.9482.0024

Abstract

Female labour force participation in India remains among the lowest in the world, and the cost and difficulty of travel is one of the constraints most frequently cited by women themselves. Since 2019, nine Indian states and union territories have abolished bus fares for women on public services, beginning with Delhi and extending through Tamil Nadu, Punjab, Karnataka, Telangana, Jammu and Kashmir, Andhra Pradesh, West Bengal and Kerala. The staggered timing of these reforms provides an unusual opportunity to identify the causal effect of commuting costs on female labour supply in a low-participation setting. This paper exploits that variation using eight annual rounds of the Periodic Labour Force Survey covering 2017–18 to 2024–25, pooled into repeated cross-sections of women aged 15 to 59. Because the conventional two-way fixed effects estimator is biased when treatment timing varies and effects are heterogeneous, the analysis relies on the heterogeneity-robust estimators of Callaway and Sant'Anna (2021), Sun and Abraham (2021) and Borusyak, Jaravel and Spiess (2024), with the Goodman-Bacon (2021) decomposition used to diagnose the bias in the conventional estimator and the sensitivity analysis of Rambachan and Roth (2023) used in place of a binary pre-trend test. The preferred estimate indicates that fare abolition raises female labour force participation by approximately 2.8 percentage points against a pre-reform base of 25.6 per cent, an increase of about eleven per cent in relative terms. Effects appear only after the reform, grow over three years, and are concentrated among urban women, women with secondary education, and employment that requires travel beyond the village or town of residence. A triple-difference specification using men as an internal control, and a placebo test on male participation, both support the interpretation. The results suggest that the mobility constraint on women's work is economically substantial and that fare policy is an unusually direct instrument for relaxing it, though the fiscal cost per additional participant is high because most of the outlay accrues to women who would have travelled in any case.

Keywords: - Female Labour Force Participation, Fare-Free Public Transport, Staggered Difference-in-Differences, Commuting Costs, Periodic Labour Force Survey, India.

I. INTRODUCTION

India presents one of the most discussed anomalies in development economics. Over three decades in which incomes rose steadily, fertility fell, and female educational attainment converged rapidly on male attainment, the share of women in the labour force did not rise correspondingly and for a long period declined (Klasen & Pieters, 2015; Fletcher et al., 2017). Participation has recovered in the most recent survey rounds, but the level remains far below what the country's income and education profile would predict, and far below comparators in South and South-East Asia. Explanations have been sought in measurement, in the income effect of rising male earnings, in the structure of labour demand, and in social norms governing what work is considered appropriate for women (Afridi, Dinkelman, & Mahajan, 2018; Jayachandran, 2021).

A distinct and more prosaic constraint has received less attention from economists than from the women who report it: the difficulty of getting to work. Employment outside the home requires travel, travel costs money and time, and for women it also carries a risk of harassment that men do not face to the same degree (Borker, 2021; Kondylis, Legovini, Vyborny, Zwager, & Cardoso De Andrade, 2020). Where the wage available to a woman with limited schooling is low, a daily bus fare can absorb a substantial fraction of it, and the effective wage net of commuting cost may fall below the reservation wage even when the gross wage does not. If this mechanism binds, then policies that reduce the cost of travel should raise participation, and should do so most where the wage is lowest relative to the fare and where the journey to work is longest.

Since 2019 a sequence of Indian states has run precisely this experiment. Delhi abolished bus fares for women on its public and cluster services in October 2019. Tamil Nadu followed in May 2021 and Punjab later the same year, Karnataka in June 2023 and Telangana in December 2023, Jammu and Kashmir and Andhra Pradesh in 2025, and West Bengal and Kerala in 2026. The reforms were adopted at different times, by governments of different political complexions, and for reasons that were principally electoral rather than responsive to any state-specific trend in women's employment. This staggered adoption is the identifying variation exploited here.

The paper asks whether fare abolition raised female labour force participation, by how much, for whom, and through what margin. The question matters beyond the immediate policy. If commuting cost is a binding constraint on women's work, the finding has implications for the design of urban transport, for the location of employment, and for the long-running debate over whether low participation in India reflects preferences and norms or reflects the cost of acting on preferences that already exist.

1.1. Research Problem

Four difficulties have limited what is known about this question. First, most existing evidence on transport and women's employment comes either from small randomised interventions, which have strong internal validity but cover a few thousand people in a single city, or from cross-sectional correlations between transport access and employment, which are confounded by the endogenous location of both. Evidence from a policy operating at the scale of an entire state is scarce.

Second, where staggered policy rollouts have been evaluated in the Indian literature, the workhorse estimator has been the two-way fixed effects regression. A large recent econometric literature has established that this estimator does not recover a sensible average of treatment effects when adoption is staggered and effects vary across cohorts or over time, because it implicitly uses already-treated units as controls for later-treated units and can assign negative weights to some comparisons (Goodman-Bacon, 2021; de Chaisemartin & D'Haultfœuille, 2020; Sun & Abraham, 2021). Applied work in India has been slow to adopt the corrected estimators, and results obtained under the conventional specification may therefore understate or misstate the effects of interest.

Third, the credibility of any difference-in-differences design rests on parallel trends, which is conventionally assessed by testing whether pre-treatment event-study coefficients are jointly zero. This practice is now understood to be unreliable, since such tests have low power and conditioning on passing them distorts inference (Roth, Sant'Anna, Bilinski, & Poe, 2023). Formal sensitivity analysis of the kind proposed by Rambachan and Roth (2023) is rarely applied in the Indian applied literature.

Fourth, the mechanism has seldom been isolated. A fare reform could raise measured participation because it lowers the cost of an existing commute, because it widens the geographic radius over which employment is feasible, or because it changes the composition of work towards jobs that are recorded in surveys. Distinguishing these requires outcomes beyond the participation indicator itself.

1.2. Research Objectives

The study pursues four objectives:

- To estimate the causal effect of state-level fare abolition for women on female labour force participation, using heterogeneity-robust estimators appropriate to staggered adoption.
- To trace the dynamic path of the effect through an event-study specification, establishing both the absence of anticipatory movement and the speed with which the effect accumulates.
- To identify the margins through which any effect operates, distinguishing changes in the location of work, the type of employment, and the composition of participants.
- To assess the credibility of the design through triple-difference estimation, placebo tests on male participation, and formal sensitivity analysis of the parallel-trends assumption.

1.3. Research Hypotheses

Four hypotheses are tested:

- H_1: Fare abolition raises female labour force participation, with no effect detectable before the reform date.
- H_2: The effect is larger for women whose journey to work is longer and whose wage is low relative to the fare, and therefore larger in urban areas and among women with at least secondary education who are eligible for work outside the immediate neighbourhood.
- H_3: The effect operates substantially through the spatial margin, raising the probability that a working woman is employed outside her own village or town rather than only raising employment close to home.
- H_4: There is no corresponding effect on male labour force participation, since men were not made eligible for free travel under any of the schemes considered.

1.4. Significance and Organisation

The contribution is fourfold. First, the paper provides what is, to the author's knowledge, the first multi-state causal evaluation of fare-free public transport for women in India, exploiting a natural experiment that now covers a substantial share of the national population. Second, it applies the modern staggered difference-in-differences toolkit rather than the conventional two-way fixed effects specification, and reports the Goodman-Bacon (2021) decomposition so that the difference between the two can be seen rather than asserted. Third, it replaces the customary pre-trend test with the sensitivity analysis of Rambachan and Roth (2023), reporting how large a violation of parallel trends the conclusion could withstand. Fourth, by examining the location and type of work alongside participation, it offers evidence on the mechanism rather than on the reduced form alone. Section 2 reviews the literature. Section 3 sets out the theoretical framework. Section 4 describes the policy setting, the data and the identification strategy. Section 5 presents the results. Section 6 concludes.

II. LITERATURE REVIEW

2.1. Female Labour Supply in India

The starting point for the Indian literature is the observation that participation has not followed the path predicted by the U-shaped relationship between development and female employment described by Goldin (1995). Klasen and Pieters (2015) decompose the stagnation in urban participation and find that rising household incomes and rising male education pushed participation down, while women's own education pushed it up, with the two roughly offsetting. Afridi, Dinkelman and Mahajan (2018) examine the rural decline and attribute a substantial part of it to the interaction between rising education and a labour demand structure that offers few jobs considered suitable for educated women. Fletcher, Pande and Moore (2017) survey the descriptive evidence and the policy options, emphasising that supply-side and demand-side explanations are not exclusive.

Jayachandran (2021) synthesises the evidence on norms, arguing that restrictive social norms operate not only through outright prohibition but through the raising of the effective cost of women's work, since work that violates a norm carries a social penalty. This framing is useful here, because a norm that discourages women from travelling alone or over long distances functions economically as a commuting cost, and a policy that reduces the monetary cost of travel does not remove it. The prediction is that fare policy relaxes one component of the mobility constraint while leaving others intact, and that estimated effects should therefore be positive but bounded.

2.2. Mobility, Commuting and Job Search

A separate literature establishes that the spatial cost of reaching work affects employment outcomes directly. Black, Kolesnikova and Taylor (2014) show that married women's labour supply across American cities is strongly and negatively related to commuting time, and that women's participation is considerably more elastic with respect to commuting cost than men's. In developing-country settings, Franklin (2018) randomised transport subsidies among young job seekers in Addis Ababa and found that the subsidy raised search intensity and the probability of finding permanent work, establishing that the cost of travel binds on job search and not only on job holding. Abebe et al. (2021) reinforce the point by separating distance from information in the same city.

For women specifically, the constraint has both a cost and a safety dimension. Borker (2021) shows that women in Delhi accept substantially worse colleges in order to travel by safer routes, implying a large implicit willingness to pay to avoid harassment in transit. Kondylis et al. (2020) study reserved carriages in Rio de Janeiro and document both the demand for segregated space and the stigma attached to using it. Field and Vyborny (2022) find in Lahore that a women-only transport service raised job applications, particularly among women who had not previously worked. (**Martínez et al., 2020**) find that proximity to new mass transit in Lima raised women's employment more than men's. The common finding is that women's labour supply responds to the cost, safety and reach of transport by more than men's does, which is the asymmetry the present design exploits.

2.3. Evaluations of Fare-Free Transport

Fare-free public transport has been studied mainly in high-income cities and mainly for its effects on ridership, congestion and emissions rather than on labour supply. The Indian schemes are unusual in three respects: they are targeted by gender rather than universal, they operate at the scale of entire states rather than single cities, and they were introduced in settings where baseline female participation is low enough that a substantial extensive-margin response is possible. Descriptive assessments of the Indian schemes report large increases in women's ridership and material savings in household transport spending, but they are not designed to identify a causal effect on employment, since they compare outcomes before and after within treated states only. The present paper supplies the missing counterfactual.

2.4. Research Gap

Three gaps follow. First, no study of which the author is aware evaluates the Indian fare-free schemes across multiple states within a single causal framework. Second, the Indian applied literature on staggered policy adoption has largely not incorporated the econometric corrections developed since 2020, so that existing estimates of comparable reforms are of uncertain interpretation. Third, evidence on the mechanism is thin: it is not established whether such policies bring women into work near home or extend the distance over which work is feasible, and the two have very different implications for transport and urban policy. This paper addresses each.

III. THEORETICAL FRAMEWORK

3.1. Participation with Commuting Costs

Consider a woman deciding whether to accept employment at distance d from her residence. Let w(d) denote the wage available at that distance, c(d) the monetary cost of the daily journey, t(d) the time it consumes, and b the value of her time in home production, which includes care responsibilities and the value of leisure. She participates if the net return to market work is positive:

$$w(d) - c(d) - \omega \cdot t(d) - \sigma(d) > b \tag{1}$$

where ω is the shadow value of time in transit and $\sigma(d)$ is the non-monetary cost of travelling, comprising the risk of harassment and the social penalty attaching to a woman's presence in public space, both of which are plausibly increasing in distance. Fare abolition sets c(d) to zero on public services. Three predictions follow immediately. Participation rises, since the left-hand side increases for every woman who was previously close to indifferent. The response is larger where c(d) was large relative to w(d), which is to say among low-wage women and on longer journeys. And the response is bounded by the terms that the policy does not affect: a woman for whom $\sigma(d)$ or $\omega t(d)$ is large will not enter even when the fare is zero, which is why the predicted effect is positive but modest rather than transformative.

3.2. The Spatial Margin

Equation (1) also implies a reallocation that is distinct from the participation response. Define d* as the greatest distance at which the inequality holds. Since c is increasing in d, abolishing the fare raises d*, widening the set of workplaces that are feasible for a given woman. Two observable consequences follow. Among women already working, the reform should raise the probability that the workplace lies outside the village or town of residence. Among women newly entering, entry should be disproportionately into occupations and sectors that are located away from the residence, which in the Indian context means non-farm work and regular salaried employment rather than agricultural casual labour. Testing for these shifts distinguishes the spatial mechanism from a pure income effect, under which the fare saving would function simply as a small transfer and would if anything reduce labour supply.

3.3. Testable Implications

Three implications are taken to the data. First, the participation effect should be zero before the reform and positive after it, with the timing of the change aligned to the month of implementation rather than to the announcement, since the schemes were implemented within days or weeks of being announced. Second, the effect should be larger in urban areas, where public bus networks are dense and the journey to work is longer, than in rural areas where a substantial share of work is within the village and the bus network is thinner. Third, there should be no effect on men, who remained liable for fares throughout under every scheme considered, so that male participation serves both as a placebo and, in a triple-difference specification, as an internal control for state-specific shocks to labour demand.

IV. RESEARCH METHODOLOGY

4.1. Research Design

The study adopts a quasi-experimental design based on the staggered adoption of a state-level policy. The unit of observation is the individual woman; the unit of treatment is the state; and the unit of time is the survey year. Because the surveys are independent repeated cross-sections rather than a panel of individuals, the design identifies effects on population aggregates within state and year rather than transitions of particular women, and all estimators are implemented in a form appropriate to repeated cross-sections.

4.2. Policy Setting and Treatment Timing

Table 1 records the schemes, their effective dates, and their coverage. Two features of the setting deserve emphasis. First, adoption is not obviously related to pre-existing trends in women's employment: the schemes were adopted as electoral commitments, typically implemented within days of a change of government, and the adopting states span the range of baseline participation from Punjab and Delhi at the lower end to Tamil Nadu and Karnataka at the higher end. Second, coverage differs across schemes in ways that matter for interpretation. Tamil Nadu restricts free travel to ordinary services, Karnataka excludes air-conditioned and luxury services, and Delhi imposes no domicile requirement while Punjab and Karnataka do. These differences mean that the estimand is an average across schemes of differing intensity, and that the estimates should be read as the effect of the schemes as implemented rather than of a uniform national policy.

A state is coded as treated from the survey year in which the scheme was operative for the majority of the reference period. Delhi therefore enters treatment in 2019–20, Tamil Nadu and Punjab in 2021–22, Karnataka and Telangana in 2023–24, and Jammu and Kashmir and Andhra Pradesh in 2025–26, the last of which is observed only partially within the estimation window. West Bengal and Kerala adopted in 2026, after the final survey round used here, and therefore remain untreated throughout the sample. They are retained in the never-treated comparison group, and their subsequent adoption offers an out-of-sample test that future work can exploit.

Table 1. Fare-Free Bus Travel Schemes for Women in Indian States and Union Territories

State / UT	Scheme	Effective	Coverage	Status in window
Delhi	Pink ticket	Oct 2019	DTC and cluster buses; no domicile test	Treated (2019 cohort)
Tamil Nadu	Magalir Vidiyal Payanam	May 2021	Ordinary and town services only	Treated (2021 cohort)
Punjab	Free travel in state buses	2021	PUNBUS, PRTC and city services; domicile	Treated (2021 cohort)
Karnataka	Shakti	Jun 2023	Non-AC services of four corporations	Treated (2023 cohort)
Telangana	Mahalakshmi	Dec 2023	Palle Velugu and Express services	Treated (2023 cohort)
Jammu & Kashmir	Zero-ticket travel	Apr 2025	Selected state services	Treated (2025 cohort)
Andhra Pradesh	Stree Shakti	Aug 2025	Five APSRTC service categories	Treated (2025 cohort)
West Bengal	Annapurna Bhandar	Jun 2026	Zero-ticket smart card	Not yet treated
Kerala	Priyadarshini	Jun 2026	Selected KSRTC ordinary services	Not yet treated

Note. Compiled by the author from state government orders, transport corporation notifications and contemporaneous press reporting. Dates should be verified against the operative government order before publication, particularly for the two schemes introduced in 2026. All remaining states and union territories are never-treated within the estimation window.

4.3. Data and Sample

The analysis uses eight consecutive annual rounds of the Periodic Labour Force Survey from 2017–18 to 2024–25 (Ministry of Statistics and Programme Implementation, various years). The survey is nationally representative, is designed to yield state-level estimates, and applies a consistent questionnaire and sampling design across the rounds used, which is essential for a design that compares states over time. The estimation sample comprises women aged 15 to 59 in the usual-status schedule, excluding full-time students, yielding a pooled sample of approximately 1.05 million woman-year observations across twenty-nine states and union territories. A parallel sample of men in the same age range is constructed for the triple-difference and placebo specifications.

The primary outcome is labour force participation on the usual status definition, including both principal and subsidiary activity, coded as unity for a woman recorded as working or as available for and seeking work. Secondary outcomes capture the margin of adjustment: whether the woman is employed at all, whether her employment is regular salaried, whether it is non-farm, and whether her workplace lies outside her village or town of residence. Individual controls comprise age and its square, years of schooling, marital status, social group, household size, and an indicator for urban residence. Table 2 defines the variables. All estimates are weighted using the survey multipliers, and standard errors are clustered at the state level, which is the level at which treatment is assigned (Abadie et al., 2023; Bertrand et al., 2004). Because twenty-nine clusters is a modest number, inference is also reported using the wild cluster bootstrap of Cameron, Gelbach and Miller (2008).

Table 2. Variable Definitions

Variable	Notation	Definition	Source
Participation	LFP	= 1 if in the labour force, usual status, principal or subsidiary	PLFS
Employment	EMP	= 1 if employed, usual status	PLFS
Salaried work	REG	= 1 if in regular salaried or wage employment	PLFS
Non-farm work	NFARM	= 1 if employed outside agriculture	PLFS
Work away	AWAY	= 1 if workplace is outside the village or town of residence	PLFS
Treatment	D	= 1 from the first survey year in which the scheme was operative	Table 1
Cohort	G	First survey year of treatment for the state; 0 if never treated	Table 1
Controls	X	Age, age squared, schooling, marital status, social group, household size, urban	PLFS

Note. PLFS denotes the Periodic Labour Force Survey, annual rounds 2017–18 to 2024–25. Sample restricted to women aged 15 to 59 excluding full-time students; a parallel male sample is used for the triple-difference and placebo specifications. Author's compilation.

4.4. Identification and Estimation

The conventional specification for a staggered rollout regresses the outcome on a treatment indicator together with unit and time fixed effects:

$$Y_{ist} = \beta\, D_{st} + \gamma' X_{ist} + \alpha_s + \lambda_t + \varepsilon_{ist} \qquad (2)$$

where i indexes individuals, s states and t survey years. Equation (2) is reported for comparability, but it is not the preferred specification. When adoption is staggered and treatment effects differ across cohorts or evolve over time, the coefficient β is a weighted average of many two-group comparisons, some of which use already-treated units as controls for later-treated units; those comparisons enter with weights that can be negative, and the resulting estimand need not lie within the range of the underlying effects (Goodman-Bacon, 2021; de Chaisemartin & D'Haultfœuille, 2020). The Goodman-Bacon decomposition is therefore reported so that the weight placed on the problematic comparisons can be inspected directly.

The preferred estimator is that of Callaway and Sant'Anna (2021), which estimates group-time average treatment effects ATT(g,t) for each adoption cohort g and period t using only never-treated and not-yet-treated units as controls, and then aggregates them with weights that are transparent and non-negative. The doubly robust version is used, so that consistency requires either the outcome regression or the propensity score to be correctly specified. Two further estimators are reported as checks on the aggregation scheme: the interaction-weighted estimator of Sun and Abraham (2021), which corrects the contamination of event-study coefficients by effects from other relative periods, and the imputation estimator of Borusyak, Jaravel and Spiess (2024), which is efficient under homoskedasticity and uses all untreated observations to fit the counterfactual. Where the three agree, the result is unlikely to be an artefact of any single aggregation choice (Roth et al., 2023). Wooldridge (2021) shows that an equivalent estimate can be recovered from a saturated two-way Mundlak regression, and this is reported as a further check.

Dynamics are examined through an event study in which effects are indexed by time relative to adoption, with the period immediately preceding adoption normalised to zero. A triple-difference specification adds men as an internal comparison group within state and year, which absorbs any state-by-year shock to labour demand that affects both sexes and leaves the female-specific component of the policy effect:

$$Y_{igst} = \theta\, (D_{st} \times Female_i) + \delta_{st} + \varphi_s \cdot Female + \psi_t \cdot Female + \gamma' X_{igst} + \varepsilon_{igst} \qquad (3)$$

4.5. Threats to Identification

Four threats are addressed explicitly. The first is differential pre-trends. Rather than relying on a test of whether pre-treatment event-study coefficients are jointly zero, which has low power and whose use as a screening device distorts subsequent inference (Roth et al., 2023), the analysis reports the sensitivity bounds of Rambachan and Roth (2023), which ask how large a post-treatment violation of parallel trends, expressed as a multiple of the largest violation observed before treatment, would be required to overturn the conclusion.

The second is policy bundling. The fare schemes were in several cases announced together with other guarantees, including cash transfers to women, subsidised cooking gas and free electricity up to a threshold. Where a co-timed cash transfer exists, its expected effect on labour supply is negative through the income effect, so that the estimate reported here is a lower bound on the fare effect rather than an upward-biased composite. Specifications that exclude the states with the largest co-timed transfers are reported in Section 5.7.

The third is the pandemic. The 2020–21 round covers a period of severe labour market disruption that affected states differentially and that coincides with the post-treatment window for the Delhi cohort. The main estimates therefore exclude that round in one robustness specification, and the Delhi cohort is dropped in another.

The fourth is migration. If women move to states that abolish fares, the composition of the treated population changes and the estimate confounds selection with behavioural response. Inter-state migration in India is low, and the schemes in Punjab, Karnataka, Telangana and Andhra Pradesh require domicile, which limits the incentive. A specification restricted to women resident in the state since birth is reported as a check.

4.6. Ethical Considerations

The study uses anonymised unit-level records from a publicly released official survey, obtained under the standard terms of access, together with publicly available policy documents. No individual is identifiable and no primary data collection was undertaken.

V. RESULTS AND DISCUSSION

5.1. Descriptive Statistics

Table 3 reports means for the pooled sample and separately for states that eventually adopt a scheme and those that do not. Female labour force participation averages 33.4 per cent across the pooled sample, and 25.6 per cent among eventually-treated states in their pre-treatment years, which is the base against which the estimated effect should be read. Adopting and non-adopting states differ in levels, as expected given that the adopters include both the highest and among the lowest participation states, but the pre-treatment trends plotted in the event study of Section 5.4 are close to parallel. Only about a fifth of working women are in regular salaried employment and fewer than one in five works outside her village or town of residence, which indicates how spatially confined women's work remains and how much room exists for a transport intervention to matter.

Table 3. Descriptive Statistics, Women Aged 15 to 59

Variable	All	Ever treated	Never treated	Diff.
Labour force participation	0.334	0.256	0.361	−0.105
Employed	0.318	0.242	0.344	−0.102
Regular salaried (if working)	0.211	0.264	0.193	0.071
Non-farm (if working)	0.383	0.451	0.359	0.092
Works outside village or town	0.187	0.226	0.173	0.053
Age (years)	34.8	35.2	34.7	0.5
Years of schooling	6.4	7.1	6.2	0.9
Currently married	0.780	0.771	0.783	−0.012
Urban residence	0.331	0.402	0.306	0.096
Observations	1,048,376	271,904	776,472	n.a.

Note. Weighted means from the Periodic Labour Force Survey, annual rounds 2017–18 to 2024–25. The ever-treated column reports pre-treatment years only. Source: Author's calculations.

5.2. The Conventional Estimator and its Decomposition

The two-way fixed effects specification of equation (2) yields a coefficient of 1.42 percentage points, significant at the five per cent level. The Goodman-Bacon (2021) decomposition shows why this figure should not be taken at face value. Comparisons of treated states against never-treated states account for 68.4 per cent of the total weight and yield an average effect of 2.91 percentage points. Comparisons of earlier-treated against later-treated states, in which the later-treated serve as controls before their own adoption, account for a further 23.1 per cent and yield 2.34 percentage points. The remaining 8.5 per cent of the weight falls on comparisons in which already-treated states serve as controls for later adopters, and these yield a strongly negative average of −4.12 percentage points, with 4.1 per cent of the total weight negative. Since the effect grows over time, as Section 5.4 shows, these forbidden comparisons subtract a growing treated outcome from the control side and pull the aggregate down. The conventional estimator therefore understates the effect substantially, and the corrected estimators are used throughout what follows.

5.3. Main Estimates

Table 4 reports the main results. The Callaway and Sant'Anna (2021) aggregate treatment effect on the treated is 2.81 percentage points with a standard error of 0.88, significant at the one per cent level. Against the pre-treatment base of 25.6 per cent among eventually-treated states, this is a relative increase of approximately eleven per cent. The Sun and Abraham (2021) estimate of 2.64 and the Borusyak, Jaravel and Spiess (2024) estimate of 3.02 bracket this figure closely, and the Wooldridge (2021) two-way Mundlak estimate of 2.88 is consistent with all three.

The agreement across estimators that differ in their aggregation and efficiency properties indicates that the result is not an artefact of a particular weighting scheme. H_1 is not rejected.The triple-difference estimate of 2.53 percentage points is close to the difference-in-differences estimates, which is informative in itself: because equation (3) absorbs all state-by-year variation common to both sexes, its agreement with the simpler specification indicates that the estimated effect is not being driven by state-specific labour demand shocks that happen to coincide with adoption. The placebo estimate on male participation is 0.31 percentage points with a standard error of 0.44 and is statistically indistinguishable from zero, consistent with H_4 and with the fact that men remained liable for fares under every scheme.

Table 4. Effect of Fare Abolition on Female Labour Force Participation

Estimator	ATT (pp)	SE	95% CI	p-value
Two-way fixed effects	1.42	0.61	[0.22, 2.62]	0.020
Callaway and Sant'Anna (2021)	2.81	0.88	[1.09, 4.53]	0.001
Sun and Abraham (2021)	2.64	0.91	[0.86, 4.42]	0.004
Borusyak et al. (2024)	3.02	0.79	[1.47, 4.57]	0.000
Wooldridge (2021) Mundlak	2.88	0.85	[1.21, 4.55]	0.001
Triple difference (women vs men)	2.53	0.84	[0.88, 4.18]	0.003
Placebo: male participation	0.31	0.44	[−0.55, 1.17]	0.481

Note. Outcome is labour force participation on the usual status definition, in percentage points. Standard errors clustered at the state level (29 clusters). All specifications include individual controls, state fixed effects and survey-year fixed effects, and are weighted by survey multipliers. Source: Author's estimation.

5.4. Event Study and Dynamics

The event study, estimated by the interaction-weighted method of Sun and Abraham (2021), shows no movement before adoption. Coefficients at four, three and two years before treatment are 0.21, −0.34 and 0.18 percentage points respectively, each small relative to its standard error, and the joint test of the pre-treatment coefficients returns a p-value of 0.61. The effect in the year of adoption is 0.94 percentage points and imprecisely estimated, which is expected since most schemes began part-way through the survey reference year. It reaches 1.93 in the first full year, 2.87 in the second and 3.61 in the third. The gradual accumulation is informative about mechanism. An effect operating purely through the household budget would appear immediately and then flatten. A gradual path is instead consistent with a search process, in which the widening of the feasible commuting radius allows women to find work that did not previously exist within reach, and with the slow adjustment of household arrangements for care that a woman's entry into work requires. It is also consistent with information diffusion, since take-up of the schemes rose over the first two years in the administrative ridership data.

5.5. Heterogeneity and Mechanism

Table 5 reports effects by subgroup and by outcome. The urban effect of 4.12 percentage points is more than twice the rural effect of 1.83, as predicted by H_2: urban bus networks are denser, urban journeys to work are longer, and a larger share of urban women's potential employment lies outside walking distance. Effects are larger among women with at least secondary education, at 3.94 against 1.64 for those with less, which is consistent with the argument of Afridi et al. (2018) that educated women in India face a constrained set of acceptable jobs, since a widening of the geographic radius has most value precisely where the locally available options are least acceptable. The outcome decomposition supports the spatial mechanism of H_3. The probability of working outside the village or town of residence rises by 2.21 percentage points against a base of 22.6 per cent among eventually-treated states, an increase of roughly a tenth. Regular salaried employment rises by 1.72 percentage points and non-farm self-employment by 1.31, while casual labour shows a small negative and statistically insignificant coefficient of −0.22. The composition of the response is therefore towards work that is located away from home and is more formally organised, which is what equation (1) predicts when the binding constraint is the cost of distance rather than the availability of work as such.

Table 5. Heterogeneity and Alternative Outcomes

Subgroup or outcome	ATT (pp)	SE	Base (%)	p-value
By location				
Urban	4.12	1.06	22.1	0.000
Rural	1.83	0.79	28.3	0.021
By education				
Below secondary	1.64	0.81	27.9	0.043
Secondary and above	3.94	1.12	22.8	0.000
By age				
15 to 29	3.47	1.09	19.4	0.001
30 to 44	2.91	0.95	30.1	0.002
45 to 59	1.62	0.88	26.8	0.066
Alternative outcomes				
Works outside village or town	2.21	0.68	22.6	0.001
Regular salaried employment	1.72	0.55	26.4	0.002
Non-farm self-employment	1.31	0.49	18.9	0.008
Casual labour	−0.22	0.41	31.2	0.591

Note. Each row is a separate Callaway and Sant'Anna (2021) estimation. Base is the pre-treatment mean of the outcome among eventually-treated states for the relevant subgroup. Standard errors clustered at the state level. Source: Author's estimation.

5.6. Sensitivity to Parallel Trends

The absence of visible pre-trends is reassuring but not decisive, and the sensitivity analysis of Rambachan and Roth (2023) is reported in preference to a binary test. Under the relative magnitudes restriction, which permits the post-treatment violation of parallel trends to be as large as a multiple M of the largest violation observed in the pre-treatment period, the estimated effect remains significantly positive for values of M up to 1.8. In other words, differential trends between treated and control states would have to be nearly twice as severe after adoption as anything observed before it in order to overturn the conclusion. Under the smoothness restriction, which bounds the extent to which any pre-existing linear trend may change, the confidence set excludes zero for departures of up to 0.6 percentage points per year.

5.7. Robustness

Seven further specifications were estimated. Excluding Delhi, whose status as a city-state makes it unrepresentative of the bus networks elsewhere, yields 2.63 percentage points. Excluding the 2020–21 pandemic round yields 2.94. Excluding the 2025 cohort, which is only partially observed, yields 2.77. Excluding the states with the largest co-timed cash transfers to women yields 2.71, which supports the argument of Section 4.5 that the bundling of schemes attenuates rather than inflates the estimate. Restricting the sample to women resident in the state since birth yields 2.69, indicating that selective migration is not driving the result. Using participation on the principal status definition alone yields a smaller 2.36, as expected given that subsidiary activity is where marginal entrants are most likely to appear. Adding state-specific linear trends yields 2.58. Inference by wild cluster bootstrap returns a p-value of 0.006, confirming that the modest number of clusters is not responsible for the significance of the result.

5.8. Discussion

Three points emerge. The first concerns magnitude. An effect of 2.8 percentage points on a base of 25.6 is substantial for a single instrument, and it is large relative to what most active labour market programmes achieve, but it is not transformative: roughly seven in ten women in the treated states remain outside the labour force after the reform. This is what the framework of Section 3 predicts. Fare abolition removes the monetary component of the cost of travel and leaves the time cost, the safety cost and the normative penalty untouched. The estimate should therefore be read as a lower bound on what a comprehensive relaxation of the mobility constraint could achieve, and as evidence that the constraint is real rather than as evidence that it is the only one.

The second concerns the mechanism. The concentration of the effect in urban areas, among educated women, and in work located outside the village or town points to the spatial margin rather than to a pure income effect. This distinction matters for policy: if the binding constraint were the household budget, an equivalent cash transfer would do as well and would be administratively simpler; because the constraint appears to be the cost of distance, an in-kind subsidy tied to travel does something a cash transfer would not.

The third concerns cost. Karnataka's scheme has involved an annual outlay of the order of four thousand crore rupees. Set against the estimated effect applied to the state's female population aged 15 to 59, this implies a gross fiscal cost per additional labour force participant of roughly seventy-five thousand rupees a year, which appears high. The figure is misleading as a measure of the cost of the employment effect, however, because the overwhelming majority of the outlay is not a resource cost at all but a transfer to women who would have travelled and paid the fare in any case, and whose welfare gain is approximately the fare saved. A complete evaluation would count that transfer as a benefit accruing to a population that is poorer than average rather than as a cost, and would set the resource cost of additional bus capacity, rather than the whole subsidy, against the employment gain. The exercise is beyond the scope of this paper but is a natural extension.

Four caveats apply. First, the estimand averages across schemes of differing generosity, and the effect of a comprehensive scheme may exceed the average reported here. Second, repeated cross-sections identify effects on population aggregates rather than transitions of particular women, so the paper cannot say whether the additional participants are the same women who gained cheaper travel or others responding to a changed local labour market. Third, general equilibrium effects are not identified: if the additional supply depresses women's wages, a partial equilibrium reading of the estimate overstates the welfare gain. Fourth, self-reported participation may respond to the salience of the scheme, although the movement in the location and type of work argues against a purely reporting-driven interpretation.

VI. CONCLUSION AND POLICY IMPLICATIONS

This paper has evaluated the abolition of bus fares for women across seven Indian states and union territories using eight rounds of the Periodic Labour Force Survey and the recent generation of estimators designed for staggered treatment adoption. Four findings emerge. Fare abolition raises female labour force participation by approximately 2.8 percentage points, an increase of about eleven per cent relative to the pre-reform base. The effect is absent before adoption and accumulates over roughly three years. It is concentrated among urban women, women with secondary education, and employment located outside the village or town of residence, which points to the widening of the feasible commuting radius as the operative mechanism. There is no corresponding effect on men, and a triple-difference specification using men as an internal control returns essentially the same estimate as the simpler design.

Four policy implications follow. First, the cost of travel is a genuine and quantitatively meaningful constraint on women's employment in India, and the persistent framing of low participation as a matter of preferences or norms alone is incomplete. Second, because the effect works through distance rather than through the household budget, an in-kind travel subsidy is not equivalent to a cash transfer of the same value, and the two should not be treated as interchangeable in fiscal debate. Third, the concentration of effects in urban areas suggests that the returns to fare policy are greatest where bus networks are dense, and that in thinly served rural areas fare abolition without an expansion of service frequency and route coverage will accomplish comparatively little; the binding constraint there is the bus that does not run rather than the fare that must be paid. Fourth, the untouched components of the mobility constraint, particularly safety in transit and at interchanges, are the natural complements to fare policy, and the evidence from Borker (2021) and Kondylis et al. (2020) suggests the returns to addressing them may be of a similar order.

Three extensions are promising. The most immediate is the adoption of schemes by West Bengal and Kerala in 2026, which falls outside the present estimation window and provides a genuine out-of-sample test: the framework developed here yields a prediction for those states that subsequent survey rounds will confirm or refute, and Kerala is of particular interest because its high female education and relatively high baseline participation place it at a different point on the distribution from the earlier adopters. The second is the use of high-frequency ridership and mobility data to measure the change in the commuting radius directly rather than inferring it from the location of work. The third is the estimation of general equilibrium effects on women's wages in the treated states, which determines whether the participation gain documented here translates into a welfare gain of comparable magnitude.

REFERENCES

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics, 138*(1), 1–35.

Abebe, G., Caria, S., Fafchamps, M., Falco, P., Franklin, S., & Quinn, S. (2021). Anonymity or distance? Job search and labour market exclusion in a growing African city. *The Review of Economic Studies, 88*(3), 1279–1310.

Afridi, F., Dinkelman, T., & Mahajan, K. (2018). Why are fewer married women joining the work force in rural India? A decomposition analysis over two decades. *Journal of Population Economics, 31*(3), 783–818.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics, 119*(1), 249–275.

Black, D. A., Kolesnikova, N., & Taylor, L. J. (2014). Why do so few women work in New York (and so many in Minneapolis)? Labor supply of married women across US cities. *Journal of Urban Economics, 79*, 59–71.

Borker, G. (2021). *Safety first: Perceived risk of street harassment and educational choices of women* (Policy Research Working Paper). World Bank.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *The Review of Economic Studies, 91*(6), 3253–3285.

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics, 90*(3), 414–427.

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996.

Field, E., & Vyborny, K. (2022). *Women's mobility and labor supply: Experimental evidence from Pakistan* (ADB Economics Working Paper Series). Asian Development Bank.

Fletcher, E., Pande, R., & Moore, C. M. T. (2017). *Women and work in India: Descriptive evidence and a review of potential policies* (Faculty Working Paper). Harvard Kennedy School.

Franklin, S. (2018). Location, search costs and youth unemployment: Experimental evidence from transport subsidies. *The Economic Journal*, *128*(614), 2353–2379.

Goldin, C. (1995). The U-shaped female labor force function in economic development and economic history. In T. P. Schultz (Ed.), *Investment in women's human capital and economic development* (pp. 61–90). University of Chicago Press.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics, 225*(2), 254–277.

Jayachandran, S. (2021). Social norms as a barrier to women's employment in developing countries. *IMF Economic Review*, *69*(3), 576–595.

Klasen, S., & Pieters, J. (2015). What explains the stagnation of female labor force participation in urban India? *The World Bank Economic Review, 29*(3), 449–478.

Kondylis, F., Legovini, A., Vyborny, K., Zwager, A. M. T., & Cardoso De Andrade, L. (2020). *Demand for safe spaces: Avoiding harassment and stigma* (Policy Research Working Paper). World Bank.

Martínez, D. F., Mitnik, O. A., Salgado, E., Scholl, L., & Yañez-Pagans, P. (2020). Connecting to economic opportunity: The role of public transport in promoting women's employment in Lima. *Journal of Economics, Race, and Policy, 3*(1), 1–23.

Ministry of Statistics and Programme Implementation. (various years). *Periodic labour force survey: Annual report*. Government of India.

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *The Review of Economic Studies, 90*(5), 2555–2591.

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics, 235*(2), 2218–2244.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics, 225*(2), 175–199.

Wooldridge, J. M. (2021). *Two-way fixed effects, the two-way Mundlak regression, and difference-in-differences estimators* (Working Paper). Michigan State University.

JOURNAL OF ECONOMIC INSIGHTS AND RESEARCH (JEIR)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3107-9482 ISSN Print: 3139-1982
EDUSCHOOL
ACADEMIC RESEARCH & PUBLISHERS

Impact of Artificial Intelligence on the Global Economy

Niharika Sharma Mahajan

Assistant Professor, Department of Economics, Goswami Ganesh Dutta Sanatan Dharma College, Chandigarh, India

Article information
Received: 18th May 2026
Received in revised form: 20th May 2026
Accepted: 17th June 2026
Available online: 25th August 2026

Volume: 2
Issue: 3
DOI: https://doi.org/10.63090/JEIR/3107.9482.0025

Abstract

Artificial Intelligence (AI) has emerged as one of the most influential technological innovations of the twenty-first century, significantly transforming economic systems, industrial structures, labor markets, and global trade mechanisms. The rapid development of machine learning, robotics, natural language processing, predictive analytics, and generative AI technologies has accelerated digital transformation across the world. This research paper critically examines the impact of Artificial Intelligence on the global economy by analyzing its role in economic growth, industrial productivity, employment generation, international trade, innovation, and income inequality. The study is based on secondary data collected from reports of international organizations such as the International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), World Economic Forum (WEF), Stanford AI Index Report, and peer-reviewed research articles published between 2024 and 2026. The study reveals that AI has become a major contributor to economic expansion by increasing efficiency, reducing operational costs, improving decision-making processes, and enabling automation across industries. Advanced economies such as the United States, China, Japan, and European countries are investing heavily in AI infrastructure and research to strengthen their competitive position in global markets. Simultaneously, AI has created challenges related to employment displacement, ethical concerns, cybersecurity risks, and widening economic inequality between technologically advanced and developing nations. The paper also discusses sector-wise impacts of AI in manufacturing, healthcare, banking, agriculture, education, and logistics. The findings suggest that although AI offers significant opportunities for economic development and innovation, its benefits remain unevenly distributed due to disparities in digital infrastructure, investment capacity, and technological preparedness. Therefore, governments and policymakers must formulate inclusive strategies focusing on education, workforce reskilling, ethical governance, cybersecurity, and sustainable digital development. The study concludes that Artificial Intelligence will continue to reshape the future of the global economy and that nations that effectively integrate AI technologies into their economic systems will emerge as leaders in the evolving digital era.

Keywords: - Artificial Intelligence, Global Economy, Automation, Economic Growth, Productivity, Employment, Innovation, Digital Transformation.

I. INTRODUCTION

AI systems enable computers and machines to perform tasks traditionally associated with human intelligence, including learning, reasoning, decision-making, language understanding, and problem-solving. Over the past decade, AI technologies have evolved rapidly and become integrated into numerous sectors of the global economy. Businesses, governments, and institutions worldwide are increasingly utilizing AI systems to improve efficiency, automate operations, and generate innovative solutions to complex economic and social problems.

The global economy has experienced major technological transformations throughout history, including the Industrial Revolution, the rise of computers, and the development of the internet. However, the emergence of AI represents a new phase of digital transformation with far-reaching economic implications. AI technologies are now capable of analyzing vast quantities of data, predicting consumer behavior, optimizing industrial production, and automating tasks previously performed by humans.

The economic significance of AI has increased substantially due to advancements in machine learning and generative AI systems. AI-driven tools are being used in sectors such as healthcare, banking, education, transportation, manufacturing, retail, and agriculture. Governments are investing heavily in AI research and infrastructure to strengthen economic competitiveness and national development strategies.

According to the IMF, AI has the potential to significantly increase global productivity and contribute to long-term economic growth (International Monetary Fund [IMF], 2025b). However, concerns regarding job displacement, ethical governance, cybersecurity, and widening inequality have also emerged as major global challenges associated with AI adoption.

The purpose of this research paper is to analyze the impact of AI on the global economy by examining its effects on economic growth, industrial productivity, labor markets, international trade, and income distribution. The paper further evaluates the opportunities and risks associated with AI-driven economic transformation and suggests policy recommendations for sustainable and inclusive development.

II. OBJECTIVES OF THE STUDY

The major objectives of the present study are as follows:

- To examine the impact of AI on global economic growth.
- To analyze the role of AI in improving industrial productivity and innovation.
- To study the influence of AI on employment and labor markets.
- To evaluate the impact of AI on international trade and economic competitiveness.
- To identify challenges associated with AI-driven economic transformation.
- To suggest policy measures for inclusive and sustainable AI development.

III. RESEARCH METHODOLOGY

The present study is descriptive and analytical in nature. The research is entirely based on secondary data collected from reliable and authentic sources such as international reports, scholarly journals, institutional publications, and research databases.

The major sources used in the study include reports published by the IMF, OECD, WEF, Stanford AI Index Report, and various peer-reviewed academic articles published between 2024 and 2026.

The collected information has been systematically analyzed to understand the economic implications of AI on global industries, employment structures, productivity systems, and international trade. Qualitative analysis has been employed to interpret economic trends and policy developments related to AI technologies.

IV. CONCEPT OF ARTIFICIAL INTELLIGENCE

AI is a branch of computer science that focuses on developing systems capable of simulating human intelligence. AI systems are designed to process information, recognize patterns, learn from experience, and make decisions with minimal human intervention.

The major categories of AI include machine learning, deep learning, natural language processing, robotics, computer vision, expert systems, and generative AI. Machine learning enables systems to improve performance through data analysis, while deep learning uses neural networks to process complex information and identify hidden patterns.

Generative AI has emerged as one of the most transformative forms of AI in recent years. These systems can generate text, images, software code, videos, and analytical content. The increasing accessibility of generative AI tools has accelerated digital innovation and transformed the operational structures of businesses and institutions globally.

AI relies heavily on data, algorithms, and computing power. Organizations with access to large datasets and advanced computing infrastructure possess a significant competitive advantage in AI development and implementation. Consequently, AI has become a strategic priority for governments and multinational corporations worldwide.

V. IMPACT OF AI ON GLOBAL ECONOMIC GROWTH

Recent advances in machine learning have become a major driver of global economic growth by increasing efficiency, reducing operational costs, and improving productivity across industries. AI technologies enable businesses to automate repetitive tasks, optimize workflows, and enhance decision-making capabilitie (Anthropic, 2025).s.

The IMF reported that AI could significantly increase global GDP over the coming decades (Gondauri, 2025; IMF, 2025c). AI-driven automation and predictive analytics allow firms to produce goods and services more efficiently, thereby contributing to higher economic output.

Countries such as the United States and China invest heavily in AI research, semiconductor manufacturing, cloud computing infrastructure, and data systems (Stanford Human-Centered Artificial Intelligence [Stanford HAI], 2025). These investments strengthen their positions as global economic and technological leaders.

AI contributes to economic growth in multiple ways. First, automation reduces dependence on manual labor in repetitive activities, thereby lowering production costs. Second, AI promotes innovation by enabling the development of new products, digital services, and business models. Third, AI improves data analysis and forecasting capabilities, allowing firms to make informed strategic decisions.

The rise of AI-based startups contributes to economic diversification. Sectors such as fintech, health technology, cybersecurity, logistics, and e-commerce are experiencing rapid growth due to AI-driven innovation (Stanford HAI, 2025). As a result, AI has become a central component of modern economic development strategies worldwide.

VI. IMPACT OF AI ON EMPLOYMENT AND LABOR MARKETS

One of the most widely debated aspects of AI is its impact on employment and labor markets. AI-driven automation has transformed the nature of work by replacing routine and repetitive tasks with intelligent systems and machines.

Historically, technological revolutions displaced some jobs while creating new employment opportunities. Similarly, AI reshapes labor markets rather than completely eliminating human employment. AI systems are increasingly being used for tasks such as customer service, bookkeeping, manufacturing operations, and administrative processing.

Low-skilled workers engaged in repetitive occupations face higher risks of job displacement due to automation (Anthropic, 2025).. At the same time, AI has created new employment opportunities in fields such as data science, machine learning, robotics, cloud computing, cybersecurity, and software development.

The growing demand for AI-related skills increases the importance of technical education and workforce reskilling programs. Governments and educational institutions increasingly focus on digital literacy and skill development initiatives.

However, AI may increase income inequality because highly skilled professionals benefit more from technological transformation than low-skilled workers (Cerutti et al., 2025) . The wage gap between technologically skilled employees and routine labor workers is likely to widen further in the coming years.

Developed economies possess stronger educational systems and digital infrastructure, enabling workers to adapt more effectively to technological change. In contrast, developing countries may struggle due to limited technological preparedness and inadequate investment in education and digital systems.

VII. AI AND INDUSTRIAL PRODUCTIVITY

Rapid digitalization has significantly improved industrial productivity by automating production systems, optimizing resource utilization, and minimizing operational inefficiencies (IMF, 2025b). AI-powered technologies are transforming manufacturing industries through smart factories, robotics, and predictive maintenance systems.

Manufacturing companies are increasingly adopting AI-driven automation to improve production speed, reduce human errors, and enhance product quality. Predictive maintenance systems use AI algorithms to identify potential equipment failures before they occur, thereby reducing downtime and operational losses.

In the banking and finance sector, AI is widely used for fraud detection, algorithmic trading, customer relationship management, and risk assessment (Organisation for Economic Co-operation and Development [OECD], 2025). AI-powered chatbots and virtual assistants have improved customer service efficiency while reducing operational expenses.

Healthcare systems worldwide utilize AI technologies for disease diagnosis, medical imaging, drug discovery, and personalized treatment planning. AI applications improve healthcare efficiency and reduce costs associated with manual processes (Stanford HAI, 2025).

Transportation and logistics industries experience substantial productivity gains through AI-enabled route optimization, warehouse automation, and intelligent supply chain management systems.

Overall, AI-driven productivity improvements are enhancing economic efficiency and strengthening global industrial competitiveness.

VIII. AI AND INTERNATIONAL TRADE

AI significantly transforms international trade by improving logistics systems, reducing transaction costs, and enhancing supply chain management. AI-powered technologies allow businesses to forecast demand, optimize inventory management, and improve transportation efficiency (World Economic Forum [WEF], 2025).

Digital commerce platforms utilizing AI expand global market access for businesses of all sizes. AI-based recommendation systems, automated customer support services, and personalized marketing strategies have increased the efficiency of international e-commerce operations.

Countries investing heavily in AI technologies gain competitive advantages in global trade through higher productivity and technological capability (WEF, 2025). AI also increases the importance of digital trade and cross-border data flows in the global economy.

However, AI-driven transformation may widen the technological gap between developed and developing nations (Cerutti et al., 2025). Countries lacking digital infrastructure and technological capabilities may struggle to compete in increasingly automated global markets.

Therefore, international cooperation and technology-sharing initiatives are essential to ensure inclusive participation in the AI-driven global economy.

IX. CHALLENGES ASSOCIATED WITH AI

Despite its numerous economic benefits, AI presents several social, ethical, and economic challenges.

One of the major concerns is job displacement caused by automation. Workers engaged in repetitive and routine occupations are particularly vulnerable to unemployment due to AI adoption.

Ethical concerns related to algorithmic bias, transparency, accountability, and decision-making fairness also remain significant challenges (OECD, 2025). AI systems may produce biased outcomes if trained on inaccurate or discriminatory datasets.

Cybersecurity risks associated with AI technologies are increasing rapidly. AI systems can potentially be exploited for cyberattacks, financial fraud, misinformation campaigns, and data manipulation.

Another major challenge is data privacy. AI systems require large quantities of personal and institutional data, raising concerns regarding surveillance and misuse of sensitive information (Stanford HAI, 2025).

Additionally, AI infrastructure requires substantial computing power and energy consumption, creating environmental and sustainability concerns (IMF, 2025a).

X. POLICY RECOMMENDATIONS

Governments and policymakers must formulate comprehensive strategies to maximize the benefits of AI while minimizing associated risks (Reuters, 2025).
The following policy measures are recommended:

- Investment in digital infrastructure and internet accessibility.
- Promotion of AI education, technical training, and workforce reskilling.
- Development of ethical and transparent AI governance frameworks.
- Strengthening cybersecurity and data protection regulations.
- Encouragement of international cooperation on AI standards and policies.
- Support for AI research and innovation in developing economies.
- Promotion of sustainable energy solutions for AI infrastructure.

XI. CONCLUSION

AI has emerged as one of the most transformative technologies influencing the modern global economy. AI-driven systems are revolutionizing industries, increasing productivity, accelerating innovation, and reshaping labor markets across the world. The study reveals that AI possesses substantial potential to contribute to global economic growth and industrial modernization. However, the benefits of AI remain unevenly distributed due to disparities in technological infrastructure, educational systems, and investment capabilities. AI-driven automation may create employment opportunities in emerging sectors while simultaneously displacing workers engaged in routine occupations. Consequently, governments and institutions must prioritize workforce reskilling, digital literacy, and inclusive technological policies. The research further demonstrates that ethical governance, cybersecurity protection, and sustainable development strategies are essential for ensuring responsible AI implementation.

In conclusion, AI will continue to shape the future trajectory of the global economy and influence patterns of productivity, trade, and labor transformation. While significant progress has been made in understanding its economic implications, further research is needed to examine long-term effects on employment structures, inequality, regulatory frameworks, and cross-country disparities in AI adoption. Future studies may also explore how emerging AI technologies can support sustainable and inclusive economic development.

REFERENCES

Anthropic. (2025). *Anthropic economic index report: Global AI adoption and labor market transformation*.

Cerutti, E., Gopinath, G., & Pizzinelli, C. (2025). *The global impact of AI: Mind the gap* (IMF Working Paper No. 2025/076). International Monetary Fund. https://www.imf.org/en/publications/wp/issues/2025/04/11/the-global-impact-of-ai-mind-the-gap-566129

Gondauri, D. (2025). The impact of artificial intelligence on gross domestic product: A global analysis. *International Journal of Economic Research and Innovation, 14*(3), 45–61.

International Monetary Fund. (2025). *AI needs more abundant power supplies to keep driving economic growth*. https://meetings.imf.org/en/IMF/Home/Blogs/Articles/2025/05/13/ai-needs-more-abundant-power-supplies-to-keep-driving-economic-growth

International Monetary Fund. (2025). *Artificial intelligence and the future of global productivity*. https://www.imf.org/en/topics/artificial-intelligence

International Monetary Fund. (2025). *World economic outlook: Global growth and AI transformation*. https://www.imf.org/en/Publications/WEO/Issues/2025/10/14/world-economic-outlook-october-2025

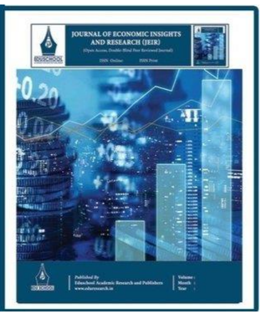
Organisation for Economic Co-operation and Development. (2025). *OECD artificial intelligence policy observatory*. https://oecd.ai

Reuters. (2025, October 13). IMF's Georgieva says countries lack regulatory foundation for AI. *Reuters*. https://www.reuters.com/business/imfs-georgieva-says-countries-lack-regulatory-ethical-foundation-ai-2025-10-13

Stanford Human-Centered Artificial Intelligence. (2025). *AI index report 2025*. Stanford University. https://hai.stanford.edu/ai-index/2025-ai-index-report

World Economic Forum. (2025). *Artificial intelligence and the future of the global economy*. https://www.weforum.org

JOURNAL OF ECONOMIC INSIGHTS AND RESEARCH (JEIR)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3107-9482 ISSN Print: 3139-1982
EDUSCHOOL ACADEMIC RESEARCH & PUBLISHERS



India-Russia Bilateral Trade: An Economic Analysis

Preety Basumatary

Research Scholar, Department of Economics, Gauhati University, Assam, India

Article information
Received: 10th July 2026
Received in revised form: 28th July 2026
Accepted: 1st August 2026
Available online: 25th August 2026

Volume: 2
Issue: 3
DOI: https://doi.org/10.63090/JEIR/3107.9482.0026

Abstract

This paper studies the bilateral trade between India and Russia from two complementary analyses. The descriptive analysis examines recent trade patterns, commodity composition and also addresses trade imbalance using data for 2020-2025, while the econometric analysis investigates the macroeconomic determinants of India-Russia bilateral trade using annual time-series data for 1991-2024. Despite geopolitical challenges and global economic fluctuations, trade between the two nations has grown significantly. This rise in trade is primarily driven by energy imports from Russia to India. However, India faces a persistent trade deficit with Russia. It also discusses the areas in which both countries can improve and strengthen their economic relations. Using annual data, we analyze the effects of India's GDP, trade openness, Brent crude oil prices, and real effective exchange rates (REER) on trade flows. To address heteroskedasticity and autocorrelation, the Newey-West estimator was used. Results indicate that oil prices and India's REER significantly influence trade, while other variables are not statistically significant. The findings underscore the importance of global oil prices and exchange rate movements in shaping the India-Russia trade relationship. This research provides insights for policymakers aiming to strengthen bilateral economic ties.

Keywords: - India, Russia, Bilateral Trade, Trade Deficit, Geopolitical Challenges, Economic Relations.

I. INTRODUCTION

The World Trade Organisation (WTO) defines trade as the exchange of goods and services, either within a country (domestic trade) or across borders (international trade). International trade is the exchange of goods and services across international borders. It allows countries to expand markets for both goods and services that might not have been available domestically. India and Russia have maintained a strong relationship since the collapse of the Soviet Union, mainly due to the Indo-Soviet partnership during the Cold War (Warren & Ganguly, 2022). The USSR assisted India in discovering and exploring energy deposits, particularly oil, gas, and coal, in some states of India (Gujarat, West Bengal, Assam, Tamil Nadu, Jharkhand, and Chhattisgarh) (Kulik, 2023). India and Russia expanded their cooperation in almost every area of the bilateral relationship since October 2000, including political, security, trade and economy, defense, science and technology, and culture. At both political and official levels, some institutionalized mechanisms operate under the strategic partnership to ensure regular interaction between the countries and pursue ongoing cooperative initiatives. In December 2010, the Strategic Partnership was upgraded to the status of 'Special and Privileged Strategic Partnership (Mishra,2018). However, Jim O'Neill (2001) coined the term BRIC to represent a group of emerging and densely populated nations consisting of Brazil, Russia, India, and China when describing their economic development and growth trends. These countries are strengthening their global trade independently and cooperatively as well. They have been characterized as the 'Southern Engines' of global growth (Wani et al., 2020).

India and Russia share a consistent economic and strategic partnership. The bilateral trade between the two economies has historically concentrated on energy, defense, and industrial products. Over the past decade, this trade has evolved in response to global market shifts, sanctions, and domestic policy changes. To date, 22 annual summits have been held alternately between India and Russia. Both countries are enhancing their trade and economic relations and set targets to increase their bilateral trade and investment to $50 billion and $100 billion by 2025 and 2030, respectively. In the financial year 2024-25, bilateral trade between India and Russia reached $68.72 billion. India's exports and imports with Russia amounted to $4.9 billion and $63.8 billion, respectively. Major exporting items from India include pharmaceuticals, organic and inorganic chemicals, iron and steel, and marine products, while major importing items include oil and petroleum products, vegetable oil (particularly sunflower oil), fertilizers, cooking coal, precious stones, and metals. Although the growth in trade between India and Russia has shown an increasing trend over the years, the growth in bilateral trade in services is still lagging despite having the potential to grow. It is widely recognized that, for various reasons, the economic cooperation between Russia and India has not advanced as rapidly as their political relationship (Kulik, 2023). Since June 2022, Russia has surpassed Iraq and Saudi Arabia as India's main crude oil supplier. During the period from January to December 2024, Russia accounted for 36% of India's crude oil imports, while Iraq accounted for 20.5% and Saudi Arabia accounted for 13%. (DGCIS). Before the Russia-Ukraine war, India imported only 2% of its overall crude oil from Russia.

India and Russia have a deep-rooted and multifaceted strategic partnership, rooted in historical connections that date back to the Soviet era and have transformed into a Special and Privileged Strategic Partnership, formally established in 2010. This partnership encompasses defense, energy, trade, technology, and investments, driven largely by mutual economic and geopolitical interests. Over recent decades, particularly since the late 2000s, energy cooperation has become the cornerstone of this bilateral relationship, serving both as a critical driver of economic trade and a platform for broader strategic collaboration (Sethy, 2020; Kulakhmetov, 2024; Anjum, 2025).

India's rising energy demand and Russia's position as a leading exporter of oil, gas, and nuclear technology have naturally aligned their interests. Indian investments in key Russian energy projects, such as Sakhalin-1, Vankor fields, and Arctic LNG initiatives, alongside Russian firms' long-term supply contracts with Indian enterprises (Rosneft, Gazprom, etc.), underscore a deep interdependence in hydrocarbons and nuclear energy (RIAC Gateway House, 2017; Sharma, 2025). Notably, the Kudankulam Nuclear Power Plant exemplifies nuclear cooperation, strengthening this partnership. This energy nexus has propelled a rapid expansion in bilateral trade volume-from approximately USD 1.4 billion in 1995 to a record USD 65.4 billion in 2023-24, with further growth projected (Kulakhmetov, 2024; Anjum, 2025).

The India-Russia trade relationship post-2022 Ukraine conflict and ensuing Western sanctions on Russia has witnessed a pronounced pivot by Russia towards Asian markets, notably India and China. India has emerged as a major importer of discounted Russian crude oil, accounting for a significant share of Russia's lost European market. Alongside hydrocarbons, bilateral trade has diversified to include fertilizers, coal, pharmaceuticals, and defense technology. Infrastructure projects such as the International North-South Transport Corridor (INSTC) and the Chennai-Vladivostok Maritime Corridor are actively enhancing trade logistics, aiming to reduce transit time and costs (Kulakhmetov, 2024; Anjum, 2025; Dolbaia et al., 2025).

Despite robust growth, several challenges persist. The trade balance remains heavily skewed in Russia's favor due to India's large energy imports and limited export diversification. Payment mechanisms leveraging rupee-ruble trade and Vostro accounts mitigate dependence on Western currencies but face operational and regulatory complexities. Non-tariff barriers, infrastructural bottlenecks, and geopolitical risks linked to sanctions continue to hinder full economic integration (PHD Chamber of Commerce, 2018; Kulik, 2023). Moreover, India is pursuing strategic autonomy by diversifying its energy sources and increasing defense indigenization, which may gradually reduce over-reliance on Russia (Dolbaia et al., 2025).

Looking forward, both nations aim to deepen their economic partnership with a target of reaching USD 100 billion trade volume by 2030 (India Brand Equity Foundation). Expanding cooperation into emerging sectors such as renewable energy, digital technologies, biotechnology, and Arctic resource development is anticipated. Enhancing people-to-people ties and regulatory harmonization may further solidify this evolving partnership, which exemplifies pragmatic alignment of economic and strategic interests amid shifting global power dynamics (Kulakhmetov, 2024; Anjum, 2025; Kulik, 2023).

The existing literature shows that most of the studies focus on the strategic and political aspects of India and Russia, with limited attention to recent trade trends and the growing trade imbalance. Therefore, this study aims to contribute by analyzing recent trade patterns in India-Russia bilateral trade and identifying key areas for future cooperation. The understanding of the dynamics of Russia-India bilateral trade is crucial. It provides insights into the changing trade patterns between India and Russia, focusing on the period after the implementation of U.S. sanctions. The purpose of this study is to examine the trade relationship between India and Russia by analysing the current status and sectoral composition of bilateral trade. In addition, the study investigates the key macroeconomic determinants of bilateral trade, including India's GDP, trade openness, Brent crude oil prices, and the real effective exchange rate (REER), using descriptive and econometric analyses. The findings aim to provide insights for strengthening bilateral trade relations and identifying potential areas for future economic cooperation giving policymakers, researchers, and businesses insights about the direction of trade between the two economies and the scope for development in trade to strengthen economic relations between them.

II. METHODOLOGY

This study is mainly based on secondary data. Data are accessed from sources like the Ministry of Commerce and Industry, India Brand Equity Foundation (IBEF), United Nations COMTRADE, Embassy of India, Moscow (Russia), etc. Descriptive statistics like rates, line graphs, and pie charts are used to analyse the key variables such as total exports, imports, and sectoral composition of trade, etc from 2020 to 2024. However, six exogenous variables such as GDP India, Trade openness India, Trade openness Russia, Oil price, REER India, and REER Russia are taken to examine the impact on the dependent variable, bilateral trade spanning from 1991 to 2024. However, data availability varied for certain variables. Trade openness (Russia) was available starting from 1996, while REER for Russia and India were available from 1994. To ensure interpretability and handle large magnitudes, the natural logarithm (log) was applied to bilateral trade, India's GDP, and trade openness for both India and Russia. This allows coefficients to be interpreted as elasticities. Other variables were left in levels. Differencing was applied based on stationarity tests. Due to the presence of multicollinearity when including GDP of both countries, Russia's GDP was dropped to avoid severe multicollinearity. The data sources are as follows: India's GDP and trade openness for both countries were sourced from the World Development Indicators. Oil prices were obtained from the World Bank Commodity Price Data (the Pink Sheet). REER data for both India and Russia were sourced from the IMF. To ensure the validity of time-series analysis, stationarity of all variables was tested using the Augmented Dickey-Fuller test. Given the differing orders of integration, the number of observations was reduced to 26. After addressing heteroskedasticity and autocorrelation, the Newey-West estimator was applied to ensure robust standard errors. Finally, multicollinearity was checked using the Variance Inflation Factor (VIF), and no further issues were found. Thus, the final model is based on these corrected and reliable estimates

III. RESULT AND DISCUSSION

India-Russia trade has shown an increasing trend over the last few years, particularly after the commencement of the Russia-Ukraine war. However, this increase in trade volume is mainly due to a rise in imports from Russia to India, which is the result of discounted crude oil prices. However, India's exports to Russia are not very significant. India needs to increase its exports and diversify its trade partners for the long -term growth of the country. Uncertain global events, such as the U.S. trade war during Trump's tenure, have impacted India, as America remains India's largest export destination. Therefore, to maintain its growth rate, India must build strong trade relationships with multiple countries. Otherwise, India may not be able to achieve its target of becoming one of the world's top three economies by 2030. In the financial year 2025, the bilateral trade between India and Russia amounted to US$ 68.72 billion. Where India's exports amounted to US$ 4.88 billion, and imports from Russia amounted to US$ 63.84 billion (Indian Embassy Moscow).

Now, to understand the current trade status between India and Russia, the following tables are used, which show India's exports to and imports from Russia starting from the financial year 2020-21 to 2024-25 with values in US$ million. Also, these tables show Russia's exports to and imports from India from 2019 to 2023. Here, we used the calendar year except for 2022, which is based on the financial year from 2021 to 2022.

Table1: India's exports to and Imports from Russia, values in US$ million (2020-21 to 2024-25)

Year	Exports to Russia	India's total exports	Imports from Russia	India's total imports
2020-21	2,655.52	291,808.48	5,485.75	394,435.88
2021-22	3,254.68	422,044.40	9,869.99	613,052.05
2022-23	3,146.95	451,070.00	46,212.71	715,968.90
2023-24	4,261.31	437,072.03	61,159.30	678,214.77
2024-25	4,881.30	437,704.58	63,811.44	712,200.22

Source: Ministry of Commerce and Industry, Government of India

From the table, we can clearly illustrate that India's export to Russia has seen a steady but relatively slow increase over the years. It is rising from US $2,655 million in 2020-21 to US $3,254 million in 2021-22. Again, it declined to US$ 3,146 million in 2022-23, then it increased to US$ 4,261million in 2023-24 and further increased to US$ 4,881 million in 2024-25. India's exports to Russia are only about 1% of India's total exports.

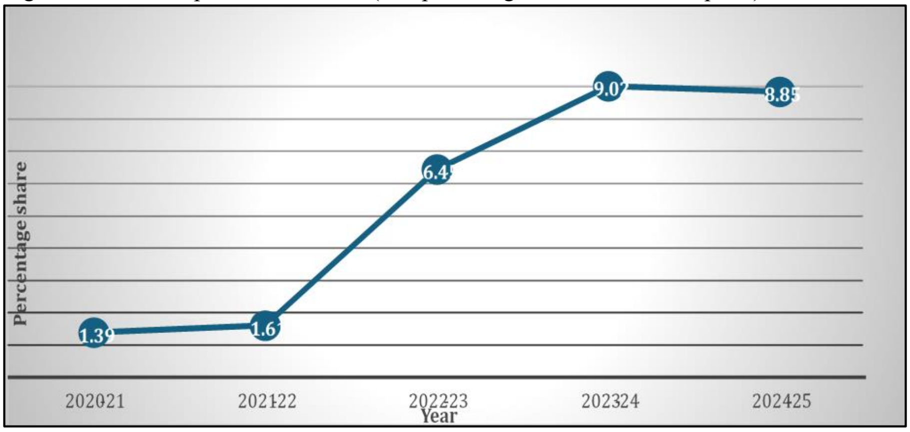
However, in the case of imports, it is relatively stable at 1% for many years. But there is a drastic change in the last three years. Imports increased from $9,869 million in 2021-22 to $46,212 million in 2022-23. In 2023-24, it increased to $61,159 million and further to $63,811 million in 2024-25. This unprecedented change is the result of an increase in the imports of discounted crude oil from Russia. Due to this extensive increase in imports, the total value of trade also increased.

Figure 1: India's exports to Russia (as a percentage of India's total exports)

Source: Ministry of Commerce and Industry, Government of India

Figure 1 depicts India's export share to Russia over the period 2020-21 to 2024-25. Here, we can see that there is a fluctuation in exports during these years. The share of exports to Russia was 0.91% in 2020-21; it declined to 0.77% in 2021-22, and further declined to 0.7% in 2022-23. However, the share of exports to Russia increases after 2022-23. It increased to 0.97% in 2023-24 and further to 1.12% in 2024-25.

Figure 2: India's imports from Russia (as a percentage of India's total imports)

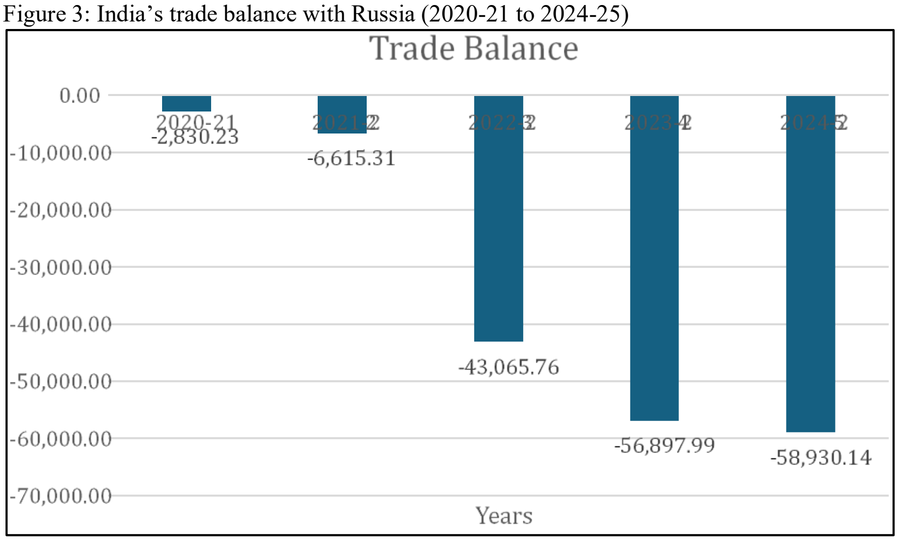


Source: Ministry of Commerce and Industry, Government of India

Figure 2 shows the share of India's imports from Russia for the same period. A significant expansion in import share is observed over the period. It was 1.39% in 2020-21 and 1.61% in 2021-22. A drastic increase in import share was observed during 2022-23, which was 6.45%. Then again, it increased to 9.02% in 2023-24, but a slight decrease was observed in the period 2024-25, which is 8.85%. Though both the share of exports to Russia and imports from Russia have increased after 2022-23, there is a huge difference between them. The import share from Russia is higher compared to the export share of India, which leads to a growing trade imbalance between the two countries. While imports from Russia increased sharply, exports to Russia did not maintain a similar pattern of growth. India has to diversify its export base to Russia and strengthen trade cooperation in sectors other than the oil sector also to correct the trade imbalance. For achieving a balanced trade between the two economies, strengthening bilateral mechanisms, promoting sectoral diversification, addressing barriers which may arise during trade, and finding solutions to these issues are essential.

From the above analyses, it can be stated that India is one of the major export destinations for Russia. India has a negligible role when it comes to import contribution. Its share is merely 1% of Russia's total imports. India has to increase its exports to Russia, focusing on increasing export potential for those goods and services in which India has a comparative advantage. India's pharmaceutical industry has cost-competitive production and a good base of generics. Also, it has the potential to increase exports in machinery and equipment, which are required by Russian industries. In agricultural products and consumer goods, India has the scope for exports. Items like fruits, vegetables, and raw tobacco have been largely exported to Russia in recent years.The export of consumer goods like clothing, footwear, toys, and appliances is rising. As India has agricultural diversity and a cost advantage in some consumable goods, India should focus on diversifying and increasing its exports of those goods. India also exports organic and inorganic chemicals and rubber to Russia. India has to specialize and also standardize its product quality, along with increasing production, to meet the desired standards. Then, only the persistent trade deficit in the country can be reduced. India's export share to Russia is comparatively lower than Russia's export share to India, so India is comparatively less dependent on Russia for its exports. However, India imported a huge amount of merchandise goods from Russia, which indicates that India heavily depends on Russia for its imports. While Russia's export share to India has captured a significant amount, it indicates that Russia's export market is heavily dependent on India. For Russia, India is an export base.

Figure 3: India's trade balance with Russia (2020-21 to 2024-25)



Source: Ministry of Commerce and Industry, Government of India

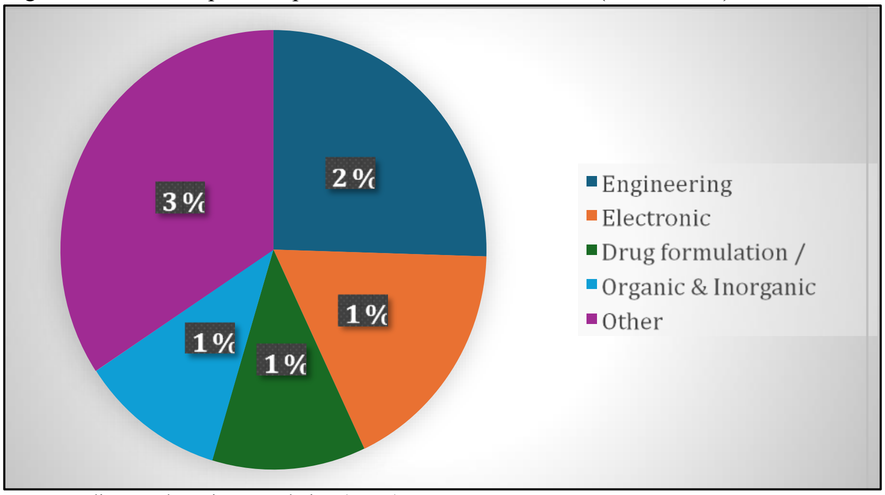
The above figure clearly shows that India is facing a huge amount of trade deficit with Russia. In 2020-21, it has US$-2,830 million trade deficit, and it further aggravated for the next four years, particularly from 2022. There is a sharp rise in the deficit after 2022, which is US$-43,065.76 million. This reflects India's growing import dependence on Russia, particularly in the energy and fertilizer sectors.

3.1. Sectoral Composition of Trade

According to the India Brand Equity Foundation (IBEF), India exported about 3,700 commodities to Russia in the financial year 2025. The major exported items from India to Russia include engineering goods, which amounted to US$ 1.26 billion; electronic goods, which amounted to US$ 862.48 million; drug formulations or pharmaceuticals, which amounted to US$ 577.22 million; organic and inorganic chemical amounted to US$ 545 million; and others, which amounted to US$ 248.40 million during the same year. India imported about 700 commodities from Russia. Major imported items from Russia include petroleum crude, which amounted to US$ 56.8 billion; fertilizers and manufactures amounted to US$ 1.84 billion; animal or vegetable fats and oils, which amounted to US$ 2.39 billion, and pearl, precious, and semi-precious stones, which amounted to US$ 433.93 million during the same financial year.

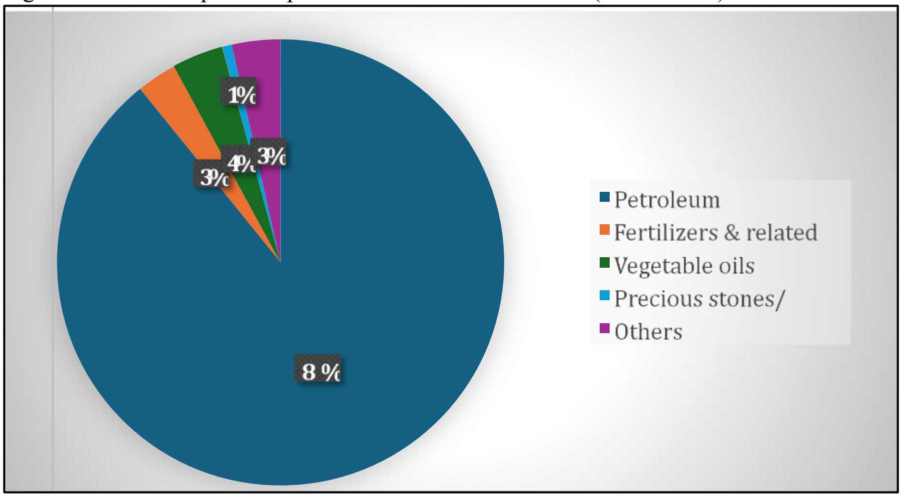
To understand the composition of trade between India and Russia, we take the top five items of exports and imports, which are contributing to bilateral trade. This aims to visualize the relative importance of different product categories in trade composition.

Figure 4: Share of top five export items from India to Russia (FY 2024-25)



Source: India Brand Equity Foundation (IBEF)

Figure 5: Share of top five import items from Russia to India (FY 2024-25)



Source: India Brand Equity Foundation (IBEF)

Figure 4 represents the share of the top five export items from India to Russia. This includes engineering goods, electronic goods, pharmaceuticals, organic and inorganic chemicals, and other export items. It is visible in the figure that engineering goods have the highest share in exports, about 26%. Followed by electronic goods (17%), pharmaceutical products (12%), and organic and inorganic chemicals (11%). Besides these items, there are other items that are also exported by India to Russia; these are machinery, nuclear reactors, boilers, iron and steel, agricultural products, consumer goods, etc. In Figure 5, it can be clearly observed that 89% of total imports consist of petroleum crude, followed by fertilizers and related chemicals, vegetable oils or fats, precious stones & metals, and others. The import share of other products is negligible compared to petroleum crude. After the war, crude oil is the highest imported commodity. India refines this imported crude oil and exports the refined oil to other countries such as the Netherlands, Singapore, the United Arab Emirates, the United States, etc.

This trade pattern between the two countries indicates a serious trade imbalance. India heavily depends on Russia for its energy resources, while the export basket of India remains limitedly diversified. The sectoral distribution shows the energy-dominant bilateral trade. However, the U.S. sanctions on major Russian oil companies, Rosneft and Lukoil, forced India to cut down its oil imports. These two companies supplied about 60% of Russia's oil to India. As a result, several Indian refiners, including both private and state-owned companies, have begun reducing their imports of Russian oil to comply with U.S. restrictions. This development is expected to significantly reduce India's oil imports from Russia and may alter the overall structure of bilateral trade between the two countries (Reuters, 2025). But recently India's oil refining companies have procured approximately 30 million barrels of crude oil from Russia following a waiver granted by the United States. This waiver enabled India to increase its purchases in response to supply disruptions in the Middle East. Major firms such as Indian Oil Corporation and Reliance Industries were involved in these transactions. The development reflects a strategic shift in India's energy sourcing pattern amid ongoing geopolitical tensions and regional conflicts (Sharma & Cheong, 2026).

Despite these opportunities, there are several challenges that hinder the expansion of India-Russia trade. The imposition of Western sanctions on Russia has complicated financial transactions and the management of logistics. Further, the absence of a direct trade route increases transportation costs and shipping times, which disrupts trade relations between Russia and its trade partners. Moreover, limited awareness among Indian exporters about Russian market dynamics, along with language and regulatory impediments, is restricting trade opportunities between the two nations. Exchange rate volatility and payment settlement issues further aggravated these restrictions. Addressing these challenges will require persistent diplomatic engagement, the development of secure financial frameworks, and the improvement of logistical infrastructure. The adoption of these measures will help to form a stable Indo-Russian trade relationship in the upcoming years.

3.2. Statistical Analysis

To understand the trade profile between the two countries, statistical analysis is also important rather than looking at it from a simple descriptive outlook. In this section, first expected signs of the key economic variables on bilateral trade are presented. After testing for stationarity, the Newey-West method was used for robust estimation due to the presence of heteroskedasticity and autocorrelation to see the impact of the explanatory variables. Finally, multicollinearity was checked, and the results are presented.

Table 2: Expected Signs of Independent Variables

Variable	Expected Sign
GDP India	+
Oil Price (Brent)	+
REER India	±
REER Russia	±
Trade Openness (India)	+
Trade Openness (Russia)	+

In the above table, GDP of India is expected to have a positive relationship, as larger economies tend to trade more. They increase their production and import demand, thereby promoting bilateral trade. Oil price is also expected to have a positive sign, given the significance of energy trade. Higher crude oil prices increase the value of India's imports from Russia, leading to higher bilateral trade. While the real effective exchange rate may have a mixed effect, depending on competitiveness. Changes in the real effective exchange rate may improve or reduce international competitiveness. Greater openness facilitates international trade and economic integration.

Following this, the stationarity test of all variables is conducted using the Augmented Dickey-Fuller (ADF) test. A total of seven variables are taken, where bilateral trade is the dependent variable, and all other variables are independent variables. It was found that only Trade Openness Russia was stationary at level (I (0)), and Trade Openness India, oil price, REER India, and REER Russia were stationary at first difference (I (1)), whereas bilateral trade was stationary at second difference (I (2)). Since, one variable was found to be integrated of order two, we could run the Autoregressive Distributed Lage (ARDL) Model. Therefore, a multiple linear regression model was initially estimated. But, after conducting diagnostic tests it was found that there was presence of heteroscedasticity and autocorrelation. To address these issues and obtain consistent standard errors for statistical inference, the Newey-West regression estimator was used. The Newey-West regression estimator provides heteroskedasticity and autocorrelation consistent standard errors, thereby improving the reliability of coefficient significance.

Regression with Newey-West standard errors:

Number of observation = 26

Maximum lag = 2

F (5, 19) = 3.92

Prob >F = 0.0131

Table 3: Newey-West Regression Results

d2lnbilateraltrade	Coefficient	std. err.	t	p>|t|	95% conf. interval
Dlngdpindia	-.837462	1.360707	-0.62	0.546	-3.685454 2.01053
Doilpricebrent	.0092655	.0036913	2.51	0.021	.0015395 .0169915
Dreerindia	0.484596	.0159374	3.04	0.007	.0151022 .081817
dreerrussia	.0116658	.0083335	-1.40	0.178	-.029108 .0057763
dtradeopennesindia	2.87e-13	1.25e-12	0.23	0.822	-2.34e-12 2.91e-12
Lntradeopennessrusia	.0018552	.0043951	0.42	0.678	-.0073438 .0110543
cons	.0108199	.1001685	0.11	0.915	-.1988352. .220475

The overall model is statistically significant as indicated by the F-statistic of 3.92. However, it is found that among the explanatory variables, only oil price and REER (India) are statistically significant. The coefficient of Brent crude oil price is 0.0093 and significant at the 5% level, indicating that an increase in international crude oil price is associated with an increase in bilateral trade between India and Russia. Similarly, India's REER has a positive coefficient of 0.0485 and is statistically significant at the 1% level. This suggests that changes in India's real effective exchange rate significantly influence bilateral trade, highlighting the importance of exchange-rate movements. It helps determine trade flows between the two countries. The significance of these variables suggests that oil prices and exchange rate fluctuations play an important role in the bilateral trade relationship

Table 4: Variance Inflation Factor (VIF) Results

Variable	VIF	1/VIF
Doilpriceb~t	1.61	0.619571
Dreerrussia	1.60	0.625628
Dlngdpindia	1.05	0.948831
Dreerindia	1.05	0.952547
Dtradeopen~a	1.03	0.968003
Mean vif	1.27	

Lastly, multicollinearity was checked by using the Variance Inflation Factor (VIF). The results indicated no multicollinearity issues, ensuring that the model's estimates are reliable. Thus, these results form the basis for the interpretation of the determinants of India-Russia bilateral trade.

IV. CONCLUSION

The analysis of India-Russia trade trends over the past five years reveals a clear shift in the structure and intensity of their bilateral trade. While India's exports to Russia have grown moderately, Russia's share in India's total imports has escalated, largely due to increased crude oil and energy imports since 2022. This trend highlights the evolving nature of India-Russia economic relations, where energy trade has become the dominant driver. However, the relatively low share of India's exports to Russia indicates the need for diversification and greater market access for Indian products. India should focus on expanding its exports where it has a competitive advantage, such as in engineering, electronics, and pharmaceutical products, organic and inorganic chemicals, agricultural products, and information technology. The strengthening of transportation and logistics networks, improving payment mechanisms, and expanding trade in non-energy sectors could help achieve a more balanced and sustainable partnership in the coming years. It should adopt a local currency trade mechanism, such as a rupee-ruble payment settlement system. Adoption of a local currency settlement system can facilitate smoother transactions and can reduce vulnerability to the U.S. dollar. Also, improving transport connectivity through the international North-South Transport Corridor (INSTC) and undertaking strategic alliances in energy, fertilizers, agriculture, and advanced technology sectors can significantly strengthen economic ties between the two countries. Joint ventures in manufacturing and green technologies can promote sustainable growth. Enhancing institutional coordination and MSME participation will make trade more inclusive. In the long term, a comprehensive Economic Partnership Agreement (CEPA) could deepen and stabilize bilateral trade relations.

Due to lack of data availability from Russia's government sources resulting from the ongoing Russia-Ukraine war, this study is therefore conducted primarily from India's export-import perspective rather than from the Russian side. Consequently, the focus of the analysis is on data accessible from the Indian viewpoint. Future studies, however, can certainly be conducted from the Russian side once the data becomes more readily available.

REFERENCES

Anjum, A. (2025, January–February). India–Russia trade and economic partnership: A comprehensive analysis. *International Journal for Multidisciplinary Research, 7*(1). https://doi.org/10.36948/ijfmr.2025.v07i01.35952

Directorate General of Commercial Intelligence and Statistics. (n.d.). *India's foreign trade statistics*. Government of India. https://www.dgciskol.gov.in

Dolbaia, T., Banerjee, V., & Southfield, A. (2025, August 22). *Guns and oil: Continuity and change in Russia–India relations*. Center for Strategic & International Studies. https://www.csis.org/analysis/guns-and-oil-continuity-and-change-russia-india-relations

India Brand Equity Foundation. (n.d.). *India-Russia trade*. https://www.ibef.org/indian-exports/india-russia-trade

Indian Embassy, Moscow, & Department of Commerce, Government of India. (2022). *India–Russia trade data (FY 2022).* https://commerce.gov.in

Kulakhmetov, T. (2024, March). *Russia's turn to the East: Energy exports to China and India* (P. J. Saunders, Ed.; Russia Energy Series Working Paper No. 7). Energy Innovation Reform Project. https://innovationreform.org/wp-content/uploads/2024/03/2024-02-Russia-energy-china-and-india.pdf

Kulik, L. (2023). Russia–India economic cooperation: Current trends and promising directions. *MGIMO Review of International Relations, 16*(2), 159–175. https://doi.org/10.24833/2071-8160-2023-2-89-159-175

Macrotrends. (2024). *India–Russia trade and economic indicators*. https://www.macrotrends.net

Ministry of Commerce and Industry, Government of India, Department of Commerce. (n.d.). *Export–import data bank*. https://tradestat.commerce.gov.in/eidb/countrywise

Mishra, A. (2018). India–Russia bilateral trade relations. *International Journal of Creative Research Thoughts, 6*(1), 134–141. https://www.ijcrt.org/papers/IJCRT1802449.pdf

Reuters. (2025, October 23). *Indian refiners review Russian oil contracts after U.S. sanctions*. Reuters. https://www.reuters.com/business/energy/indian-refiners-review-russian-oil-contracts-after-us-sanctions-source-says-2025-10-23/

Sethy, B. (2020). India–Russia energy cooperation: Blooming ties beyond the horizon. *International Journal of Advanced Research, 8*(11), 705–709. https://doi.org/10.21474/IJAR01/12059

Sharma, R., & Cheong, S. (2026). *India buys 30 million barrels of Russian oil after US waiver*. https://economictimes.indiatimes.com/industry/energy/oil-gas/india-buys-30-million-barrels-of-russian-oil-after-us-waiver/articleshow/129444471.cms?from=mdr

UN Comtrade. (n.d.). *International trade statistics database*. United Nations. https://data.un.org

UN Comtrade, & World Integrated Trade Solution. (n.d.). *Russian Federation trade indicators*. World Bank. https://wits.worldbank.org

Wani, N. U. H., Dhami, J. K., & Sidana, N. (2020). Indo–Russia trade: An evaluation of symmetry, complementarity, intensity and similarity. *Kardan Journal of Economics and Management Sciences, 3*(1), 14–34. https://doi.org/10.31841/KJEMS.2021.36

Warren, S. A., & Ganguly, S. (2022). India–Russia relations after Ukraine. *Asian Survey, 62*(5), 811–837. https://doi.org/10.1525/as.2022.1799235

EDUSCHOOL
ACADEMIC RESEARCH & PUBLISHERS
JOURNAL OF ECONOMIC INSIGHTS AND RESEARCH (JEIR)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3107-9482 ISSN Print: 3139-1982

Public Health Expenditure, Governance Quality, and Infant Mortality in Indian States: A Dynamic Panel Analysis

K Shivashankar Bhat

Research Supervisor, College of Management and Commerce, Srinivas University, Mangalore, India

Article information

Received: 12th May 2026
Received in revised form: 16th June 2026
Accepted: 20th July 2026
Available online: 25th August 2026

Volume: 2
Issue: 3
DOI: https://doi.org/10.63090/JEIR/3107.9482.0027

Abstract

India has achieved a substantial reduction in infant mortality over the past three decades, yet the pace of decline has been strikingly uneven across states, and public spending on health has remained close to one per cent of gross domestic product against a National Health Policy target of 2.5 per cent. Whether additional public health expenditure translates into improved child survival, and whether the translation depends on the administrative capacity of the spending state, is a question of direct fiscal consequence. Drawing on an unbalanced panel of twenty-one major Indian states over the financial years 1995–96 to 2022–23, this paper estimates the elasticity of the infant mortality rate with respect to real per capita public health expenditure within a dynamic panel framework. Because health outcomes are a slowly adjusting stock and because budget allocations respond to observed health deficits, the specification includes a lagged dependent variable and is estimated by two-step system generalised method of moments (Arellano & Bover, 1995; Blundell & Bond, 1998) with Windmeijer-corrected standard errors, a collapsed instrument matrix, and expenditure treated as endogenous. Results indicate a short-run elasticity of approximately −0.15 and a long-run elasticity of approximately −0.38 at mean governance quality. The effect is strongly conditional on administrative capacity: in the lowest tercile of the constructed governance index the elasticity is small and statistically indistinguishable from zero, while in the highest tercile it approaches −0.25. Elasticities are roughly twice as large in the Empowered Action Group states as elsewhere, and markedly weaker for neonatal than for post-neonatal mortality. Estimates satisfy the Arellano–Bond AR(2) and Hansen over-identification criteria and survive alternative instrument configurations, alternative outcome measures, and the exclusion of the pandemic years. The findings suggest that fiscal expansion alone is an incomplete instrument: the return to the health rupee is governed by the administrative machinery that spends it.

Keywords: - Public Health Expenditure, Infant Mortality, Dynamic Panel Data, System GMM, Governance Quality, India.

I. INTRODUCTION

Health occupies a dual position in development economics. It is constitutive of well-being in its own right, and it is instrumental to the accumulation of human capital and to labour productivity (Sen, 1999; Deaton, 2013). Among the many summary measures of population health, the infant mortality rate has retained a privileged position in applied work because it is sensitive to a wide range of underlying determinants (maternal nutrition, sanitation, immunisation coverage, skilled birth attendance, and the quality of primary and secondary care), and because it is measured with reasonable consistency over long horizons. India's aggregate record on this indicator is one of steady improvement: infant deaths per thousand live births fell from roughly seventy-four in the mid-1990s to the high twenties by the early 2020s. The aggregate, however, conceals an extraordinary dispersion. Kerala's infant mortality rate has been in the single digits for some years, comparable to that of upper-middle-income countries, while several states in central and eastern India continue to record rates four to six times higher. Convergence has occurred, but slowly and incompletely.

The fiscal context is equally striking. Health is constitutionally a subject of the State List, and the states account for the large majority of public health spending in India, with the Union government contributing through centrally sponsored schemes, notably the National Rural Health Mission from 2005 and its successor architecture. Combined public spending on health has nonetheless remained close to one per cent of gross domestic product for most of the study period, well below the 2.5 per cent target articulated in the National Health Policy of 2017 (Government of India, 2017) and below the level typical of comparator economies. The corollary is one of the highest out-of-pocket shares of total health expenditure in the world, with well-documented consequences for catastrophic health spending and impoverishment.

These two observations, unequal outcomes and constrained budgets, motivate the central empirical question of this paper. If a state increases its per capita public health expenditure, by how much does infant mortality fall? The question is less settled than it may appear. A cross-country literature beginning with Filmer and Pritchett (1999) has repeatedly found that variation in public health spending explains remarkably little of the variation in child mortality once income, education, and other covariates are controlled. A subsequent literature has argued that this null result is not a statement about the futility of spending but about the heterogeneity of institutional environments: where governance is weak, budgeted rupees do not become delivered services (Rajkumar & Swaroop, 2008; Farag et al., 2013). This paper takes that conditional hypothesis to the Indian sub-national context, where the wide dispersion in administrative capacity across states offers useful identifying variation and where the policy implication, whether to prioritise the size of the health budget or the machinery that executes it, is immediate.

1.1. Research Problem

Four gaps in the existing literature motivate the study. First, the majority of Indian sub-national studies specify static models in which current mortality depends only on current inputs. This is difficult to reconcile with the underlying epidemiology: the infant mortality rate is a manifestation of a slowly accumulating stock of maternal health, facility infrastructure, trained personnel, and household practice, and its adjustment to a change in inputs is partial within any single year. Omitting the dynamics both misstates the speed of adjustment and, if the omitted lag is correlated with the regressors, biases the estimated elasticities.

Second, public health expenditure is not plausibly exogenous. Legislatures and finance departments observe mortality outcomes and respond to them; conditional grants under national missions have been explicitly weighted towards high-burden states. Reverse causality of this kind biases the estimated elasticity towards zero and may explain part of the weak association reported in the earlier literature. Studies that do not confront the endogeneity of the budget with credible instruments are therefore difficult to interpret causally.

Third, the conditional-effectiveness hypothesis has been tested extensively across countries but rarely within India, despite the fact that Indian states differ from one another in administrative capacity by margins comparable to those separating countries. Where governance has entered Indian health econometrics, it has typically appeared as a descriptive contrast between well-performing and poorly performing states rather than as an interaction term inside a single estimating equation.

Fourth, the available panel evidence for India largely predates the period in which the most consequential recent policy changes occurred: the maturing of the National Health Mission, the sanitation campaign of the late 2010s, the introduction of publicly financed hospital insurance, and the disruption of routine services during the COVID-19 pandemic. Extending the panel through 2022–23 is therefore of substantive and not merely arithmetical interest.

1.2. Research Objectives

The study pursues four objectives:

- To estimate the short-run and long-run elasticity of the infant mortality rate with respect to real per capita public health expenditure across Indian states over 1995–2023, allowing for partial adjustment of the health stock.
- To test whether the effectiveness of public health expenditure is conditional on the administrative and governance capacity of the spending state.
- To assess heterogeneity in the expenditure elasticity between the Empowered Action Group states and the remaining major states, and between neonatal and post-neonatal mortality.
- To establish the robustness of the estimates to the endogeneity of budget allocations, to instrument proliferation, and to the anomalous pandemic years.

1.3. Research Hypotheses

Four hypotheses are tested:

- H_1: The infant mortality rate exhibits substantial persistence, with an autoregressive coefficient bounded strictly between zero and one, so that the long-run elasticity of expenditure exceeds the short-run elasticity.
- H_2: Real per capita public health expenditure exerts a negative and statistically significant effect on the infant mortality rate once its endogeneity is instrumented.
- H_3: The magnitude of that effect is increasing in governance quality, so that the interaction between expenditure and the governance index is negative in the mortality equation.
- H_4: The expenditure elasticity is larger in the Empowered Action Group states, where baseline mortality is high and the marginal cost of averting an infant death is correspondingly lower.

1.4. Significance and Organisation

The contribution of the paper is fourfold. First, it provides dynamic panel estimates of the health-expenditure elasticity for Indian states using a panel that extends to 2022–23, encompassing the entire National Health Mission period and the pandemic disruption. Second, it treats the budget as endogenous and identifies the elasticity through internal instruments in a system generalised method of moments framework, rather than assuming allocations to be conditionally random. Third, it operationalises governance at the sub-national level through a composite index of administrative capacity and embeds it as an interaction inside the estimating equation, permitting a direct test of the conditional-effectiveness hypothesis for India. Fourth, by decomposing the outcome into neonatal and post-neonatal components it offers indirect evidence on which margin of the health system responds to marginal expenditure. The remainder of the paper is organised as follows. Section 2 reviews the literature. Section 3 develops the theoretical framework. Section 4 describes the data and the econometric strategy. Section 5 presents and discusses the results. Section 6 concludes.

II. LITERATURE REVIEW

2.1. Health Production Functions

The analytical foundation of the empirical literature is the health production function, in which health status is an output produced from medical inputs, environmental conditions, and household behaviour. Auster, Leveson and Sarachek (1969) provided the first systematic estimation of such a function across American states, reporting that medical inputs mattered considerably less than education and lifestyle. Grossman (1972) supplied the microeconomic foundation by modelling health as a durable capital stock that depreciates with age and is augmented by investment, an insight that motivates the dynamic specification adopted here. At the macroeconomic level, Preston (1975) documented the concave relation between national income and life expectancy and, crucially, the outward shift of that relation over time, implying that factors other than income, namely the diffusion of medical knowledge and public health infrastructure, accounted for much of the twentieth-century mortality decline. Caldwell (1979) identified maternal education as a decisive determinant of child survival, a finding replicated in almost every subsequent study and one that the present specification accommodates directly. Cutler, Deaton and Lleras-Muney (2006) synthesise this tradition, concluding that the historical association between income and mortality is mediated substantially by the collective provision of public health goods. Anand and Ravallion (1993) make the corresponding point for cross-country data: private income raises health outcomes largely through its effect on poverty reduction and on public spending on health, not independently of them.

2.2. The Expenditure-Effectiveness Debate

The most influential sceptical contribution is that of Filmer and Pritchett (1999), who found that variation in public health spending explained a negligible share of cross-country variation in child mortality once income, female education, ethnic fragmentation, and cultural variables were included. Their interpretation was not that health care is ineffective but that public spending is frequently allocated to interventions of low marginal value and delivered through systems of low technical efficiency. Gupta, Verhoeven and Tiongson (2002) offered a partial rejoinder, showing that public spending on health matters considerably more for the poor, whose access to private substitutes is limited. Bokhari, Gai and Gottret (2007), instrumenting government health expenditure with variables capturing the fiscal capacity of the state, recovered elasticities of under-five and maternal mortality that were both statistically significant and economically substantial, suggesting that the earlier nulls owed much to simultaneity bias.

The most directly relevant strand is the conditional one. Rajkumar and Swaroop (2008) demonstrated that the effect of public health spending on under-five mortality depends on the quality of governance: in countries with strong bureaucratic quality and low corruption the elasticity is large and precisely estimated, whereas in poorly governed countries it is indistinguishable from zero. Farag et al. (2013) confirmed this pattern using alternative governance measures and a larger sample. Moreno-Serra and Smith (2012) showed that expansions in publicly financed coverage, as distinct from expenditure per se, generate measurable mortality reductions. The quality dimension has been sharpened by Kruk et al. (2018), who argue that a substantial share of amenable mortality in low- and middle-income countries arises not from lack of access but from poor-quality care within facilities that patients do reach, an argument that bears directly on the interpretation of neonatal mortality below.

2.3. Indian Evidence

Indian studies have converged on a modest but genuine effect. Bhalotra (2007) combined state expenditure data with household survey microdata and found that public health spending reduces infant mortality, with the effect concentrated among rural and poorer households and operating substantially through the funding of front-line rather than tertiary services. Farahani, Subramanian and Canning (2010) estimated that a ten per cent increase in state public health spending is associated with a decline of roughly two per cent in the mortality probability, with effects detectable across age groups rather than confined to infancy. Barenberg, Basu and Soylu (2017), using a long state panel, reported a robust negative association between public health expenditure and infant mortality, and found the effect to be larger for rural and for female infants, evidence that public provisioning matters most where private substitutes are least available.

Alongside this, a literature on service delivery documents the conversion problem directly. Banerjee, Deaton and Duflo (2004), auditing public health facilities in rural Rajasthan, found extremely high rates of provider absence and correspondingly low utilisation of nominally free public services, notwithstanding the budgetary provision behind them. Drèze and Sen (2013) attribute the divergence between Kerala and Tamil Nadu on the one hand and the large northern states on the other principally to differences in the effectiveness of public action rather than to differences in per capita income. These observations supply the microeconomic content of the governance interaction estimated here.

2.4. Research Gap

Three gaps follow from this survey. First, the dynamic structure implied by Grossman's stock formulation is rarely imposed in the Indian sub-national literature, and the estimators used are consequently vulnerable to omitted-dynamics bias. Second, the conditional-effectiveness result of Rajkumar and Swaroop (2008) has not been systematically tested across Indian states within a unified econometric framework, despite the availability of considerable variation in administrative capacity.

Third, no published state panel of which the author is aware extends through the pandemic period, leaving the recent policy era unexamined. The present paper addresses each in turn.

III. THEORETICAL FRAMEWORK

3.1. A Stock-Adjustment Health Production Function

Let the desired, or steady-state, level of infant survival in state i and year t be produced according to a Cobb–Douglas technology in the principal determinants identified above:

$$H^*_{it} = A_{it} \cdot PHE_{it}^{\beta_1} \cdot Y_{it}^{\beta_2} \cdot E_{it}^{\beta_3} \cdot S_{it}^{\beta_4} \tag{1}$$

where PHE denotes real public health expenditure per capita, Y real income per capita, E the stock of female education, S environmental sanitation, and A a state-and-time-specific efficiency term capturing medical technology, disease environment, and cultural practice. Because the health stock adjusts only partially within a period (facilities take time to build and staff, cohorts of mothers acquire health capital slowly, and household practices change with a lag), the observed outcome follows a partial-adjustment rule:

$$\ln H_{it} - \ln H_{i,t-1} = \varphi\ (\ln H^*_{it} - \ln H_{i,t-1}), \quad 0 < \varphi \le 1 \tag{2}$$

Substituting (1) into (2) and rearranging yields an estimating equation in which the lagged outcome appears on the right-hand side with coefficient $\delta = 1 - \varphi$, and in which each structural elasticity β is scaled by φ. The long-run elasticity is recovered as the short-run coefficient divided by $(1 - \delta)$. A value of δ close to unity implies a health system in which the effects of any given year's spending are realised only over a long horizon; a value close to zero implies immediate adjustment. Testing H_1 amounts to testing whether δ lies strictly inside the unit interval.

3.2. Governance as a Conversion Parameter

The sceptical results of the cross-country literature are naturally represented by distinguishing budgeted expenditure from effective expenditure. Let a proportion θ of each rupee allocated be converted into delivered service, with the remainder dissipated through leakage in procurement, provider absence, unfilled sanctioned posts, delayed fund release, and the misallocation of expenditure towards low-value uses. Conversion efficiency is taken to be increasing in governance quality G, so that effective expenditure is $\theta(G) \cdot PHE$ with $\theta'(G) > 0$. Taking logarithms, the effective input enters as $\ln \theta(G) + \ln PHE$; a first-order approximation of $\ln \theta(G)$ around the sample mean of G generates an interaction term between ln PHE and G in the estimating equation. The sign of this interaction is the object of H_3: in an equation for log infant mortality it should be negative, indicating that the same proportional increase in spending buys a larger mortality reduction where the administrative machinery is stronger.

It is worth being explicit about what this formulation does and does not claim. It does not claim that governance is exogenous to health outcomes over long horizons, nor that the index used below captures conversion efficiency without error. It claims only that, conditional on state and year effects, cross-state and within-state variation in measured administrative capacity is informative about the marginal productivity of the health rupee. Section 5.7 examines the sensitivity of the interaction to alternative constructions of the index.

3.3. Testable Implications

Three implications follow and are examined in Section 5. First, the autoregressive coefficient should be positive, significant, and bounded above by unity, and the estimated δ should lie between the pooled ordinary least squares estimate, which is biased upward by the correlation between the lagged outcome and the state effect, and the within-groups estimate, which is biased downward by the Nickell (1981) transformation, which is the bracketing check recommended by Bond (2002). Second, the expenditure coefficient should be negative, and larger in absolute value once endogeneity is instrumented than under ordinary least squares, since simultaneity biases it towards zero. Third, the marginal effect of expenditure evaluated across the observed governance distribution should be monotonically increasing in absolute value, and may be statistically indistinguishable from zero at the lower end of that distribution.

IV. RESEARCH METHODOLOGY

4.1. Research Design and Approach

The study adopts a quantitative, deductive, panel-data design (Wooldridge, 2010). The cross-sectional unit is the major Indian state and the temporal unit is the financial year running from April to March. The design is dictated by the research question: identifying the effect of a state's own spending decisions on its own mortality trajectory requires within-state variation over time, while the governance hypothesis requires that this variation be observed across states of differing administrative capacity. A purely cross-sectional design would confound expenditure with the many time-invariant characteristics (disease environment, topography, historical patterns of public action) that determine both.

4.2. Data Sources and Sample

The analysis uses an unbalanced panel of twenty-one major Indian 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, and West Bengal) for the financial years 1995–96 to 2022–23. States created by the reorganisations of 2000 and 2014 enter the panel from their year of formation, which together with occasional gaps in the reporting of state-level mortality produces the unbalanced structure.

Infant, neonatal and under-five mortality rates are taken from the Sample Registration System bulletins and statistical reports (Registrar General of India, various years). Public health expenditure is constructed as the sum of revenue and capital expenditure under the heads of medical and public health, family welfare, and water supply and sanitation as reported in the annual study of state budgets (Reserve Bank of India, various years); the series is deflated to constant 2011–12 prices using the state-level gross domestic product deflator and expressed in per capita terms using interpolated projected population. Net state domestic product per capita at constant prices is taken from the national accounts of the Ministry of Statistics and Programme Implementation (various years). Female literacy is interpolated between the decennial censuses of 1991, 2001 and 2011 (Census Commissioner of India, various years) and extended using the education modules of the National Sample Survey and the Periodic Labour Force Survey (Ministry of Statistics and Programme Implementation, various years). Urbanisation is interpolated between censuses from the same source. Household access to improved sanitation is constructed from successive rounds of the National Family Health Survey (International Institute for Population Sciences, various years) supplemented by administrative coverage data.

The governance variable is a composite administrative-capacity index constructed by principal-components analysis of three annual state-level indicators: the ratio of actual to budgeted health expenditure, which captures absorptive capacity and the timeliness of fund release; the proportion of sanctioned positions actually filled at primary health centres and community health centres, taken from the annual Rural Health Statistics (Ministry of Health and Family Welfare, various years); and the timeliness with which state accounts are audited and laid before the legislature (Comptroller and Auditor General of India, various years). The first principal component, which accounts for the majority of the common variance, is rescaled to the unit interval. The index is deliberately constructed from inputs and processes rather than from health outcomes, so as to avoid mechanical correlation with the dependent variable. Table 1 summarises the variables and their sources.

Table 1. Variable Definitions and Sources

Variable	Notation	Definition	Source
Infant mortality	lnIMR	Infant deaths per 1,000 live births (natural log)	SRS, Registrar General of India
Health expenditure	lnPHE	Real public health expenditure per capita (₹ at 2011–12 prices, natural log)	RBI, State Finances
Governance	GOV	Composite administrative-capacity index (0–1), first principal component	RBI; Rural Health Statistics; CAG
Income	lnNSDP	Real net state domestic product per capita (₹ at 2011–12 prices, natural log)	MoSPI
Female education	lnFLIT	Female literacy rate, population aged 7 and above (%, natural log)	Census; NSS; PLFS
Sanitation	SAN	Households with access to improved sanitation (%)	NFHS; administrative data
Urbanisation	URB	Share of population resident in urban areas (%)	Census (interpolated)

Note. Panel of 21 states × 28 financial years, unbalanced. Monetary variables deflated using state GSDP deflators (base: 2011–12). Author's compilation.

4.3. Empirical Specification

The estimating equation follows directly from Section 3 and takes the dynamic form:

$$ln\ IMR_{it} = \delta\ ln\ IMR_{i,t-1} + \beta_1\ ln\ PHE_{it} + \beta_2\ (ln\ PHE_{it} \times \tilde{G}_{it}) + \beta_3\ G_{it} + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (3)$$

where $\tilde{G}$ denotes the governance index expressed as a deviation from its sample mean, so that β_1 is interpretable as the expenditure elasticity evaluated at average administrative capacity; X is a vector of controls comprising log real income per capita, log female literacy, sanitation access, and urbanisation; μ denotes unobserved state heterogeneity; and λ denotes year effects absorbing national shocks such as the roll-out of centrally sponsored schemes, national immunisation drives, and the pandemic. The marginal effect of expenditure at any level of governance is $\beta_1 + \beta_2\ \tilde{G}$, and the corresponding long-run effect is that quantity divided by $(1 - \delta)$.

4.4. Estimation Procedure and Diagnostic Strategy

Equation (3) cannot be estimated consistently by ordinary least squares, since the lagged dependent variable is correlated with the state effect, nor by within-groups, since the demeaning transformation induces the Nickell (1981) bias, which is of order 1/T and non-negligible in a panel of this length. The estimation therefore proceeds in the following sequence.

First, pooled ordinary least squares and fixed-effects estimates are reported to establish the bracket within which a consistent estimate of δ should lie (Bond, 2002). Second, the Arellano and Bond (1991) difference generalised method of moments estimator is applied, instrumenting the differenced lagged outcome with suitable lags in levels. Third, because infant mortality is highly persistent and lagged levels are consequently weak instruments for differences, the preferred estimator is the system generalised method of moments of Arellano and Bover (1995) and Blundell and Bond (1998), which augments the differenced equations with a levels equation instrumented by lagged differences. Public health expenditure and the governance index are treated as endogenous and instrumented with their own lags; income and female literacy are treated as predetermined; year dummies are treated as strictly exogenous.

Estimation is by the two-step procedure with the Windmeijer (2005) finite-sample correction to the standard errors, without which two-step standard errors are severely downward-biased in panels of this size. Following Roodman (2009), the instrument matrix is collapsed and the lag depth restricted, keeping the instrument count comfortably below the number of cross-sectional units so that the Hansen test retains power. Specification is assessed on three criteria: the Arellano–Bond test should reject the absence of first-order serial correlation in differenced residuals but fail to reject the absence of second-order correlation; the Hansen (1982) test of over-identifying restrictions should fail to reject the joint validity of the instrument set; and the difference-in-Hansen test should fail to reject the additional moment conditions imposed by the levels equation. Cross-sectional dependence is examined using the Pesaran (2021) test and stationarity using the Pesaran (2007) CIPS test.

Robustness is assessed through re-estimation with the neonatal and under-five mortality rates as alternative dependent variables; exclusion of the pandemic years 2020–21 and 2021–22; alternative instrument lag structures; substitution of a single-indicator governance measure for the composite index; exclusion of the reorganised states; and exclusion of Kerala and Tamil Nadu, whose outlying performance might otherwise drive the governance interaction.

4.5. Ethical Considerations

The study uses publicly available aggregate secondary data published by government statistical agencies. No human subjects are involved and no individual-level records are accessed. All sources are acknowledged in the references and in Table 1.

V. EMPIRICAL RESULTS AND DISCUSSION

5.1. Descriptive Statistics

Table 2 reports descriptive statistics for the estimation sample of 542 state-year observations. The infant mortality rate averages 42.6 per thousand live births across the panel, with a standard deviation of 18.9 and a range extending from 6 to 91, a spread that itself testifies to the magnitude of the sub-national divergence noted in the introduction. Real public health expenditure per capita averages ₹612 at 2011–12 prices, but the highest observation is roughly twenty times the lowest, reflecting both differences in fiscal capacity and differences in the priority accorded to the sector. Female literacy averages 62.3 per cent and access to improved sanitation 47.9 per cent, both rising steeply over the period. The governance index has a mean of 0.482 and a standard deviation of 0.163, with substantial within-state as well as between-state variation, an important feature, since identification of the interaction term in a specification with state fixed effects relies on the former.

Table 2. Descriptive Statistics (State-Year Panel, 1995–2023)

Variable	Mean	Std. Dev.	Min	Max
Infant mortality rate (per 1,000)	42.6	18.9	6.0	91.0
Neonatal mortality rate (per 1,000)	29.4	12.1	4.0	60.0
Public health expenditure per capita (₹)	612	401	118	2,340
NSDP per capita (₹'000)	68.4	41.2	14.6	231.7
Female literacy (%)	62.3	14.1	27.4	93.6
Improved sanitation access (%)	47.9	26.3	6.2	99.1
Urbanisation (%)	29.8	12.4	10.1	51.7
Governance index (0–1)	0.482	0.163	0.121	0.874

Note. N = 542 state-year observations (unbalanced panel). Monetary variables at constant 2011–12 prices. Author's calculations from SRS, RBI, MoSPI, Census, NFHS and Rural Health Statistics data.

5.2. Preliminary Diagnostics

The Pesaran (2021) test rejects cross-sectional independence for all series at the one per cent level, which is unsurprising given the common national policy environment and shared macroeconomic shocks; year effects are retained throughout partly on this ground. The Pesaran (2007) CIPS test, which is robust to such dependence, fails to reject a unit root in the level of log infant mortality but rejects decisively in first differences, indicating a highly persistent though not explosive series. This persistence is precisely what motivates the dynamic specification, and it also has an important implication for estimator choice: with a near-unit root, lagged levels are weak instruments for first differences, which is the standard argument for preferring system to difference generalised method of moments in this setting.

The bracketing check of Bond (2002) is reported in the first two columns of Table 3. Pooled ordinary least squares yields an autoregressive coefficient of 0.884, while within-groups yields 0.412. A consistent estimate should lie between these bounds. The system generalised method of moments estimate of 0.617 satisfies this criterion, whereas the difference estimate of 0.548, though also within the bracket, sits closer to the downward-biased fixed-effects value, which is consistent with the weak-instrument concern noted above and reinforces the choice of the system estimator as the preferred specification.

5.3. Baseline Dynamic Estimates

Column (4) of Table 3 reports the preferred baseline. The autoregressive coefficient of 0.617 is significant at the one per cent level and bounded well away from unity, confirming H_1: roughly thirty-eight per cent of the gap between actual and steady-state mortality is closed within a single year, so that the full effect of a permanent change in inputs is realised over approximately five to seven years. The elasticity of infant mortality with respect to real per capita public health expenditure is −0.147 and significant at the one per cent level. A ten per cent increase in per capita health spending is thus associated with a decline in infant mortality of about 1.5 per cent within the year and of approximately 3.8 per cent in the long run once the dynamics have worked through. Evaluated at the panel mean infant mortality rate, the long-run effect corresponds to a reduction of roughly 1.6 infant deaths per thousand live births. H_2 is accordingly not rejected. It is instructive that the instrumented elasticity is substantially larger in absolute value than the pooled ordinary least squares estimate of −0.041. This is the pattern predicted by the simultaneity argument of Section 1.1: because states with poor health outcomes attract larger allocations, both through their own political processes and through the burden-weighted formulae of national missions, the uncorrected association understates the causal effect. Studies that report weak associations without confronting this endogeneity are therefore likely to be reporting attenuated coefficients rather than genuinely small effects.

The controls behave as the literature predicts. The elasticity of infant mortality with respect to female literacy is −0.312, roughly twice the magnitude of the expenditure elasticity, corroborating the Caldwell (1979) tradition and underlining that health spending operates alongside rather than in place of broader human-development investment. Sanitation access enters negatively and significantly, consistent with the enteric-disease pathway to infant death. The income elasticity of −0.186 is smaller than that typically recovered in cross-country work, which is to be expected once state and year effects absorb the slow-moving component of the income–mortality relation. Urbanisation is statistically insignificant, plausibly because the health advantages of urban infrastructure are offset by the conditions of peri-urban informal settlement.

Table 3. Dynamic Panel Estimates (Dependent Variable: lnIMR)

Regressor	(1) OLS	(2) FE	(3) Diff-GMM	(4) Sys-GMM	(5) + Interaction
lnIMR (t−1)	0.884***	0.412***	0.548***	0.617***	0.604***
	(0.021)	(0.048)	(0.072)	(0.058)	(0.061)
lnPHE	−0.041**	−0.198***	−0.161***	−0.147***	−0.152***
	(0.018)	(0.046)	(0.052)	(0.044)	(0.046)
lnPHE × GOV (dev.)	n.a.	n.a.	n.a.	n.a.	−0.479**
					(0.206)
GOV	−0.062	−0.118*	−0.097*	−0.104*	−0.111*
	(0.049)	(0.064)	(0.058)	(0.057)	(0.059)
lnNSDP	−0.058**	−0.224***	−0.201***	−0.186***	−0.181***
	(0.024)	(0.061)	(0.068)	(0.055)	(0.056)
lnFLIT	−0.147***	−0.368***	−0.329***	−0.312***	−0.308***
	(0.041)	(0.092)	(0.104)	(0.086)	(0.088)
SAN	−0.0021***	−0.0058***	−0.0050***	−0.0053***	−0.0051***
	(0.0006)	(0.0014)	(0.0017)	(0.0014)	(0.0014)
URB	−0.0009	−0.0041	−0.0036	−0.0040	−0.0038
	(0.0011)	(0.0033)	(0.0039)	(0.0032)	(0.0033)
State FE	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Long-run elasticity	−0.353	−0.337	−0.356	−0.384	−0.384
AR(1) p-value	n.a.	n.a.	0.002	0.001	0.001
AR(2) p-value	n.a.	n.a.	0.294	0.327	0.341
Hansen J p-value	n.a.	n.a.	0.216	0.290	0.274
Instruments	n.a.	n.a.	24	27	29
Observations	542	542	500	521	521

Note. Standard errors in parentheses; two-step GMM standard errors incorporate the Windmeijer (2005) finite-sample correction. Instrument matrix collapsed. lnPHE and GOV treated as endogenous; lnNSDP and lnFLIT as predetermined. ***, **, * denote significance at 1%, 5% and 10% respectively. Long-run elasticity is the coefficient on lnPHE divided by $(1 - \delta)$, evaluated at mean governance. Source: Author's estimation.

5.4. The Governance Interaction

Column (5) augments the baseline with the interaction between log expenditure and the mean-deviated governance index. The interaction coefficient is −0.479 and significant at the five per cent level, supporting H_3: the same proportional increase in health spending purchases a larger mortality reduction in states with stronger administrative capacity. Because the index is entered as a deviation from its mean, the coefficient on log expenditure in column (5) retains its interpretation as the elasticity at average capacity, and is essentially unchanged from the baseline at −0.152.

Table 4 makes the economic magnitude explicit by evaluating the marginal effect at the tercile midpoints of the governance distribution. In the lowest tercile the short-run elasticity is −0.065 with a standard error large enough that the effect cannot be distinguished from zero at conventional levels; the corresponding long-run elasticity of −0.164 is likewise imprecise. In the highest tercile the short-run elasticity is −0.247 and the long-run elasticity −0.624, both significant at the one per cent level. The point estimate of the marginal product of the health rupee is thus close to four times larger at the top of the administrative-capacity distribution than at the bottom.

This is the Indian sub-national analogue of the cross-country result of Rajkumar and Swaroop (2008), and it reconciles two apparently opposed strands of the literature. The Filmer and Pritchett (1999) null is recovered here as the low-governance case; the substantial elasticities of Bokhari et al. (2007) are recovered as the high-governance case. What appears in pooled estimation as a weak average effect is better understood as an average taken over a distribution in which the effect is large in some places and close to zero in others.

Table 4. Marginal Effect of Public Health Expenditure across the Governance Distribution and by State Group

Group	Mean GOV	Short-run	Long-run	p-value
Governance terciles				
Lowest tercile	0.300	−0.065	−0.164	0.312
Middle tercile	0.482	−0.152	−0.384	0.001
Highest tercile	0.680	−0.247	−0.624	0.000
State groups				
EAG states and Assam	0.391	−0.213	−0.531	0.001
Non-EAG states	0.564	−0.094	−0.238	0.037
Alternative outcomes				
Neonatal mortality	n.a.	−0.083	−0.212	0.048
Under-five mortality	n.a.	−0.161	−0.409	0.000

Note. Marginal effects computed from column (5) of Table 3 as $\beta_1 + \beta_2\bar{G}$ at the stated value of the governance index; long-run effects divide by $(1 - \delta)$. State-group and alternative-outcome rows are from separate re-estimations of the full specification. EAG denotes the Empowered Action Group states (Bihar, Chhattisgarh, Jharkhand, Madhya Pradesh, Odisha, Rajasthan, Uttar Pradesh, Uttarakhand). Source: Author's estimation.

5.5. Heterogeneity across State Groups

Re-estimating the full specification separately for the Empowered Action Group states together with Assam, and for the remaining states, reveals the pattern anticipated in H_4. The short-run elasticity in the high-burden group is −0.213 against −0.094 elsewhere, and the long-run elasticities are −0.531 and −0.238 respectively. The marginal rupee therefore averts substantially more infant deaths where baseline mortality is high, which is what a concave health production function implies and what an allocation rule maximising lives saved would exploit.

This result stands in some tension with the governance finding, and the tension is policy-relevant rather than merely econometric. The states with the highest marginal returns to expenditure are on average also the states with the weakest measured administrative capacity: mean governance in the high-burden group is 0.391 against 0.564 elsewhere. The efficient allocation of a marginal health rupee therefore cannot be read off either result alone. Directing funds towards high-burden states raises the potential return; realising that return requires simultaneous investment in the capacity to spend.

5.6. Neonatal versus Post-Neonatal Mortality

Substituting the neonatal mortality rate for the infant mortality rate yields a markedly smaller elasticity of −0.083, significant only at the five per cent level, whereas the under-five rate yields −0.161. Since neonatal deaths account for a large and rising share of all infant deaths in India, this attenuation is substantively important: it implies that the component of infant mortality most responsive to aggregate public health spending is the post-neonatal component, driven by infectious disease, diarrhoeal illness and undernutrition, rather than the neonatal component, driven by prematurity, low birthweight, intrapartum complications and sepsis.

The natural interpretation is that aggregate expenditure expands the reach of services (immunisation, oral rehydration, outreach visits, referral transport) more readily than it raises the clinical quality of intrapartum and immediate newborn care, which depends on trained personnel, functioning newborn care units and adherence to protocol. This is the quantity-versus-quality distinction emphasised by Kruk et al. (2018), and it suggests that further reductions in Indian infant mortality will depend increasingly on the harder problem of facility quality rather than on the more tractable problem of coverage.

5.7. Diagnostic and Robustness Tests

The specification tests support the estimated models. The Arellano–Bond test rejects the absence of first-order serial correlation in differenced residuals ($p = 0.001$) and fails to reject the absence of second-order correlation ($p = 0.327$), so the moment conditions based on lags of order two and higher are valid. The Hansen test of over-identifying restrictions returns a p-value of 0.290 in the baseline, comfortably within the range that indicates neither rejection of the instrument set nor the implausibly high values that signal instrument proliferation. With the instrument matrix collapsed, the instrument count of twenty-seven remains above but close to the number of cross-sectional units, and the difference-in-Hansen test of the additional levels moment conditions returns a p-value of 0.412, supporting the system rather than the difference estimator.

Six robustness exercises were conducted. Restricting the instrument lag depth to the second and third lags yields an expenditure elasticity of −0.151, statistically indistinguishable from the baseline. Excluding the pandemic years 2020–21 and 2021–22, during which routine service delivery and civil registration were both disrupted, yields −0.139. Replacing the composite governance index with the budget-utilisation ratio alone yields an interaction coefficient of −0.402, still significant at the five per cent level, indicating that the conditional result is not an artefact of the principal-components construction. Excluding the four states affected by the reorganisations of 2000 and 2014 yields estimates within seven per cent of the baseline. Excluding Kerala and Tamil Nadu, which combine unusually low mortality with unusually high measured capacity, attenuates the interaction to −0.391 but leaves it significant, confirming that the governance result is not driven by these two observations alone. Finally, the difference generalised method of moments estimate of −0.161 lies close to the preferred estimate, indicating that the conclusion does not hinge on the additional levels moment conditions.

5.8. Discussion

Three implications follow from the results taken together. The first is that public health expenditure in India does reduce infant mortality, and by an economically meaningful margin, but that the effect is realised slowly and is easily understated. The gap between the pooled elasticity of −0.041 and the instrumented long-run elasticity of −0.384 is almost an order of magnitude, and it arises from two sources that operate in the same direction: the simultaneity of budget allocation with health need, and the omission of the adjustment dynamics. Empirical work that neglects either will tend to conclude, wrongly, that health budgets do not matter much.

The second is that the size of the health budget and the capacity of the administration that executes it are complements rather than substitutes. The near-zero elasticity in the lowest governance tercile is not a statement that spending is wasted in poor states; it is a statement that in the absence of filled posts, timely fund release and functional procurement, additional appropriation does not become additional service. This has an uncomfortable corollary for fiscal design. Conditional grants that are formula-weighted towards high-burden states will, on the evidence here, be directed disproportionately towards states in which the conversion rate is lowest, unless the grants are themselves designed to raise that conversion rate.

The third is that the composition of the remaining mortality burden has shifted in a way that reduces the leverage of aggregate expenditure. When most infant deaths were post-neonatal and infectious in origin, coverage expansion was close to a sufficient instrument. As the neonatal share rises, the binding constraint moves towards clinical quality at the point of delivery, which responds to expenditure only when that expenditure is specifically directed to personnel, training, equipment and protocol adherence.

Four caveats merit acknowledgement. First, the aggregate expenditure series does not distinguish primary from secondary and tertiary spending, nor salary from non-salary components, and the average elasticity estimated here conceals potentially large differences between them; Bhalotra (2007) suggests that the front-line component does most of the work. Second, the governance index is a proxy constructed from a small number of administrative indicators and is measured with error, which will attenuate the estimated interaction and means that the reported tercile contrast is, if anything, conservative. Third, state aggregates conceal within-state variation that district-level data would reveal, and the largest states in this panel contain populations exceeding those of most countries. Fourth, generalised method of moments estimation with internal instruments identifies the parameter from a particular source of variation, and the resulting estimate should be read as a local rather than a global average effect.

VI. CONCLUSION AND POLICY IMPLICATIONS

This study has estimated the effect of public health expenditure on infant mortality across twenty-one major Indian states over 1995–2023 within a dynamic panel framework that treats budget allocations as endogenous. Four findings emerge. First, infant mortality is highly persistent, with roughly thirty-eight per cent of the adjustment to a change in inputs completed within a year, so that long-run elasticities substantially exceed short-run ones. Second, the long-run elasticity of infant mortality with respect to real per capita public health expenditure is approximately −0.38 at average administrative capacity, and is substantially larger in absolute value than uninstrumented estimates suggest. Third, this effect is strongly conditional on governance quality: it is close to zero and statistically insignificant in the lowest tercile of measured administrative capacity and approaches −0.62 in the highest. Fourth, elasticities are roughly twice as large in the high-burden Empowered Action Group states as elsewhere, but markedly weaker for neonatal than for post-neonatal mortality.

Four policy implications follow. First, the case for raising public health spending towards the National Health Policy target is supported by the evidence, but the return to doing so is not uniform, and a fiscal strategy that treats the expenditure ratio as a sufficient statistic for health policy will disappoint. Second, and correspondingly, investment in absorptive capacity (filling sanctioned positions at primary and community health centres, reforming procurement, and ensuring timely release of funds to implementing units) is not an administrative footnote to health financing but a determinant of its productivity; on these estimates, capacity-building measures may raise the return to the existing budget by more than a plausible increase in the budget itself would deliver. Third, transfer design should recognise the tension between need and capacity: burden-weighted grants direct resources where the potential return is highest and the realised return currently lowest, which argues for pairing them with capacity-linked components and technical support rather than with output conditionalities alone. Fourth, the attenuation of the expenditure effect for neonatal mortality implies that further gains will depend on targeted investment in the quality of intrapartum and immediate newborn care (skilled birth attendance, functional special newborn care units, and protocol adherence) rather than on further expansion of coverage alone.

Three avenues for future research are promising. First, district-level analysis linking health accounts to facility surveys would permit finer identification and would reveal the within-state heterogeneity that state aggregates conceal. Second, decomposing expenditure by function (primary versus tertiary, salary versus non-salary, preventive versus curative) would identify which components of the budget carry the estimated average effect. Third, the endogenous evolution of administrative capacity itself deserves attention: if governance quality responds over time to sustained investment, the conditional elasticities estimated here understate the cumulative return to a policy that raises spending and capacity together.

REFERENCES

Anand, S., & Ravallion, M. (1993). Human development in poor countries: On the role of private incomes and public services. *Journal of Economic Perspectives, 7*(1), 133–150.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies, 58*(2), 277–297.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics, 68*(1), 29–51.

Auster, R., Leveson, I., & Sarachek, D. (1969). The production of health, an exploratory study. *The Journal of Human Resources, 4*(4), 411–436.

Banerjee, A., Deaton, A., & Duflo, E. (2004). Wealth, health, and health services in rural Rajasthan. *American Economic Review, 94*(2), 326–330.

Barenberg, A. J., Basu, D., & Soylu, C. (2017). The effect of public health expenditure on infant mortality: Evidence from a panel of Indian states, 1983–84 to 2011–12. *The Journal of Development Studies, 53*(10), 1765–1784.

Bhalotra, S. (2007). Spending to save? State health expenditure and infant mortality in India. *Health Economics, 16*(9), 911–928.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87*(1), 115–143.

Bokhari, F. A. S., Gai, Y., & Gottret, P. (2007). Government health expenditures and health outcomes. *Health Economics, 16*(3), 257–273.

Bond, S. R. (2002). Dynamic panel data models: A guide to micro data methods and practice. *Portuguese Economic Journal, 1*(2), 141–162.

Caldwell, J. C. (1979). Education as a factor in mortality decline: An examination of Nigerian data. *Population Studies, 33*(3), 395–413.

Census Commissioner of India. (various years). *Census of India*. Office of the Registrar General & Census Commissioner, Government of India.

Comptroller and Auditor General of India. (various years). *State finances audit reports*. Government of India.

Cutler, D., Deaton, A., & Lleras-Muney, A. (2006). The determinants of mortality. *Journal of Economic Perspectives, 20*(3), 97–120.

Deaton, A. (2013). *The great escape: Health, wealth, and the origins of inequality*. Princeton University Press.

Drèze, J., & Sen, A. (2013). *An uncertain glory: India and its contradictions*. Princeton University Press.

Farag, M., Nandakumar, A. K., Wallack, S., Hodgkin, D., Gaumer, G., & Erbil, C. (2013). Health expenditures, health outcomes and the role of good governance. *International Journal of Health Care Finance and Economics, 13*(1), 33–52.

Farahani, M., Subramanian, S. V., & Canning, D. (2010). Effects of state-level public spending on health on the mortality probability in India. *Health Economics, 19*(11), 1361–1376.

Filmer, D., & Pritchett, L. (1999). The impact of public spending on health: Does money matter? *Social Science & Medicine, 49*(10), 1309–1323.

Government of India. (2017). *National health policy 2017*. Ministry of Health and Family Welfare.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy, 80*(2), 223–255.

Gupta, S., Verhoeven, M., & Tiongson, E. R. (2002). The effectiveness of government spending on education and health care in developing and transition economies. *European Journal of Political Economy, 18*(4), 717–737.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica, 50*(4), 1029–1054.

International Institute for Population Sciences. (various years). *National family health survey (NFHS), India*. IIPS.

Kruk, M. E., Gage, A. D., Arsenault, C., Jordan, K., Leslie, H. H., Roder-DeWan, S., … Pate, M. (2018). High-quality health systems in the Sustainable Development Goals era: Time for a revolution. *The Lancet Global Health, 6*(11), e1196–e1252.

Ministry of Health and Family Welfare. (various years). *Rural health statistics*. Government of India.

Ministry of Statistics and Programme Implementation. (various years). *State domestic product statistics, National Sample Survey reports and Periodic Labour Force Survey reports*. Government of India.

Moreno-Serra, R., & Smith, P. C. (2012). Does progress towards universal health coverage improve population health? *The Lancet, 380*(9845), 917–923.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica, 49*(6), 1417–1426.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics, 22*(2), 265–312.

Pesaran, M. H. (2021). General diagnostic tests for cross-sectional dependence in panels. *Empirical Economics, 60*(1), 13–50.

Preston, S. H. (1975). The changing relation between mortality and level of economic development. *Population Studies, 29*(2), 231–248.

Rajkumar, A. S., & Swaroop, V. (2008). Public spending and outcomes: Does governance matter? *Journal of Development Economics, 86*(1), 96–111.

Registrar General of India. (various years). *Sample registration system statistical report*. Office of the Registrar General & Census Commissioner.

Reserve Bank of India. (various years). *State finances: A study of budgets*. Reserve Bank of India.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal, 9*(1), 86–136.

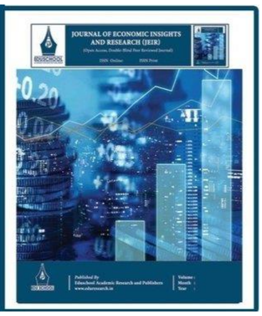
Sen, A. (1999). *Development as freedom*. Oxford University Press.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics, 126*(1), 25–51.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). MIT Press.

EDUSCHOOL
ACADEMIC RESEARCH & PUBLISHERS

JOURNAL OF ECONOMIC INSIGHTS AND RESEARCH (JEIR)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3107-9482 ISSN Print: 3139-1982



Size and Composition of Gig Workforce in India and Estimation of Its Size in the State of Assam and Guwahati Metropolitan Area

Mir Hussain[1*], Mofidul Hassan[2]

[1] Ph.D. Research Scholar, Department of Economics, Gauhati University, Assam, India
[2] Assistant Professor, Department of Economics, Gauhati University, Assam, India

Article information

Received: 19th July 2026
Received in revised form: 6th August 2026
Accepted: 10th August 2026
Available online: 25th August 2026

Volume: 2
Issue: 3
DOI: https://doi.org/10.63090/JEIR/3107.9482.0028

Abstract

In recent years, the gig work has emerged as a vital component of India's labour market, reshaping employment relations and creating new avenues for income generation, particularly in the service sector. However, the details about the size, growth, and composition of these emerging sectors have yet to be widely known. While some data are available at the national level, the state-level pictures are still sketchy. This paper examines the size and composition of the gig workforce at the all-India level and attempts to estimate its size in the state of Assam and the Guwahati Metropolitan Area. Further, the composition of the gig workforce, its sectoral, occupational, and skill level-based classifications, as well as their formal, informal, organised, and unorganised characteristics, have been examined at the all-India level. To estimate the gig workforce in Assam and Guwahati, a methodological framework has been developed using ratios of the state versus the national level of workforce and population. The paper uses data from the NITI Aayog (2022) report for analysis. The findings show that the gig workforce has been a growing component of the country's labour force.

Keywords: - Labour Market, Classification, Estimation, Projection, Assam & Guwahati Metropolitan Area.

I. INTRODUCTION

The rapid expansion and development of Information and Communication Technology and emergence of the digital world have given rise to a new type of job opportunities in the economy, collectively referred to as the gig economy, and the individuals involved in this segment are known as gig workers. Code on Social Security of the Government of India (2020) defines a gig worker as "a person who performs work or participates in a work arrangement and earns from such activities outside of traditional employer-employee relationships". According to NITI Aayog, the gig economy is characterized by flexibility of work and income, where long-term employer-employee relationships do not bind the worker. It includes all work arrangements outside the traditional employer-employee relationship. In short, in the gig economy, individuals engage in short-term, freelance tasks mediated through digital platforms, characterized by flexibility, autonomy, and task-specific arrangements. It offers flexibility, autonomy, and diverse income opportunities for workers while enabling firms to reduce costs and access talent efficiently (ASSOCHAM, 2020; Augustinraj & Bajaj, 2021). However, it is also associated with job insecurity, lack of social protection, and income instability, raising concerns about the quality and sustainability of such employment (Zelma, 2024). The global financial crisis of 2008 marked a pivotal moment in the growth of the gig economy, leading to an increased reliance on temporary and alternative forms of employment. The expansion of the gig economy is also evident in the Indian economy, where the COVID-19 pandemic resulted in a surge in food delivery, e-commerce, and freelance work. Platforms such as Swiggy, Zomato, and Urban Company experienced increased demand due to lockdowns (Business Standard, 2020). Mohanty & Jethy's (2022) study reveals a significant positive association between COVID cases and the number of gig workers in India. The present landscape, specifically the 2023-25 gig economy, has developed into a well-established sector of the labour market that is experiencing structural changes. It significantly contributes to workforce growth, skill specialization, technological integration, and sector expansion (Sharma & Sharma, 2025). Although several studies and reports provide estimates of the size and composition of the gig workforce at the national level in India, comparable evidence at the state and city levels remains limited. In particular, no reliable estimates are available for the gig workforce in Assam or the Guwahati Metropolitan Area. The present study seeks to address this gap. The specific objectives of the study being:

- To understand the size, composition, and growth trajectory of the gig workforce in India.
- To estimate the size of the gig workforce in the state of Assam, with special reference to the Guwahati Metropolitan Area, Guwahati being the largest metropolis in the north eastern part of the country.

II. REVIEW OF LITERATURE

Datta et al. (2023) estimate that online gig work accounts for 4.4% to 12.5% of the global workforce, corresponding to a broad range of approximately 154 million to 435 million workers worldwide, using 545 online gig work platforms across 186 countries. Compared to industrialized nations, emerging nations have a higher participation rate in the gig economy. About 1 to 4 percent of workers in developed countries such as the USA, Germany, the UK, and Sweden rely on gig work as their main source of income. In contrast, between 5 to 12 percent of workers in developing countries like China, India, and Brazil depend on gig work as their primary income source, with around 30 percent using it as a secondary source of income (Augustinraj & Bajaj, 2021). V. V. Giri National Labour Institute (2025) projected the total number of gig workers for the year 2047. The point estimate predicts that the number will increase to 6.16 crores by 2047, representing 14.89 percent of India's non-agricultural workforce. By 2047, the sector will be sufficient to create 9.08 crores of employment. Also, according to Augustinraj & Bajaj (2021), the gig economy has the potential to support up to 90 million jobs in the non-farm sector in about 8 to 10 years in India's economy. It is expected that around 24 million jobs may transition to technology-based gig platforms in the near to medium term. Additionally, it is estimated that nearly one million net new jobs could be created over the next two to three years to address constraints in specific operational roles and to meet the significant 'latent demand' for services among households. In the coming decade, the gig economy in India is projected to witness USD 250 billion in transactions, accounting for 1.25% of the nation’s gross domestic product. These projections should not be interpreted as directly comparable because they are based on different definitions of gig work, estimation methodologies, data sources, and projection horizons. The V. V. Giri projection is a long-term forecast extending to 2047, whereas the BCG estimate reflects a different forecast period and assumptions regarding the growth of the digital platform economy. Consequently, differences in the projected workforce size reflect methodological and definitional differences. However, studies of gig work at the sub-national level are few and far between. The present paper is an attempt to fill the void.

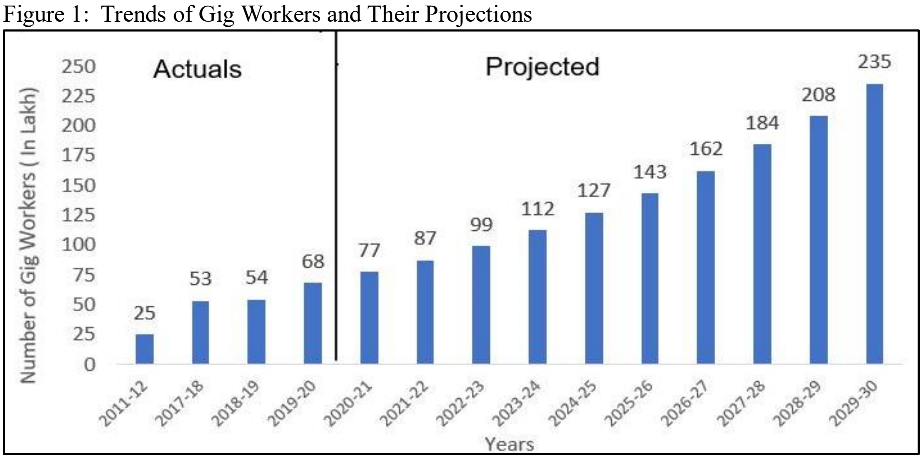
III. METHODOLOGY

The paper consists of two broad parts. The first part utilizes available secondary data to understand the size and composition of the gig workforce in India, as well as its projected growth. The basic data is taken from the NITI Aayog (2022) report, “India’s booming gig and platform economy: Perspectives and recommendations on the future of work”. The analysis is done by calculating ratios and making graphical presentation. The patterns emerging from the graphical presentations have induced us to seek plausible explanations of the undergoing compositional shifts within the gig workforce. The second part of the study makes assumptions for estimating the component of the all-India gig workforce, i.e., serving in the state of Assam, with special reference to the Guwahati Metropolitan Area. The details of its assumptions and the methodology for estimating them are laid out in Section 5.

IV. SIZE AND COMPOSITION OF GIG WORKFORCE IN INDIA

Estimating the size of the gig economy is a tricky problem. NITI Aayog has attempted to estimate its size and growth by sector based on the number of workers engaged. They have used the Periodic Labour Force Survey data of the NSSO to extract the number of workers engaged in gig work from the total number of workers. Their broad estimates are for the years 2011-12, 2017-18, 2018-19, and 2019-20. Based on these basic estimates, they have computed not only a time series of the number of gig workers in the Indian economy but also projected the series up to the year 2029-30. The data is presented in Figure 1. Given this obvious growth in the size of the gig economy in terms of workers engaged in it, it is pertinent to ask what proportion of the total workforce and total non-agricultural workforce gig workers constitute. This has been estimated, and the results are presented in Figure 2.

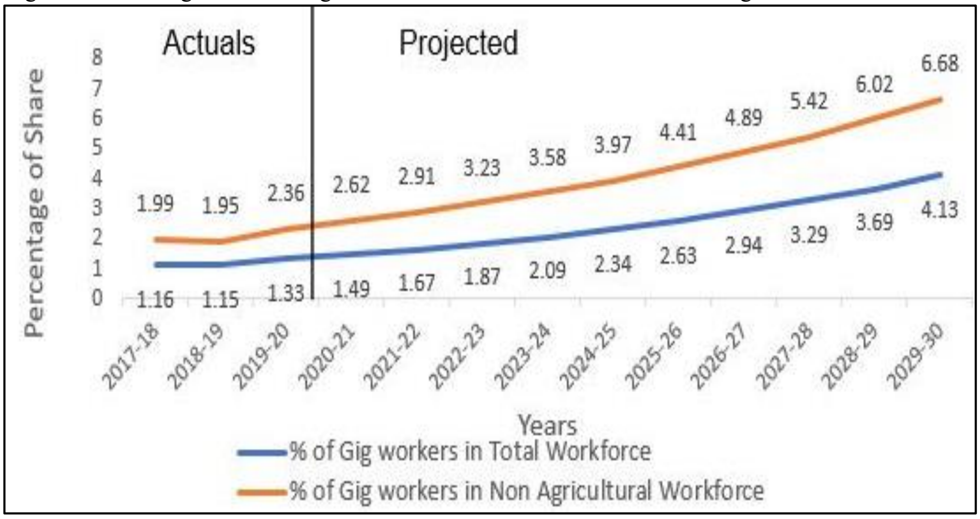
Figure 1: Trends of Gig Workers and Their Projections



Source: Author's Construction Using Data from NITI Aayog (2022).

From Figure 1, it is clear that the gig economy has the potential to increase from 25 lakh in 2011 to 2.35 crore in 2029-30. The next step involves determining the proportion of gig workers within the total workforce and specifically within the non-agricultural workforce. This can also be demonstrated from the data (actual and projected) estimated by NITI Aayog. The share of gig workers in the total workforce and the non-agricultural workforce in 2011-12 is 0.54% and 1.03%, respectively.

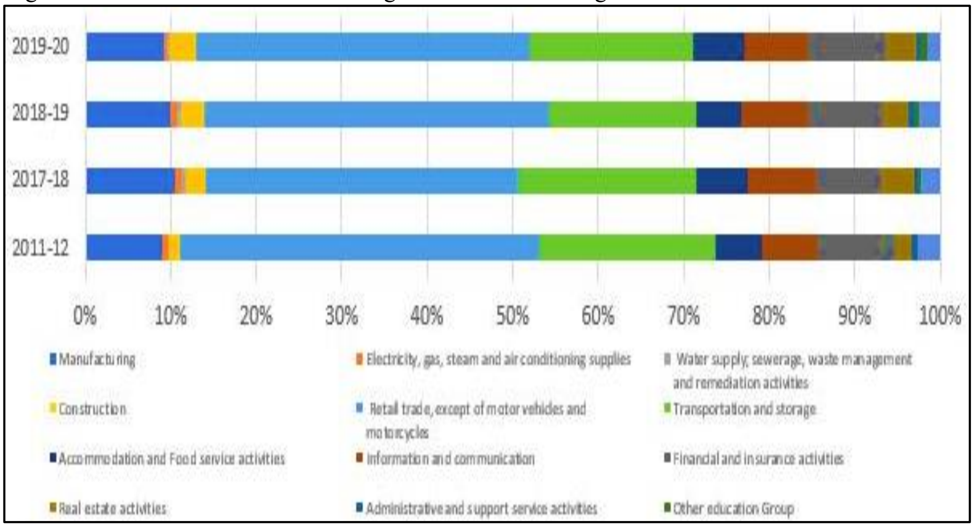
Figure 2: Percentage Share of Gig Workforce in Total Workforce and Non-Agricultural Workforce



Source: Authors' Construction Using Data from NITI Aayog (2022)

Figure 2 shows that in 2017-18, the share of gig workers in the total and non-agricultural workforce was 1.16% and 1.99%, respectively. However, in the next year, i.e., 2018-19, this share decreased to 1.15% and 1.95%, respectively. But in 2019-20, the share in both categories increased significantly to 1.33% and 2.36%, respectively. Platform-mediated gig work, which integrates the gig economy with mainstream economies, has risen sharply, mostly as a result of the COVID pandemic and lockdowns around the world (Mohanty & Jethy, 2022; International Labour Organization, 2024). NITI Aayog has projected that the share of gig workers in the total workforce and the non-agricultural workforce will increase to 4.13% and 6.68%, respectively, by 2029-30. Since gig work is virtually absent in the agricultural sector, it may be of interest to look at their presence in the non-agricultural workforce of the countries.

Figure 3: Industrial Classification of Gig Workers in Percentage Share

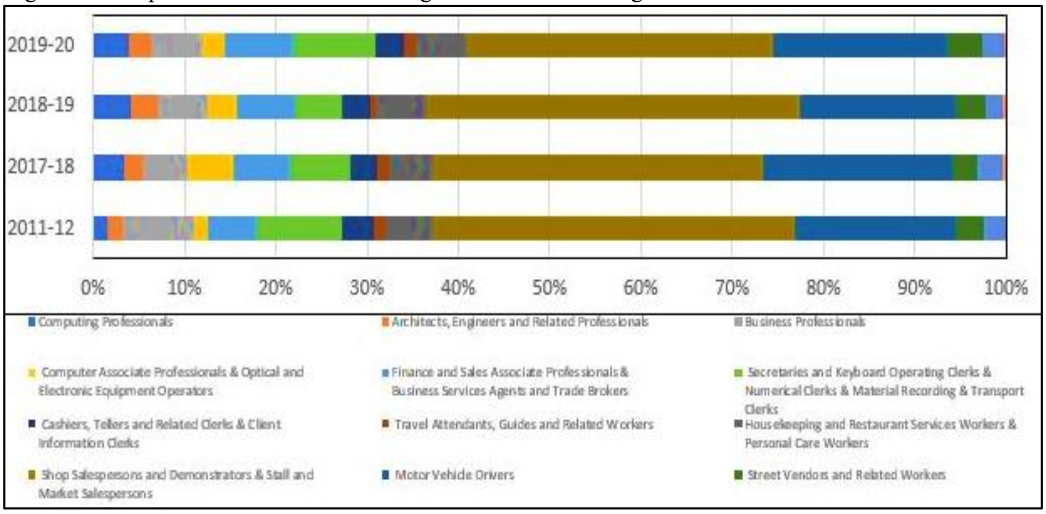


;

Source: Authors' Construction Using Data from NITI Aayog (2022)

The following industries and occupations are identified based on the probable characteristics of gig workers by the NITI Aayog. The share of gig workers by sectors and occupational categories are presented in Figures 3 and 4. Figure 3 presents the industrial distribution of gig workers in India over the following four years, i.e., 2011–12, 2017–18, 2018–19, and 2019–20. It shows that the retail trade sector has continuously employed the largest share of gig workers during this time, with a range of 36% to 42%. This indicates how prevalent sales, delivery, and service-related activities are in India's gig economy. Following closely behind with 17–21% of gig jobs is the transportation and storage industry, which has been greatly impacted by the rapid growth of ride-hailing and logistics platforms like Ola, Swiggy, and Amazon. The construction industry saw a notable increase from 1.24% to 3.06%, reflecting the growing use of project-based gig contracts, while the manufacturing sector maintained a stable share, rising marginally from 9.09% in 2011–12 to 9.17% in 2019–20. High-skilled, digital industries like communication and information, as well as financial and insurance operations, have consistently contributed between 7% and 9% over time. The real estate sector expanded from 1.72% in 2011–12 to 3.53% in 2019–20, and education-related activities rose modestly from 0.08% to 0.67%, showing diversification in gig opportunities. Additionally, the proportion of gig workers is increasing in sectors that previously had little or no presence of such workers.

Figure 4: Occupational Classification of Gig Workers in Percentage Share

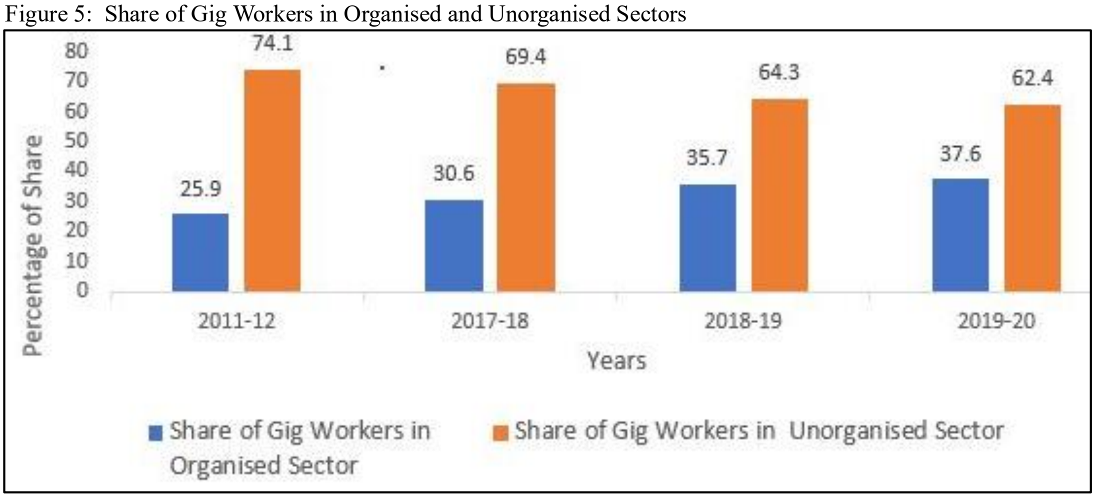


Source: Authors' Construction Using Data from NITI Aayog (2022)

When examined from an occupational perspective, Figure 4 illustrates the key occupations where the gig workforce is predominantly concentrated. Figure 4 shows the occupational distribution of gig workers in India for the years 2011–12, 2017-18, 2018-19, and 2019–20. Three occupational segments, specifically drivers, sales personnel, and restaurant service workers, collectively represented 58%-60% of gig workers across all four years of the study. This indicates a steady growth in the gig economy within these occupations. At the same time, there has been a gradual increase in the share of high-skilled gig workers, particularly computing, business, and finance professionals, reflecting the expansion of digital and knowledge-based work. In contrast, clerical and administrative jobs such as secretarial and data-entry roles have declined due to automation and digitalisation.

Despite these changes, low- and medium-skilled service jobs, like driving and sales, continue to dominate the gig economy. Overall, this trend indicates a slow but steady diversification of gig work in India, moving from purely service-based activities to a broader mix of traditional and technology-driven occupations. This industrial and occupational classification gives an idea about the nature of gig workers and their diversification. The data shows that the majority of gig workers are engaged in occupations such as sales, driving, and restaurant services, with sectoral concentration in transportation, trade, accommodation, and food services. In contrast, sectors like information and communication technology (ICT), financial services, real estate, and administrative activities, as well as occupations such as computing professionals and computer associates, though part of the broader service sector, account for a relatively smaller share of gig workers. These areas generally demand specialized skills, long-term engagement, and formal contractual arrangements. This distinction underscores that gig work primarily expands in activities that can be easily divided, outsourced, or digitally mediated, rather than in structured and professionally regulated domains of the service economy. Also, to capture the share of gig workers in the organised and unorganised sectors, the distribution is presented in Figure 5.

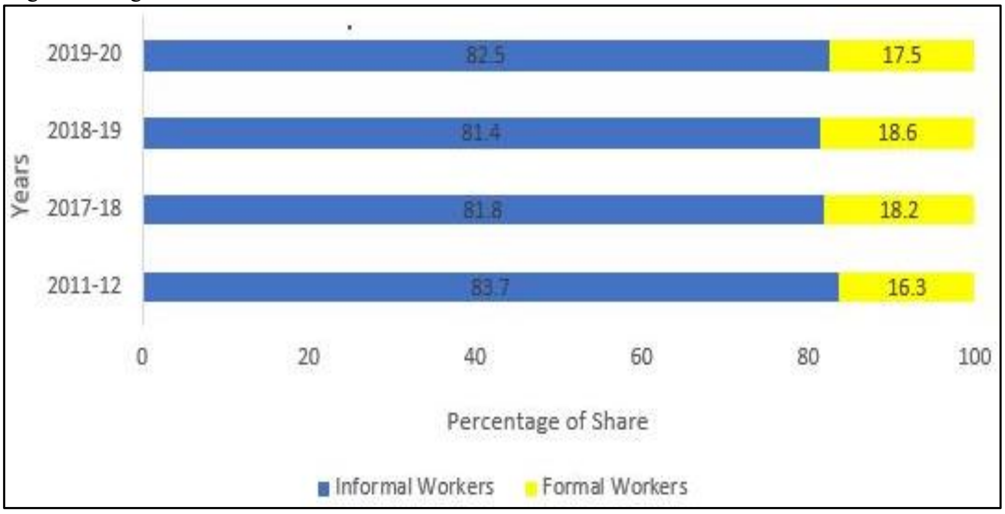
Figure 5: Share of Gig Workers in Organised and Unorganised Sectors



Source: Authors' Construction Using Data from NITI Aayog (2022)

Figure 5 illustrates the share of gig workers across the organised and unorganised sectors for 2011–12, 2017-18, 2018-19, and 2019–20. The data reveal that the majority of gig workers remain concentrated in the unorganised sector, accounting for about 62.4% of the total in 2019–20. However, the share of gig workers in the organised sector has been steadily increasing, indicating that more of these workers are now associated with larger firms. The proportion rose from 26% in 2011–12 to 37.6% in 2019–20, with a corresponding decline in the share of the unorganised sector. Between 2011–12 and 2019–20, India underwent a major transformation driven by rapid digitalisation, widespread smartphone adoption, and increased internet accessibility. These changes facilitated the emergence of platform-based businesses such as Swiggy, Zomato, Ola, Uber, Amazon, and Urban Company, which operate as large, organised firms. Before 2011, most “gig-type” jobs, such as delivery, driving, or household repair services, were largely concentrated in the unorganised economy, typically carried out through small-scale or local arrangements. The growing share of the organised sector in the gig workforce is apparently a positive development. However, this positive development needs to be qualified by the fact that although gig workers now operate within the organised sector, their labour relations, particularly in terms of contracts and employment conditions, remain largely informal, even though the company structures themselves are formal and organised. This is illustrated in Figure 6.

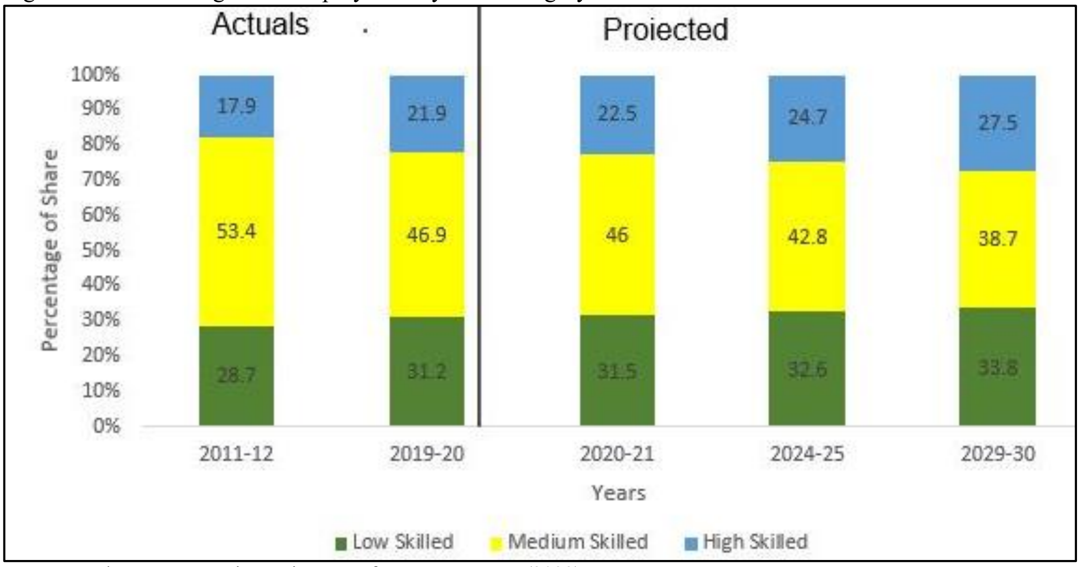
Figure 6: Gig Workers: Formal and Informal Workers Share



Source: Authors' Construction Using Data from NITI Aayog (2022)

Figure 6 presents the distribution of gig workers between formal and informal employment categories for the years 2011–12, 2017-18, 2018-19, and 2019–20, based on NITI Aayog's estimates. Across all four years, informal workers consistently dominate the gig workforce, accounting for more than 80% of all gig workers. This indicates that the majority of gig workers continue to operate without formal contracts, social security coverage, or employment-related benefits, reflecting the largely informal nature of work within India's gig economy. However, a portion, 17%–19%, of gig workers fall under the formal category, mainly representing those engaged through organised platforms or firms where some level of documentation, registration, or regulatory compliance exists, even though their employment relationships often remain contractual or semi-formal in practice. Also, high-skilled professional gig work, such as specialized consultants, software developers, and finance professionals, operates under formal, written contracts for projects, which further aligns this segment with the characteristics of formal employment despite its flexible and non-permanent nature. If we see gig work by skill-wise, which is presented in Figure 7.

Figure 7: Share of Gig Work Employment by Skill Category



Source: Authors' Construction Using Data from NITI Aayog (2022)

Figure 7 presents the actual and projected distribution of gig work by skill category for the years 2011–12, 2019-20 (Actuals) to 2020-21, 2024-25, and 2029-30 (Projected). In 2011–12, the majority of gig workers were medium-skilled (53.4%), followed by low-skilled (28.7%) and high-skilled (17.9%) workers. However, over time, the share of medium-skilled workers is expected to decline steadily to 38.7% by 2029–30. In contrast, high-skilled gig work is projected to rise from 17.9% to 27.5%, and low-skilled work from 28.7% to 33.8% during the same period. This trend suggests a polarisation of the gig economy, where over time, the proportions of high-skilled and low-skilled gig workers increase but the medium-skilled workers constitute the majority in the skilled profile, and their size is reduced. The increase in the share of high-skilled components is presumably because many medium-skilled workers upgrade their skills through learning by doing and move to the high-skilled category. As for the rising share of the low-skilled category, the possible explanation seems to be the push factor of the overcrowded unorganised sector and the pull factor of somewhat better earnings in the gig economy. The low entry barriers for the lower skill segments of the gig economy presumably keep replenishing this segment more than the number of those who upgrade themselves to the medium-skilled category.

V. SIZE OF GIG WORKFORCE IN ASSAM AND GUWAHATI METROPOLITAN AREA

Direct estimation of the gig workforce for Assam and the Guwahati Metropolitan Area using the PLFS is constrained by the relatively small sample size available for these geographical units. Therefore, this study adopts an indirect estimation approach, using NITI Aayog's national estimates as the benchmark and adapting them to the state and metropolitan levels through a set of assumptions. The methodology for estimating the state-level and area-level size of the gig workers is outlined.

5.1. Estimation of Gig Workers in Assam for the Year 2023-24

Although gig work is largely concentrated in metropolitan and other urban centres, the estimation of the gig workforce in this study follows the methodology adopted by NITI Aayog (2022), which estimates the national gig-worker share using the total workforce identified from both urban and rural samples of the PLFS. Accordingly, the total workforce of Assam is taken as the base for estimating the state's gig workforce. Therefore, at first, we need to estimate the total workforce of Assam for the year 2023-24. The following formula is used to estimate the total workforce of Assam in 2023-24.

Total Workforce of Assam in 2023–24 = Total Population of Assam in 2023–24 × Worker Population Ratio (WPR) (Usual Status, PS+SS, All Ages) for Assam in 2023–24.

The projected population for Assam in 2023–24 is 35,991,500(StatisticsTimes.com, 2024). According to the Periodic Labour Force Survey (PLFS) Annual Report 2023–24, the Worker Population Ratio (WPR) in usual status (ps+ss) for all ages for Assam is 47.2%. Therefore,

$$\begin{aligned}\text{Total Workforce of Assam in } 2023-24 &= 35{,}991{,}500 \times 47.2\% \\ &= 16{,}987{,}988\end{aligned}$$

The total workforce of Assam in 2023-24 is 16,987,988. For estimating the gig workforce of Assam for the year 2023-24, this study considers the gig-worker proportion reported by NITI Aayog (2022), which indicates that gig workers constituted approximately 2.09% of India's total workforce in 2023–24 (as presented in Figure 2). If the national gig-worker share is applied directly, the estimated gig workforce of Assam is calculated as:

Estimation of Gig Workforce of Assam in 2023-24 = Total Workforce of Assam in 2023-24 × Share of Gig Workforce in Total Workforce in 2023-24

$$\begin{aligned}\text{Estimation of Gig Workforce of Assam in } 2023-24 &= 16{,}987{,}988 \times 2.09\% \\ &\approx 355{,}049\end{aligned}$$

However, applying this national-level proportion directly to Assam may not accurately reflect the state's actual gig-workforce size, given the significant variation in urbanisation patterns between Assam and the country as a whole. Since gig work is generally more concentrated in urban and metropolitan areas, the higher urbanisation level of India compared to Assam may result in an upward bias if the national gig-worker ratio is applied without adjustment. Therefore, an urbanisation-adjusted approach is adopted to estimate. The adjusted gig-worker estimate for Assam is derived by scaling the national gig-worker share according to Assam's urban population proportion relative to the national urban population proportion. The formula is

$$\begin{gathered}\text{Estimated Gig Workforce of Assam in 2023–24} = \\ \text{Total Workforce of Assam in } 2023-24 \times \text{Adjusted Gig} - \\ \text{Worker Share of Assam in Total Workforce in 2023–24}\end{gathered} \tag{1}$$

Where,

$$\begin{gathered}\text{Adjusted Gig} - \text{Worker Share of Assam in Total Workforce in 2023–24} = \\ \text{National Gig} - \text{Worker Share in } 2023-24 \times \\ \text{Urbanisation Adjustment Ratio}\end{gathered} \tag{2}$$

$$\text{Urbanisation Adjustment Ratio} = \frac{\text{Urbanisation Rate of Assam in } 2023{-}24}{\text{Urbanisation Rate of India in } 2023{-}24}$$

Where,

$$\text{Urbanisation Rate of Assam in } 2023-24 = \frac{\text{Urban Population of Assam in } 2023-24}{\text{Total Population of Assam in } 2023-24} \times 100$$

$$\text{Urbanisation Rate of Assam in } 2023-24 = \frac{5{,}896{,}388}{35{,}991{,}500} \times 100$$

$$Urbanisation\ Rate\ of\ Assam\ in\ 2023-24 = 16.38\%$$

Also,

$$\text{Urbanisation Rate of India in } 2023-24 = \frac{\text{Urban Population of Indiain } 2023-24}{\text{Total Population of India in } 2023-24} \times 100$$

Since the urban population of India in 2023-24 is 508,811,448 (Macrotrends, n.d.), and the total population of India in 2023-24 is 1,448,163,288 (Macrotrends, n.d.) Therefore,

$$\text{Urbanisation Rate of India in } 2023-24 = \frac{508{,}811{,}448}{1{,}448{,}163{,}288} \times 100$$

$$\text{Urbanisation Rate of India in } 2023-24 = 35.13\%$$

Now,

$$\text{Urbanisation Adjustment Ratio} = \frac{16.38}{35.13}$$

Urbanisation Adjustment Ratio = 0.4663 or 46.63%

Substituting the values into equation (2)

$$\text{Adjusted Gig} - \text{Worker Share of Assam in Total Workforce in } 2023\text{–}24 = 2.09\ \% \times 46.63\%$$
$$= 0.974\%$$

Now, substituting the values into equation (1)

$$\text{Estimated Gig Workforce of Assam in } 2023\text{–}24 = 16{,}987{,}988 \times 0.974\%$$
$$\approx 165{,}463$$

5.2. Estimation of Gig Workforce in Guwahati Metropolitan Area for the Year 2023-24

Since official estimates of the spatial distribution of gig workers across districts and urban centres in Assam are unavailable, the estimated gig workforce of the state is allocated to the Guwahati Metropolitan Area based on its share in Assam's total urban population. This allocation criterion is adopted because the urban population serves as a proxy for the spatial distribution of gig work, reflecting the greater concentration of economic activity, consumer demand, and digital infrastructure in urban areas. In the absence of district- or city-level gig-workforce data, this approach provides a practical basis for estimating the gig workforce of the Guwahati Metropolitan Area. However, the estimate should be interpreted as an approximation, as it assumes that the spatial distribution of gig workers is proportional to the distribution of the urban population. Accordingly, the first step is to estimate the share of the Guwahati Metropolitan Area in Assam's total urban population for the year 2023–24, which is calculated as follows:

-

$$\begin{gathered}\text{Share of Urban population of Guwahati Metropolitan Area in} \\ \text{Assam's Total Urban Population in } 2023-24 = \text{Urban Population of Guwahati Metropolitan Area} / \\ \text{Total Urban Population of Assam in } 2023-24 \times 100\end{gathered} \quad (3)$$

Since the official annual population projections for the Guwahati Metropolitan Area (GMDA) are not available. Therefore, the projected population of the Guwahati Urban Agglomeration (UA), estimated at 1,345,500 for 2023–24 (Census2011.co.in, n.d.), is used as a proxy for estimating the size of the gig workforce in the Guwahati Metropolitan Area. Since the UA substantially overlaps the metropolitan urban core, this proxy is considered appropriate, although it may introduce some measurement error. The total urban population of Assam for the year 2023-24 is not directly available. Therefore, we need to estimate this number. To estimate the urban population of Assam for the year 2023–24, the compound growth rate method was used, based on the decadal growth rate of Assam's urban population for the 2001–2011 period and taking Assam's urban population of 2011 as the base year.

According to 2011 Census data, Assam's urban population is 4,388,756 (Government of Assam, Directorate of Town & Country Planning, n.d.), and the decadal growth rate of Assam's urban population from 2001 to 2011 is 27.9% or 0.279 (Census2011.co.in, n.d). For the projection, it has been assumed that the 2001–11 decadal growth rate continues at a uniform rate in the following years.

The projection formula used is as follows:

$$P_t = P_{2011} \times (1 + r)^{t/10} \quad (4)$$

Where,

P_t = Projected urban population of Assam in 2023–24

P_{2011} = Assam's urban population in 2011 according to census data

r = Decadal growth rate of Assam's urban population from 2001 to 2011

t = Time period since 2011 (12 years)

Substituting the values in equation (4)

$$P_{2023-24} = 4{,}388{,}756 \times (1 + 0.279)^{12/10}$$

$$P_{2023-24} = 4{,}388{,}756 \times (1.279)^{1.2}$$

$$P_{2023-24} = 4{,}388{,}756 \times 1.3436$$

$$P_{2023-24} = 5{,}896{,}388$$

Hence, the projected urban population of Assam for the year 2023–24 is 5,896,388
Now, substituting the value in equation (3)
Share of Urban population of Guwahati Metropolitan Area in Assam's Total Urban Population in 2023-24

$$= 1{,}345{,}500 / 5{,}896{,}388 \times 100$$

$$= 0.2282 \times 100$$

$$= 22.82\%$$

Now, to estimate the Gig workforce of the Guwahati Metropolitan Area for the year 2023-24, the following formula is used. i.e.
Gig Workforce in the Guwahati Metropolitan Area for the Year 2023-24
= Estimated Gig Workforce of Assam in 2023-24 × Share of Urban population of Guwahati Metropolitan Area in Assam's Total Urban Population in 2023-24

$$= 165{,}463 \times 22.82\%$$

$$\approx 37{,}758$$

Hence, the estimated gig workforce of the Guwahati Metropolitan Area for the year 2023–24 is 37,758.

The urban population share-based estimate assumes that gig workers are distributed across urban areas in proportion to their share of the urban population, implying uniform penetration across urban centres. However, it is common knowledge that the penetration of gig work increases more than proportionally with the population size of urban centres. Guwahati Metropolitan Area, being the principal metropolitan, commercial, and service centre of Assam, the concentration of gig workers here will be much higher than in the other urban centres of the state. Therefore, an alternative estimate is developed by adjusting Guwahati's urban population share using the Urban Primacy Index, which measures the dominance of the largest urban agglomeration relative to the second-largest urban agglomeration (Jefferson, 1939; Dick & Rimmer, 2019; Tewari, n.d.). This Index is used to account for the likelihood that gig workers are disproportionately concentrated in Guwahati due to its dominant position within Assam's urban system. It is calculated as:

$$\text{Actual Urban Primacy Index} = \frac{\text{Population of the Largest Urban Agglomeration in } 2023-24}{\text{Population of the Second} - \text{largest Urban Agglomeration in } 2023-24}$$

According to the Census of India 2011, Guwahati and Silchar are the first- and second-largest Urban Agglomerations in Assam, respectively. The projected population of the Silchar Urban Agglomeration for 2023–24 is estimated at 318,500 (Census2011.co.in, n.d.).The Actual Urban Primacy Index is therefore calculated as follows:

$$\text{Actual Urban Primacy Index} = \frac{\text{Population of Guwahati Urban Agglomeration in } 2023-24}{\text{Population of Silchar Urban Agglomeration in } 2023-24}$$

Substituting the projected population values:

$$\text{Actual Urban Primacy Index} = \frac{1{,}345{,}500}{318{,}500} = 4.22$$

The calculated Urban Primacy Index of 4.22 indicates that the population of the Guwahati Urban Agglomeration is approximately 4.22 times larger than that of the Silchar Urban Agglomeration. This highlights the dominant position of Guwahati within Assam's urban hierarchy. According to the rank-size rule (Zipf's Law), a balanced urban hierarchy is expected to have a primacy index of approximately 2.0, where the largest urban agglomeration is about twice the size of the second-largest urban agglomeration (Zipf, 1949; Dick & Rimmer, 2019; Tewari, n.d.). The relative urban primacy factor is therefore calculated as:

$$\text{Relative Urban Primacy Factor} = \frac{\text{Actual Urban Primacy Index}}{\text{Zipf} - \text{expected Primacy Index}}$$

$$\text{Relative Urban Primacy Factor} = \frac{4.22}{2} = 2.11$$

A value greater than one indicates that Guwahati's dominance exceeds the level expected under a balanced urban hierarchy. The Relative Urban Primacy Factor of 2.11 suggests that Guwahati's dominance is approximately twice the level implied by the rank-size rule. Accordingly, this factor is used to adjust Guwahati's urban population share to examine the possible implications of a higher concentration of gig workers in the Guwahati Metropolitan Area. The adjusted allocation share of the Guwahati Metropolitan Area is calculated as follows:

Adjusted Allocation Share = Share of Guwahati's Urban Population in Assam's Total Urban Population × Relative Urban Primacy Factor

$$= 22.82\% \times 2.11$$

$$= 48.15\%$$

Accordingly, the estimated gig workforce of the Guwahati Metropolitan Area based on the adjusted allocation share is calculated as follows:

= Estimated Gig Workforce of Assam in 2023–24 × Adjusted Allocation Share

= 165,463 × 48.15%

= 79,670

Thus, the estimated gig workforce of the Guwahati Metropolitan Area ranges from 37,758 workers under the urban population share-based estimate and approximately 79,670 workers based on the adjusted allocation share. The latter is presented as an alternative estimate to illustrate the possible implications of a higher concentration of gig workers in the Guwahati Metropolitan Area.

5.3. Projected Gig Workforce in Assam and the Guwahati Metropolitan Area up to 2030

Based on the estimates for the year 2023–24, the gig workforce is projected for the period 2024–25 to 2029–30 using the forward projection method. Annual growth rates specific to these geographical units are not available; however, NITI Aayog (2022) provides projected estimates of India's gig workforce up to 2029–30, from which the corresponding year-to-year national growth rates can be derived. In the absence of state- and city-level growth projections, it is assumed that the gig workforce in Assam and the Guwahati Metropolitan Area follows the same annual growth pattern as the national gig workforce. Accordingly, the estimated gig workforce for 2023–24 is projected forward to obtain estimates for the period 2024–25 to 2029–30.

The forward projection is carried out using the following formula:

$$P_t = P_{t-1} \times (1 + r_t) \tag{5}$$

Where,

P_t = Projected population in year t

P_{t-1} = Population in the previous year

r_t= Annual growth rate (NITI Aayog's year-specific national growth rate for year t)

t = Number of years into the future (Per year)

Based on this, we calculate the gig workforce for Assam and the Guwahati Metropolitan Area for the years 2024-25 to 2029-30. This will be shown in Table 1.

Table 1. Projected Size of Gig Workforce in Assam and Guwahati Metropolitan Area

Year	Growth Rate (%)	Projected Gig Workforce: Assam	Projected Gig Workforce: Guwahati Metropolitan Area	
			Urban Population Share-Based Estimate	Estimate Based on Adjusted Allocation Share
2023-24	-	165,463	37,758	79,670
2024-25	13.39%	188,000	43,000	90,000
2025-26	12.60%	211,000	48,000	102,000
2026-27	13.29%	239,000	55,000	115,000
2027-28	13.58%	272,000	62,000	131,000
2028-29	13.04%	307,000	70,000	148,000

From Table 1, which presents the projected gig workforce for Assam and the Guwahati Metropolitan Area, it is observed that, for the year 2026–27, the projected gig workforce is approximately 239,000 in Assam, and between 55,000 and 115,000 in the Guwahati Metropolitan Area.

VI. LIMITATIONS

In the absence of reliable state- and city-level data on the gig workforce, the authors had to make a number of assumptions to estimate the size of the gig workforce in Assam and the Guwahati Metropolitan Area. As more comprehensive survey-based or administrative data become available in the future, these estimates can be further revised and refined. The principal limitations of the present estimates are as follows:

First, the estimated gig workforce is derived indirectly using the national gig-worker share projected by NITI Aayog (2022), which is based on occupation-code proxies rather than direct identification of gig workers. Consequently, all subsequent estimates inherit the assumptions underlying the NITI Aayog methodology.

Second, although the study adjusts the national gig-worker share for Assam using the urbanisation ratio, the estimation assumes that urbanisation is appropriate for the spatial concentration of gig work. Actual penetration may differ across states and cities because of variations in digital infrastructure, economic activity, and platform availability.

Third, the allocation of Assam's estimated gig workforce to the Guwahati Metropolitan Area is based primarily on its share of the state's urban population. This approach assumes a proportional distribution of gig workers across urban centres. Since larger metropolitan areas generally exhibit higher concentrations of economic activities, an alternative estimate is presented to illustrate the possible implications of a higher concentration of gig workers in the Guwahati Metropolitan Area.

VII. CONCLUSION

Though the gig workers as yet comprise a small part of the total workforce in India, their share in the workforce has been rising rapidly and is expected to rise even faster in the coming years. This gig workforce is primarily concentrated in sectors such as transportation, retail trade, accommodation, and food services. Occupation-wise, it is largely engaged in driving, sales, and restaurant service activities, etc. The majority of gig workers are concentrated in the unorganised sector. But the share of gig workers in the organised sector has been rising in the recent years with the rise of the digital economy and the emergence of platforms such as Swiggy, Zomato, Ola, and Uber etc. These platforms act as intermediaries between sellers and buyers of goods or services and are classified under organised sector activities. However, it is important to note that despite serving the organised sector, the employment relationships and working conditions of these workers remain largely informal in nature. It is therefore no surprise that an overwhelming number of gig workers are found located in the informal sector when we do a formal-informal classification of the gig workforce. Skill-wise, the gig workforce mostly falls in the medium-skilled category; however, the skill-wise composition of these workers has changed over time. There have been increases in the shares of high-skilled and low-skilled workers, resulting in some decline in the share of medium-skilled workers. The increased shares of the high-skilled category can be understood in terms of upward mobility of the skill level among workers over time from the process of "Learning by Doing". Increased share of the low-skill segment, on the other hand, can be ascribed to the push factor of the overcrowding in the general unorganised sector and the pull factor of somewhat better labour earnings in the gig economy. The estimate of the size of the gig workforce in Assam in general and the Guwahati Metropolitan Area in particular, for the year 2023-24, and projections up to the year 2030, are expected to assist policymakers and researchers in designing informed strategies and policies related to the gig economy there.

REFERENCES

ASSOCHAM. (2020, January). *Gig economy aligning consumer preferences: The way forward* (Background note for conference). The Associated Chambers of Commerce and Industry of India. https://www.assocham.org/uploads/files/1628143386.pdf

Augustinraj, R., & Bajaj, S. (2021). *Unlocking the potential of the gig economy in India*. BCG & Michael & Susan Dell Foundation.

Business Standard. (2020, September 30). India gig workers, hit by COVID, need tech for financial resilience: Report. https://www.business-standard.com/article/economy-policy/india-gig-workers-hit-by-covid-need-tech-for-financial-resilience-report-120093001053_1.html

Census2011.co.in. (n.d.-a). *Assam population 2025: Sex ratio & literacy rate 2025*. https://www.census2011.co.in/census/state/assam.html

Census2011.co.in. (n.d.-b). *Guwahati city population 2025, literacy, and Hindu Muslim population*. Retrieved November 9, 2025, from https://www.census2011.co.in/census/city/191-guwahati.html

Census2011.co.in. (n.d.-c). *Silchar metropolitan urban region population 2011–2026 census*. Retrieved August 5, 2026, from https://www.census2011.co.in/census/metropolitan/178-silchar.html

Datta, N., Rong, C., Singh, S., Stinshoff, C., Iacob, N., Nigatu, N. S., … Klimaviciute, L. (2023). *Working without borders: The promise and peril of online gig work* (No. 32573393). World Bank.

Dick, H., & Rimmer, P. (2019). Primate city. In A. M. Orum (Ed.), *The Wiley Blackwell encyclopedia of urban and regional studies*. Wiley. https://doi.org/10.1002/9781118568446.eurs0247

Government of Assam, Directorate of Town & Country Planning. (n.d.). *General urbanization*. Directorate of Town & Country Planning, Government of Assam. Retrieved August 1, 2025, from https://tcp.assam.gov.in/information-services/general-urbanization

Government of India. (2020). *The Code on Social Security, 2020* (No. 36 of 2020). The Gazette of India. https://www.indiacode.nic.in/bitstream/123456789/16823/1/aA2020-36.pdf

International Labour Organization. (2024). *Expansion of the gig and platform economy in India: Opportunities for employer and business member organizations*. International Labour Office.

Jefferson, M. (1939). The law of the primate city. *Geographical Review, 29*(2), 226–232. https://doi.org/10.2307/209944

Macrotrends. (n.d.-a). *India population (1950–2026)*. Retrieved August 4, 2026, from https://www.macrotrends.net/global-metrics/countries/ind/india/population

Macrotrends. (n.d.-b). *India urban population (1960–2025)*. Retrieved August 4, 2026, from https://www.macrotrends.net/global-metrics/countries/ind/india/urban-population

Mohanty, S., & Jethy, J. (2022). COVID-19: A salient driver of the emerging gig economy in India. *Odisha Journal of Commerce and Management, 15*.

NITI Aayog. (2022, June). *India's booming gig and platform economy: Perspectives and recommendations on the future of work* (Policy Brief).

Sharma, A. K., & Sharma, R. (2025). The gig economy and the evolving nature of work in India: Employment, policy, and platform realities in the age of convenience. *Journal of Digital Economy, 4*, 156–167. https://doi.org/10.1016/j.jdec.2025.07.005

StatisticsTimes.com. (2024, April 30). *Assam population 2024*. https://statisticstimes.com/demographics/india/assam-population.php

Tewari, D. B. (n.d.). *Urban hierarchy, primate city and rank size rule*. In *Urban geography*. e-PG Pathshala, UGC–INFLIBNET Centre.

V. V. Giri National Labour Institute. (2025). *Gig and platform workers: Vision 2047* (NLI Research Studies Series No. 173/2025). V. V. Giri National Labour Institute. https://www.scribd.com/document/959150940/Dhanya-MB

Zelma, E. (2024). The gig economy from the freelancer's perspective: The risk of precarianization. *International Journal of Contemporary Management, 60*(1), 211–226.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Addison-Wesley Press.