The Heterogeneous Impact of COVID-19 Lockdowns: Evidence from India's Scheduled Castes and Tribes

Dhananjay Balakrishnan
dhananjb@stanford.edu

Vedant Sahu
vsahu@stanford.edu

1 Introduction

The COVID-19 pandemic created an unprecedented health crisis, with government mandated lockdowns and supply chain disruptions causing economic shocks, further exacerbating people's hardships. Its impact was particularly severe in developing countries such as India, where GDP contracted 7.3% in the fiscal year 2020-21 [Deshpande and Bhardwaj, 2023]. Natural shocks often have disproportionate impacts on marginalized populations. In India's context, people belonging to Scheduled Castes and Scheduled Tribes (SC/ST), historically disadvantaged groups with constitutional protections, face structural vulnerabilities that amplify their exposure to such crises [Deshpande and Ramachandran, 2023]. This project exploits a natural experiment in India's National Family Health Survey (NFHS-5), which surveyed different households before versus during the pandemic, to estimate whether the COVID-19 pandemic had differential impacts across caste groups. Using a difference-in-differences framework, we will examine how the pandemic's economic shocks were distributed across India's socioeconomic and caste hierarchy. See Appendix A for related work.

2 Data

We propose using the Demographic and Health Surveys (DHS) data for India, also referred to as the National Family Health Survey (NFHS) [The Demographic and Health Surveys Program, 2025]. The National Family Health Survey (NFHS) is a large-scale, multi-round survey conducted in a representative sample of households throughout India, which provides information on various factors covering child mortality, family planning, maternal and child health, reproductive health, nutrition, and quality of health services [Ministry of Health & Family Welfare, Government of India, 2025]. In particular, we plan on using the NFHS-4 and NFHS-5 data, which correspond to nationwide surveys conducted in 2014-15 and 2019-21 (split into Phase 1 and 2) respectively. See Appendix B for details.

3 Research Question

This project addresses two interrelated questions about the causal impact of COVID-19 lockdowns in India:

- **Identification:** Can pre-pandemic trends from NFHS-4 (2015-16) to NFHS-5 Phase 1 (2019-20) be used to construct valid counterfactual outcomes for houselholds surveyed only in NFHS-5 Phase 2 (2020-21), thereby isolating the causal effect of COVID-19 lockdowns?
- **Heterogenous Treatment Effects:** Did SC/ST populations experience significantly larger pandemic-induced declines in socio-economic and health outcomes compared to privileged caste groups, as identified using these counterfactual estimates?

4 Proposed Methodology

The NFHS-5 survey was conducted in two phases: Phase 1 (pre-pandemic) and Phase 2 (post-pandemic). Let $s \in S$ index districts, where $S = S_1 \cup S_2$ with S_1 denoting Phase 1 districts and S_2 denoting Phase 2 districts. Directly comparing these two sets is problematic because districts differ substantially in social structure, political context, and local policies. To enable valid causal inference, we consider districts as the unit of analysis and aggregate household-level indicators to the district level.

We propose two complementary identification strategies:

- Strategy 1: Matched Pair Design
- Strategy 2: Two-Step Counterfactual Construction

Our primary analysis will use difference-in-differences to estimate lockdown effects on the difference in outcomes. Time permitting, we will conduct additional analyses including event studies to examine dynamic treatment effects. The NFHS survey's comprehensive questionnaire provides multiple relevant outcomes for analysis. See Appendix C for details.

References

T. Banerjee and T. Mishra. Covid pandemic and spousal violence against indian women: An exploration of nfhs-4 (2015–2016) and nfhs-5 (2019–2021) data sets. *Millennial Asia*, 2023. doi: 10.1177/09763996231169702.

T. Banerjee and T. Mishra. Sexual and reproductive health of indian women, intimate partner violence and covid-19 pandemic: an analysis using nfhs-iv and nfhs-v data sets. *SN Social Sciences*, 4, 2024. doi: 10.1007/s43545-024-00960-3.

A. Deshpande and A. Bhardwaj. Here is what nfhs-5 tells us about india. Centre for Economic Data and Analysis (CEDA), Ashoka University, 2021. URL `https://ceda.ashoka.edu.in/here-is-what-nfhs-5-tells-us-about-india/`.

A. Deshpande and A. Bhardwaj. Covid-19 and the indian economy: Constructing roadmaps to an inclusive recovery. In *Global Perspectives of COVID-19 Pandemic on Health, Education, and Role of Media*, pages 99–146. Springer, Singapore, 2023.

A. Deshpande and R. Ramachandran. Covid-19 and caste inequalities in india: The critical role of social identity in pandemic-induced job losses. *American Journal of Agricultural Economics*, 2023. doi: 10.1002/aepp.13384. First published online: 07 June 2023.

Ministry of Health & Family Welfare, Government of India. National family health survey, 2025. URL `https://www.nfhsiips.in/nfhsuser/index.php`. Accessed: 2025-11-13.

The Demographic and Health Surveys Program. The dhs program, 2025. URL `https://dhsprogram.com/`. Accessed: 2025-11-13.

A Related Work

A small but growing literature exploits the NFHS-5's unique two-phase survey design to study COVID-19's impacts in India. Banerjee and Mishra [2023] examine spousal violence against Indian women using both NFHS-4 and NFHS-5 datasets, exploiting the fact that Phase 1 states were surveyed before the pandemic while Phase 2 states were surveyed during it. In related work, Banerjee and Mishra [2024] extend this analysis to examine the structural relationships between intimate partner violence, women's empowerment, and sexual and reproductive health outcomes during COVID-19.

Beyond formal causal studies, several analyses provide descriptive insights into NFHS-5 patterns. Deshpande and Bhardwaj [2021] at the Centre for Economic Data and Analysis note a puzzling anomaly: Phase 1 states (surveyed pre-COVID) showed worsening malnutrition indicators, while Phase 2 states (surveyed during COVID) showed improvements in stunting. They raise important questions about whether the pandemic affected survey quality or whether Phase 2 states happened to be better performers by chance.

B Data Details

Due to the onset of the COVID-19 pandemic and the subsequent government lockdowns, the survey collection for NFHS-5 had to be split up into two phases. Phase 1, which was conducted pre-pandemic (June 2019 to March 2020), covered 22 states & Union Territories (UTs) and included 439,306 households. Meanwhile, Phase 2, which was conducted post-pandemic (November 2020 to May 2021), covered 14 states & UTs and included 197,393 households. While the NFHS-4 data gives us a certain baseline for how different groups were doing in that time-period through the measured socio-economic and health indicators, the NFHS-5 data being split into two parts: one pre-COVID and one post-COVID gives us a clean and intuitive way of isolating the effect due to the pandemic shock on various groups of people.

C Methodology Details

Strategy 1: Matched Pair Design. Using Phase 1 districts as controls and Phase 2 districts as treated units (exposed to the pandemic shock), we match districts on pre-pandemic socio-economic and demographic characteristics. For each treated district, we select the closest control match, forming matched pairs that approximate a natural experiment. Estimating treatment effects then reduces to a classical matched-pair comparison of pre- and post-shock outcomes.

Strategy 2: Two-Step Counterfactual Construction. We use NFHS-4 (2015–16) as baseline data for all districts. The temporal shift from NFHS-4 to NFHS-5 Phase 1 for districts $s \in S_1$ provides an empirical mapping of pre-pandemic evolution. We apply this shift to NFHS-4 values for districts $s \in S_2$ to construct counterfactual pre-pandemic outcomes for districts not surveyed in Phase 1. This yields a synthetic pre-pandemic profile for S_2, after which standard treatment–control comparisons become feasible.

Potential choices of outcomes to measure include:

- **Health & Morbidity:** illness prevalence, chronic disease indicators, access to treatment.
- **Maternal & Reproductive Health:** antenatal care (ANC) visits, institutional delivery, postnatal care, contraceptive use.
- **Child Health & Nutrition:** stunting/wasting, immunization coverage, breastfeeding practices, child anemia.
- **Economic & Welfare Indicators:** employment status, wealth index, food insecurity, use of public distribution systems (PDS).
- **Health Access & Infrastructure:** distance to facilities, barriers to care, transport availability.
- **Sanitation & Environment:** access to clean drinking water, improved sanitation, clean cooking fuel.

Based on some exploratory analysis, time periods of the shock, and background research about outcomes that are likely to be affected on such a short timeline (1 year), we will choose an outcome variable that fits our modeling exercise.