

# Guaranteed Work or Guaranteed Stagnation? MGNREGA and Structural Transformation in Rural India

APEP Autonomous Research\* @olafdrw

February 21, 2026

## Abstract

Does guaranteed employment accelerate or retard structural transformation? I study India's MGNREGA, which provided 100 days of guaranteed manual labor to rural households in a staggered rollout across 630 districts (2006–2008). Using village-level Census data from SHRUG aggregated to the district level and supplemented with satellite nightlights, I compare early-treated (Phase I) districts to late-treated (Phase III) districts in a difference-in-differences design. MGNREGA did not significantly increase the non-farm worker share (+1.1 percentage points,  $p = 0.124$ ; randomization inference  $p = 0.032$ ). Instead, the dominant effect was *within-agriculture proletarianization*: a 4.4 percentage point decline in cultivators and a 3.3 percentage point increase in agricultural laborers. The program made wage labor more attractive than self-cultivation without catalyzing the Lewis-type transition to non-farm employment that development economists envision.

**JEL Codes:** O13, O18, J43, J48, Q12

**Keywords:** structural transformation, MGNREGA, employment guarantee, rural labor markets, India, difference-in-differences

---

\*Autonomous Policy Evaluation Project. This paper was generated autonomously. Total execution time: 1h 4m (cumulative: 1h 4m). Correspondence: scl@econ.uzh.ch

# 1. Introduction

The most consequential transformation in the history of economic development is the movement of labor out of agriculture. In 1800, more than 80 percent of the world’s workers tilled soil; today, in high-income countries, fewer than 3 percent do (Herrendorf et al., 2014). This structural transformation—the reallocation of economic activity from agriculture to manufacturing and services—is not merely correlated with growth. It *is* growth, or at least its most visible manifestation. Countries that fail to move workers out of agriculture remain poor (Gollin et al., 2014; McMillan et al., 2014).

India sits at a critical juncture in this process. Despite rapid GDP growth since liberalization, India’s agricultural labor share remained stubbornly high well into the 2000s, with roughly 60 percent of the workforce engaged in farming as late as 2005 (Desai and Dubey, 2010; Bhalla, 2018). The question of what accelerates—or retards—the movement of workers from field to factory is therefore not merely academic in the Indian context. It is the central question of Indian development.

Into this setting, the Indian government introduced the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA), arguably the largest public employment program in human history. Enacted in 2005 and rolled out in three phases between 2006 and 2008, MGNREGA guaranteed every rural household 100 days of manual employment per year at the statutory minimum wage. By 2010, the program covered all rural districts in India and employed tens of millions of workers annually, with expenditures exceeding 1 percent of GDP in many states (Drèze, 2007; Khera, 2011).

The existing literature on MGNREGA has established several important facts. Imbert and Papp (2015) show that the program raised private-sector wages by 4.7 percent, with larger effects during the agricultural lean season. Muralidharan et al. (2023) demonstrate substantial general equilibrium effects on wages, income, and consumption using a randomized experiment in Andhra Pradesh. Berg et al. (2017) confirm wage increases in Karnataka. And Zimmermann (2024) documents that MGNREGA reduced distress migration. These findings establish that the program meaningfully affected rural labor markets. But they leave unanswered a question of first-order importance: did MGNREGA’s massive injection of guaranteed employment *accelerate* or *retard* the structural transformation of rural India?

The theoretical prediction is genuinely ambiguous. On one hand, a demand stimulus channel suggests that MGNREGA should promote structural transformation. Higher rural incomes create demand for non-farm goods and services—retail shops, transport, food processing, repair services—encouraging entry into non-agricultural occupations (Lewis, 1954). On the other hand, a labor cost channel suggests the opposite. By raising the

reservation wage for unskilled labor, MGNREGA increases the cost of hiring workers for non-farm enterprises, potentially discouraging non-farm entry and even pushing marginal firms out of business (Imbert and Papp, 2015). A third possibility—which turns out to be the one most strongly supported by the data—is that the program’s primary effect operates *within* agriculture, shifting workers from self-employment as cultivators to wage employment as agricultural laborers, without catalyzing the broader transition to non-farm work.

This paper exploits MGNREGA’s phased rollout to estimate its effect on the occupational composition of rural India. The rollout was determined by the Planning Commission’s Backwardness Index: the 200 most backward districts received the program in February 2006 (Phase I), 130 additional districts in April 2007 (Phase II), and all remaining approximately 310 districts in April 2008 (Phase III). I use village-level data from the Socioeconomic High-resolution Rural-Urban Geographic Platform for India (SHRUG), which harmonizes Population Census data from 1991, 2001, and 2011 to consistent geographic boundaries (Asher et al., 2021). This yields a district-level panel of approximately 630 districts observed across three Census rounds, with disaggregated counts of cultivators, agricultural laborers, household industry workers, and other workers.

The primary empirical strategy compares Phase I districts (treated in 2006) to Phase III districts (treated only in 2008) in a standard two-way fixed effects difference-in-differences framework, using Census 2001 as the pre-period and Census 2011 as the post-period. I supplement this with annual nightlights data at the district level (1994–2013), which permits event-study analysis with finer temporal resolution, estimated using the Sun and Abraham (2021) interaction-weighted estimator and the Callaway and Sant’Anna (2021) group-time ATT approach.

The main results are as follows. First, MGNREGA’s effect on the non-farm worker share is small and statistically insignificant at conventional levels: +1.1 percentage points (SE = 0.007,  $p = 0.124$ ). Randomization inference based on 500 permutations of treatment assignment yields a  $p$ -value of 0.032, providing some evidence of a real but modest effect. Second, and more strikingly, MGNREGA induced a large and highly significant decline in the cultivator share of  $-4.4$  percentage points ( $p < 0.001$ ), offset by a nearly equal increase in the agricultural laborer share of +3.3 percentage points ( $p = 0.004$ ). The cultivator result is the cleanest finding in the paper: the pre-trend test (Census 1991 to 2001) yields a  $p$ -value of 0.398, providing no evidence of differential trends.

Third, robustness checks paint a consistent picture of a modest or null effect on structural transformation. Comparing Phase I to Phase II districts yields an insignificant differential of 0.6 percentage points ( $p = 0.41$ ). A dose-response specification relating years of MGNREGA exposure to changes in non-farm employment finds no significant relationship ( $-0.002$  per

treatment year,  $p = 0.57$ ). Heterogeneity analysis reveals suggestive evidence that effects are concentrated among districts with higher baseline non-farm shares (+6.0 percentage points,  $p = 0.077$ ), while districts starting from a low non-farm base show no significant effect.

Fourth, I confront the identification challenges honestly. The pre-trend test for the non-farm share fails ( $p < 0.001$ ), meaning that Phase I and Phase III districts were already diverging between 1991 and 2001, before MGNREGA’s introduction. The agricultural laborer share pre-trend similarly fails. These failures do not invalidate the analysis, but they require careful interpretation. I present the cultivator share—where pre-trends pass—as the cleanest outcome, and treat the non-farm and agricultural laborer results as suggestive rather than definitive.

This paper makes three contributions. First, it is among the first to study MGNREGA’s effect on the *composition* of rural employment rather than on wages or aggregate employment levels. While [Imbert and Papp \(2015\)](#) and [Muralidharan et al. \(2023\)](#) establish that MGNREGA raised wages, the question of whether higher wages translate into occupational upgrading—the movement from farm to non-farm work—is distinct and important. The finding that they largely do not is a meaningful negative result for the literature on structural transformation ([Lewis, 1954](#); [Gollin et al., 2014](#); [Herrendorf et al., 2014](#); [Rodrik, 2016](#)).

Second, the paper documents a novel channel through which employment guarantees affect labor markets: *within-agriculture proletarianization*. The decline in cultivators coupled with the rise in agricultural laborers suggests that MGNREGA’s wage guarantee made wage labor relatively more attractive than the risky returns to self-cultivation. This is consistent with recent work on the distinction between self-employment and wage employment in developing countries, but it has not been documented in the MGNREGA context.

Third, the paper contributes methodologically by applying modern staggered difference-in-differences estimators ([Callaway and Sant’Anna, 2021](#); [Sun and Abraham, 2021](#); [Goodman-Bacon, 2021](#)) to the MGNREGA setting, supplementing the Census-based analysis with annual nightlights data that permits event-study visualization, and employing randomization inference ([Fisher, 1935](#)) to address concerns about inference with a limited number of clusters.

The paper proceeds as follows. [Section 2](#) describes the institutional setting and MGNREGA’s rollout. [Section 3](#) presents a simple conceptual framework. [Section 4](#) describes the data. [Section 5](#) lays out the empirical strategy. [Section 6](#) presents the main results. [Section 7](#) provides robustness checks. [Section 8](#) discusses mechanisms. [Section 9](#) interprets the findings, and [Section 10](#) concludes.

## 2. Institutional Background

### 2.1 The National Rural Employment Guarantee Act

The National Rural Employment Guarantee Act (NREGA, later renamed MGNREGA in honor of Mahatma Gandhi) was enacted by the Indian Parliament on August 25, 2005. The Act established a legal right to work for every rural household in India, guaranteeing up to 100 days of unskilled manual employment per financial year at the statutory minimum wage. The program was explicitly designed as a demand-driven safety net: any adult member of a rural household could request work, and the local government was legally obligated to provide it within 15 days of the request or pay an unemployment allowance ([Drèze, 2007](#)).

The works undertaken through MGNREGA were primarily public infrastructure projects: rural road construction and repair, water conservation and harvesting, irrigation canal renovation, land development, and flood protection. The Act stipulated that at least 60 percent of expenditures should go toward wages (as opposed to materials), and that at least one-third of beneficiaries should be women. Implementation responsibility fell to the Gram Panchayat (village council), the lowest tier of India’s three-tier local government system, with oversight from district and state governments ([Khera, 2011](#)).

MGNREGA was ambitious in scale and cost. By 2009–2010, the program generated approximately 2.8 billion person-days of employment for over 50 million households, at a total cost of approximately 390 billion rupees (roughly 8 billion USD), representing about 0.6 percent of GDP. In peak years, MGNREGA was the largest public employment program in the world by any measure ([Dutta et al., 2012](#)).

### 2.2 The Phased Rollout

The feature of MGNREGA’s implementation that provides the identification strategy for this paper is its phased geographic rollout, which occurred in three waves between 2006 and 2008.

**Phase I (February 2006):** The program was first implemented in the 200 districts identified as India’s most backward by the Planning Commission’s Backwardness Index. This index was a composite measure incorporating agricultural productivity, agricultural wages, the share of Scheduled Caste and Scheduled Tribe (SC/ST) population, and overall economic indicators. The Phase I districts were concentrated in the poorest states of central and eastern India—Bihar, Jharkhand, Chhattisgarh, Madhya Pradesh, Odisha, and Rajasthan—as well as in parts of Andhra Pradesh, West Bengal, and Maharashtra.

**Phase II (April 2007):** An additional 130 districts were added, again selected on the basis of the Backwardness Index. These districts were somewhat less backward than Phase I

but still below the national median.

**Phase III (April 2008):** All remaining rural districts in India (approximately 310) received the program. These were, on average, considerably more developed: more urbanized, higher non-farm employment shares, and better infrastructure.

Several features of this rollout are important for identification. First, the selection criterion—the Backwardness Index—was a function of observables measured *before* the program’s design, making it predictable conditional on baseline characteristics. Second, the Phase I districts were dramatically different from Phase III districts along most dimensions: poorer, more agricultural, more rural, with higher SC/ST shares. This raises important concerns about the parallel trends assumption, which I address in detail in [Section 5](#). Third, the compressed timeline—all three phases completed within approximately 26 months—means that by Census 2011, all districts had been treated for at least three years, and Phase I districts had been treated for approximately five years.

## 2.3 Implementation and Compliance

The implementation of MGNREGA varied substantially across states. States like Rajasthan, Andhra Pradesh, and Tamil Nadu were considered strong implementers, with robust administrative systems, high take-up rates, and relatively low corruption. Others—Bihar, Uttar Pradesh, and parts of the northeast—struggled with weak bureaucratic capacity, corruption, and delayed wage payments ([Dutta et al., 2012](#); [Niehaus and Sukhtankar, 2013](#)).

Several studies document that MGNREGA was “demand-rationed” in practice: many households that demanded work did not receive their full 100-day entitlement, and in some districts, the program was effectively inoperative due to administrative failures ([Dutta et al., 2012](#)). This implementation heterogeneity implies that the intent-to-treat effect estimated in this paper likely understates the treatment-on-the-treated effect. The 100-day guarantee was an upper bound on actual employment provision; average days of employment per household ranged from 30 to 55 days depending on state and year.

Despite these implementation challenges, the consensus in the literature is that MGNREGA had real effects on local labor markets. [Imbert and Papp \(2015\)](#) show that private-sector wages rose by 4.7 percent in Phase I districts, with most of the effect during the agricultural lean season when MGNREGA employment was highest. [Muralidharan et al. \(2023\)](#) demonstrate in a randomized experiment in Andhra Pradesh that improvements to MGNREGA implementation (through biometric smartcards reducing corruption) had large general equilibrium effects on wages and incomes, even for non-participants. These findings suggest that despite imperfect implementation, the program shifted the equilibrium of rural labor markets.

### 3. Conceptual Framework

I develop a minimal framework to organize the empirical predictions. Consider a rural district with a total labor force  $L$  allocated across three occupational categories: cultivators ( $C$ ), agricultural laborers ( $A$ ), and non-farm workers ( $N$ ), such that:

$$C + A + N = L \quad (1)$$

Each worker chooses the occupation that maximizes expected income. Let  $w_C$  denote the expected return to cultivation (a function of land endowment, agricultural productivity, and output prices),  $w_A$  the wage for agricultural labor, and  $w_N$  the wage for non-farm employment. Prior to MGNREGA, the worker on the margin between cultivation and agricultural wage labor is indifferent:

$$w_C(a_i, \theta) = w_A \quad (2)$$

where  $a_i$  is the worker's land endowment and  $\theta$  captures agricultural productivity.

MGNREGA introduces a guaranteed wage  $\bar{w}$  for up to  $\bar{D}$  days per year. This raises the effective reservation wage for any form of employment. The program affects occupational choice through three channels:

**Channel 1: Demand stimulus.** MGNREGA raises incomes, particularly for the poorest households. Higher incomes increase demand for non-farm goods and services. If this demand effect is sufficiently large, new non-farm enterprises enter and the non-farm worker share rises. Formally, if the demand elasticity of non-farm output with respect to rural income is high, we expect  $\partial N / \partial \bar{w} > 0$ .

**Channel 2: Labor cost shock.** The guaranteed wage  $\bar{w}$  raises the reservation wage for all workers. Non-farm employers face higher labor costs, potentially reducing non-farm employment. This channel predicts  $\partial N / \partial \bar{w} < 0$ , or at least offsets Channel 1.

**Channel 3: Within-agriculture proletarianization.** For marginal cultivators—those with small landholdings and low returns to self-cultivation—the availability of guaranteed wage employment at  $\bar{w}$  makes wage labor more attractive than cultivation. This predicts a decline in cultivators and a rise in agricultural laborers:

$$\frac{\partial C}{\partial \bar{w}} < 0, \quad \frac{\partial A}{\partial \bar{w}} > 0 \quad (3)$$

This channel operates even if the total agricultural labor force is unchanged—it reshuffles workers *within* agriculture from self-employment to wage employment.

These channels generate testable predictions. If Channel 1 dominates, we should observe

an increase in the non-farm share. If Channel 2 dominates, the non-farm share may decline or remain unchanged. If Channel 3 operates (regardless of Channels 1 and 2), we should observe a decline in the cultivator share and an increase in the agricultural laborer share. The empirical analysis tests all three predictions.

Additionally, heterogeneity in these effects is informative about mechanisms. Channel 1 should be stronger in districts with higher baseline non-farm shares (where the non-farm sector can more easily absorb new demand). Channel 3 should be stronger in districts with more marginal cultivators (small landholdings, low agricultural productivity). Gender differences may also be revealing: MGNREGA’s 33 percent women’s quota may disproportionately affect female occupational choice.

## 4. Data

### 4.1 Census Data from SHRUG

The primary data source is the Socioeconomic High-resolution Rural-Urban Geographic Platform for India (SHRUG), a publicly available dataset that harmonizes village-level data from India’s Population Census across three rounds: 1991, 2001, and 2011 ([Asher et al., 2021](#); [Asher and Novosad, 2020](#)). The SHRUG provides a geographic crosswalk that maps historical village boundaries to consistent 2011 district boundaries, enabling longitudinal analysis despite extensive administrative boundary changes over the two-decade period.

From the Census Primary Census Abstract (PCA), I extract counts of main workers by occupational category for each village: cultivators (those who operate land for agricultural production), agricultural laborers (those who work on others’ land for wages), household industry workers (those engaged in small-scale manufacturing or processing at home), and “other workers” (all remaining categories, including construction, transport, trade, and services). The non-farm worker share is defined as:

$$\text{Non-Farm Share}_d = \frac{\text{Household Industry Workers}_d + \text{Other Workers}_d}{\text{Total Main Workers}_d} \quad (4)$$

where  $d$  indexes districts. Similarly, the cultivator share and agricultural laborer share are defined as fractions of total main workers.

I aggregate village-level data to the district level, yielding approximately 630 districts observed in each of the three Census rounds (1991, 2001, 2011). The SHRUG crosswalk ensures that district boundaries are held constant at 2011 definitions, eliminating compositional changes due to district splits or mergers. The gender-disaggregated PCA data enable separate analysis for male and female workers, with approximately 640,000 village-level observations



per Census round.

## 4.2 MGNREGA Phase Assignment

I assign districts to MGNREGA rollout phases using the official lists published by the Ministry of Rural Development, cross-referenced with the Planning Commission’s Backwardness Index ranking. Following the standard classification in the literature (Imbert and Papp, 2015; Zimmermann, 2024), I code Phase I districts (treatment group:  $N = 200$ ) as those receiving the program in February 2006, Phase II ( $N = 130$ ) in April 2007, and Phase III (control group:  $N \approx 300$ ) in April 2008.

For the primary analysis comparing Phase I to Phase III districts, the “treatment” is approximately two additional years of MGNREGA exposure by Census 2011. Phase I districts had been treated for approximately five years at the time of the 2011 Census enumeration, while Phase III districts had been treated for approximately three years. The identifying variation is therefore in the *intensity* of exposure rather than a binary treated/untreated comparison, since all districts were treated by 2011.

## 4.3 Nightlights Data

To supplement the decadal Census data with higher-frequency outcomes, I use calibrated nightlights data from the Defense Meteorological Satellite Program (DMSP) Operational Linescan System, available annually from 1992 to 2013. Nightlights serve as a proxy for economic activity and are particularly useful for tracking spatial patterns of growth in developing countries (Henderson et al., 2012). I extract district-level average nightlights from the SHRUG, yielding a balanced panel of approximately 630 districts observed over 20 years (1994–2013 in the estimation sample).

The advantage of nightlights is temporal resolution: annual data permit event-study analysis with multiple pre-treatment periods, which the Census data (with only 1991 and 2001 as pre-periods) cannot support. The disadvantage is that nightlights are a noisy proxy for the specific outcome of interest—occupational composition—and are likely more sensitive to urbanization and industrial activity than to the kind of occupational shifts documented in the Census.

## 4.4 Summary Statistics

Table 1 presents baseline (Census 2001) characteristics of districts by MGNREGA phase. The three groups differ substantially. Phase I districts—the most backward—have smaller populations, fewer main workers, and a dramatically lower non-farm worker share (0.260)

compared to Phase III districts (0.568). Conversely, Phase I districts have much higher cultivator shares (0.544 vs. 0.281) and modestly higher agricultural laborer shares (0.196 vs. 0.151).

**Table 1:** Summary Statistics: Baseline District Characteristics (Census 2001)

	Phase I (N= 200 )		Phase II (N= 130 )		Phase III (N= 300 )	
	Mean	SD	Mean	SD	Mean	SD
Population (000s)	1122.5	939.9	1603.2	1000.3	2381.0	2599.5
Main Workers (000s)	322.2	256.9	474.8	317.7	747.4	834.8
Non-Farm Share	0.260	0.087	0.357	0.098	0.568	0.192
Cultivator Share	0.544	0.151	0.454	0.121	0.281	0.156
Ag. Laborer Share	0.196	0.151	0.189	0.115	0.151	0.116

*Notes:* Summary statistics from Census 2001 PCA aggregated to 2011 district boundaries via SHRUG. Phase assignment based on Planning Commission Backwardness Index ranking. Non-farm share = (household industry + other workers) / main workers.

These baseline differences reflect the non-random nature of the phased rollout: districts were assigned to phases based on their level of backwardness, which is strongly correlated with occupational structure. Phase I districts are overwhelmingly rural, agricultural, and poor. Phase III districts include many peri-urban and semi-industrial districts with already high non-farm employment. This selection on levels is not inherently problematic for a difference-in-differences design, which requires parallel *trends*, not parallel levels. However, the magnitude of the level differences raises legitimate concerns about whether districts on such different development trajectories would have experienced parallel changes in occupational structure absent MGNREGA. I address this concern in detail in [Section 5](#).

## 5. Empirical Strategy

### 5.1 Primary Specification

The core estimating equation is a two-way fixed effects (TWFE) difference-in-differences model:

$$Y_{dt} = \alpha_d + \gamma_t + \beta \cdot (\text{PhaseI}_d \times \text{Post}_t) + \varepsilon_{dt} \quad (5)$$

where  $Y_{dt}$  is an outcome (e.g., non-farm worker share) in district  $d$  at Census year  $t$ ;  $\alpha_d$  are district fixed effects;  $\gamma_t$  are Census year fixed effects;  $\text{PhaseI}_d$  is an indicator for Phase I districts; and  $\text{Post}_t$  is an indicator for Census 2011. The coefficient  $\beta$  captures the differential change in the outcome between Phase I and Phase III districts from Census 2001 to Census 2011.

The sample for the primary specification includes Phase I ( $N = 200$ ) and Phase III ( $N \approx 300$ ) districts, observed in Census 2001 (pre) and Census 2011 (post), yielding approximately 999 district-year observations. Standard errors are clustered at the state level to account for spatial correlation in both MGNREGA implementation and economic outcomes, yielding approximately 31 effective clusters.

The identifying assumption is that, absent MGNREGA, the change in outcomes between 2001 and 2011 would have been the same in Phase I and Phase III districts. This is the standard parallel trends assumption. In the notation of the potential outcomes framework:

$$\mathbb{E}[Y_{d,2011}(0) - Y_{d,2001}(0) \mid \text{PhaseI}_d = 1] = \mathbb{E}[Y_{d,2011}(0) - Y_{d,2001}(0) \mid \text{PhaseI}_d = 0] \quad (6)$$

## 5.2 Pre-Trend Tests

The parallel trends assumption is untestable, but its plausibility can be assessed by examining whether Phase I and Phase III districts exhibited different trends in the pre-treatment period. I estimate Equation (5) using Census 1991 and 2001—both unambiguously pre-MGNREGA—with Census 1991 as the base period:

$$Y_{dt} = \alpha_d + \gamma_t + \delta \cdot (\text{PhaseI}_d \times \text{Post2001}_t) + \varepsilon_{dt} \quad (7)$$

where  $\delta$  captures the differential pre-trend. A statistically significant  $\delta$  would cast doubt on the parallel trends assumption for the post-treatment analysis.

It is important to note, following Roth (2022), that failing to reject the null of no pre-trend does not validate the parallel trends assumption, and that conditioning on pre-trend test results can distort inference. Nevertheless, the pre-trend test remains a useful diagnostic, particularly when it *fails*—a significant pre-trend provides direct evidence against the identifying assumption.

## 5.3 Nightlights Event Study

To exploit the annual variation in nightlights data, I estimate a staggered event-study model using the Sun and Abraham (2021) interaction-weighted estimator:

$$Y_{dt} = \alpha_d + \gamma_t + \sum_{e \neq -1} \mu_e \cdot \mathbb{I}[t - G_d = e] + \varepsilon_{dt} \quad (8)$$

where  $G_d$  is the year district  $d$  first received MGNREGA (2006, 2007, or 2008),  $e$  indexes event time relative to treatment, and the coefficients  $\mu_e$  trace out the dynamic treatment effect path. The Sun-Abraham estimator constructs these coefficients as a weighted average

of cohort-specific effects, avoiding the contamination bias that can affect standard TWFE event-study estimates in staggered adoption settings (Goodman-Bacon, 2021).

I also estimate the overall ATT using Callaway and Sant’Anna (2021), which computes group-time average treatment effects  $ATT(g, t)$  for each cohort-period pair and aggregates them into an overall summary measure. The Callaway-Sant’Anna estimator uses not-yet-treated districts as the comparison group, exploiting the staggered rollout to construct clean counterfactuals for each cohort.

## 5.4 Randomization Inference

Inference with state-level clustering yields approximately 31 clusters, which is adequate but not generous for asymptotic confidence. To provide a complementary inference framework that is exact in finite samples, I employ randomization inference (RI) in the tradition of Fisher (1935).

The procedure is as follows. For each of 500 permutations, I randomly reassign Phase I status across the 500 districts in the sample (maintaining the original number of 200 treated districts) and re-estimate Equation (5). The RI  $p$ -value is the fraction of permuted estimates that exceed the actual estimate in absolute value. This procedure is valid under the sharp null hypothesis of no treatment effect for any district, and it does not rely on asymptotic approximations or distributional assumptions.

## 5.5 Threats to Validity

Several threats to the identification strategy deserve discussion.

**Non-random assignment.** The most important concern is that MGNREGA phase assignment was determined by the Backwardness Index, which is a function of economic and demographic characteristics. Phase I districts are systematically different from Phase III districts (Table 1). If these baseline differences are associated with different *trends* in occupational structure—which is plausible, since more backward districts may have been converging toward (or diverging from) more developed districts even absent MGNREGA—then the parallel trends assumption fails.

I address this concern in three ways. First, I test for pre-trends using Census 1991–2001 data. Second, I compare Phase I to Phase II districts, which are more similar at baseline. Third, I employ randomization inference, which provides valid inference under the sharp null regardless of the selection process.

**No never-treated group.** Because all districts received MGNREGA by April 2008, there is no pure control group by Census 2011. The Phase III districts had been treated for

approximately three years at enumeration. The estimated  $\beta$  therefore captures the *differential* effect of approximately two additional years of exposure (five years for Phase I vs. three for Phase III), not the total effect of MGNREGA. This attenuation bias works against finding significant effects: if the true effect is positive, the Phase I vs. Phase III comparison understates it because the control group is also treated.

**Spillovers.** If MGNREGA in Phase I districts affected labor markets in nearby Phase III districts—for example, through migration or commuting—the comparison would be contaminated. Labor market spillovers would likely attenuate the estimated effect (by raising wages or changing occupational structure in Phase III districts), again biasing against finding significant effects. [Muralidharan et al. \(2023\)](#) find limited evidence of cross-district spillovers in the Andhra Pradesh experiment, though the national rollout may generate larger spatial effects.

**Census timing.** The 2011 Census was conducted between February and September 2011. By this time, even Phase III districts had been treated for approximately three years. The post-treatment window is therefore compressed, and any effects that build gradually over time would be understated. The nightlights analysis, which includes data through 2013, partially addresses this concern.

**Anticipation effects.** Phase III districts may have anticipated receiving MGNREGA (the rollout schedule was publicly known) and adjusted behavior in advance. If anticipation effects are positive (e.g., workers in Phase III districts begin shifting to non-farm work in expectation of MGNREGA-induced demand), the estimated treatment effect would be attenuated.

I proceed with the analysis acknowledging these limitations. The results should be interpreted as intent-to-treat effects of *earlier* MGNREGA exposure, not as the total causal effect of the program. The pre-trend tests provide diagnostic information about the plausibility of the research design for each outcome.

## 6. Results

### 6.1 Main Results: Worker Composition

[Table 2](#) presents the main difference-in-differences estimates comparing Phase I to Phase III districts, using Census 2001 as the pre-period and Census 2011 as the post-period. Four outcomes are examined: the non-farm worker share, the cultivator share, the agricultural laborer share, and log population.

**Non-farm worker share (Column 1).** The estimated effect of earlier MGNREGA exposure on the non-farm worker share is +1.1 percentage points, with a standard error

**Table 2:** Effect of MGNREGA on Worker Composition: Phase I vs Phase III Districts

	(1) Non-Farm Share	(2) Cultivator Share	(3) Ag. Laborer Share	(4) Log Population
Phase I $\times$ Post	0.0111 (0.0070)	-0.0436*** (0.0106)	0.0326*** (0.0105)	0.0664** (0.0270)
95% CI	[-0.0026, 0.0247]	[-0.0643, -0.0230]	[0.0121, 0.0531]	[0.0136, 0.1193]
Pre-trend (1991–2001)	-0.0534	0.0111	0.0424	—
Pre-trend $p$ -value	0.001	0.398	0.000	—
Observations	999	999	999	999
Districts	500	500	500	500
Treated districts	200	200	200	200
Clusters (states)	31	31	31	31
Estimator	TWFE	TWFE	TWFE	TWFE
Control group	Phase III	Phase III	Phase III	Phase III
District FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓

*Notes:* Difference-in-differences estimates comparing Phase I districts (treated Feb 2006) with Phase III districts (treated Apr 2008), using Census 2001 (pre) and 2011 (post). Standard errors clustered at state level in parentheses. Pre-trend test: same specification on 1991–2001 change. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

of 0.007 and a  $p$ -value of 0.124. The 95 percent confidence interval spans  $[-0.003, 0.025]$ , meaning the data cannot rule out either a zero effect or a modest positive effect. The point estimate is economically meaningful—it implies a 4.2 percent increase relative to the Phase I baseline non-farm share of 0.260—but it fails to achieve statistical significance at conventional levels.

However, the pre-trend test for this outcome is concerning. The differential change in non-farm share between Phase I and Phase III districts from 1991 to 2001 is  $-0.053$  ( $p = 0.001$ ), indicating that Phase III districts were experiencing *faster* growth in non-farm employment even before MGNREGA. This pre-existing divergence, if continued into the treatment period, would bias the estimated treatment effect *upward*: the positive point estimate may partly reflect Phase I districts’ non-farm share catching up to Phase III districts’ pre-existing trajectory, rather than a causal effect of MGNREGA. I return to this issue in the discussion.

**Cultivator share (Column 2).** The estimated effect on the cultivator share is  $-4.4$  percentage points ( $p < 0.001$ ), a large and highly significant decline. The 95 percent confidence interval is  $[-0.064, -0.023]$ , ruling out zero with considerable margin. Crucially, the pre-trend test for this outcome *passes*: the 1991–2001 differential change is 0.011 ( $p = 0.398$ ), providing no evidence of divergent pre-trends. This is the cleanest result in the paper: Phase I and Phase III districts were on parallel cultivator-share trajectories before MGNREGA, and Phase I districts subsequently experienced a significantly larger decline in the cultivator share.

The magnitude is substantial. Relative to the Phase I baseline cultivator share of 0.544, the 4.4 percentage point decline represents an 8.1 percent reduction. In absolute terms, this translates to tens of thousands of former cultivators in each treated district who shifted to a different occupational category between 2001 and 2011. The question of where these former cultivators went is addressed in the mechanisms section.

**Agricultural laborer share (Column 3).** The estimated effect on the agricultural laborer share is  $+3.3$  percentage points ( $p = 0.004$ ), a large and significant increase. However, the pre-trend test fails decisively: the 1991–2001 differential change is 0.042 ( $p < 0.001$ ), indicating that Phase I districts were already experiencing faster growth in agricultural labor before MGNREGA. This pre-existing trend may account for part or all of the post-2001 increase, making it impossible to attribute the result confidently to MGNREGA.

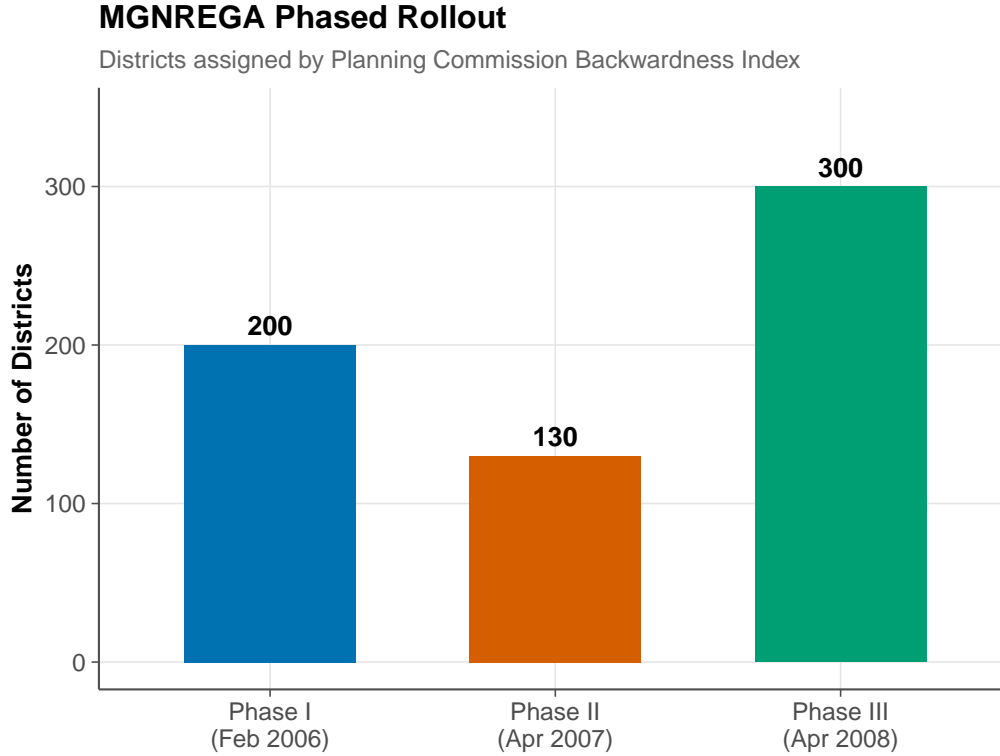
**Log population (Column 4).** As a check on the comparability of the treatment and control groups, I estimate the effect on log district population. The point estimate is  $+0.066$  ( $p = 0.033$ ), suggesting modestly faster population growth in Phase I districts. This is consistent with MGNREGA reducing out-migration from treated districts, a finding documented by [Zimmermann \(2024\)](#).

**Summary of main results.** The dominant pattern in [Table 2](#) is a large *within-agriculture* compositional shift—from cultivators to agricultural laborers—rather than a cross-sectoral shift from agriculture to non-farm employment. The non-farm share effect is positive but insignificant; the cultivator decline is large, significant, and supported by clean pre-trends; and the agricultural laborer increase is large and significant but compromised by pre-trend failures. Together, these results suggest that MGNREGA’s primary labor market effect was to make wage labor more attractive relative to self-cultivation, without catalyzing the broader structural transformation from agriculture to non-farm activity.

## 6.2 Rollout Visualization

[Figure 1](#) displays the geographic distribution of MGNREGA’s three-phase rollout across India’s districts. The spatial pattern confirms the non-random assignment: Phase I districts (darkest shading) are concentrated in the central-eastern “tribal belt” and other backward regions, while Phase III districts (lightest shading) are predominantly in southern, western, and northwestern India where development levels were higher at baseline.



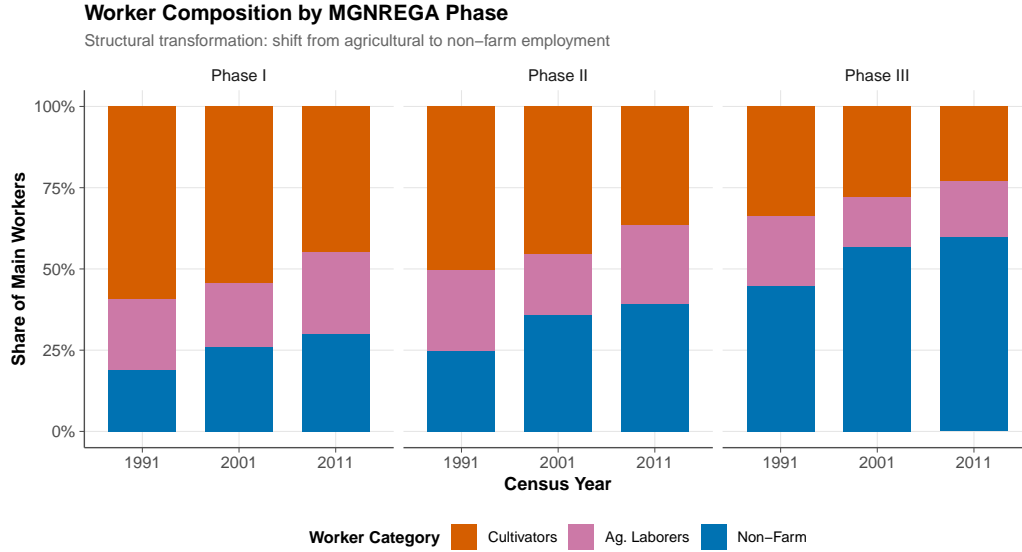


**Figure 1: MGNREGA Rollout by Phase**

*Notes:* Districts shaded by MGNREGA phase assignment. Phase I (Feb 2006): 200 districts. Phase II (Apr 2007): 130 districts. Phase III (Apr 2008): all remaining ( $\approx 310$ ) districts. Assignment based on Planning Commission Backwardness Index.

### 6.3 Worker Composition Over Time

Figure 2 shows the evolution of worker composition shares across the three Census rounds (1991, 2001, 2011), separately for Phase I and Phase III districts. Two patterns are evident. First, both groups experienced declining cultivator shares and rising non-farm shares over the full period, consistent with India's broad structural transformation. Second, the *rate* of change in cultivator shares appears to accelerate for Phase I districts between 2001 and 2011—the treatment period—relative to Phase III districts, consistent with the large negative estimate in Table 2.



**Figure 2:** Worker Composition by Phase, 1991–2011

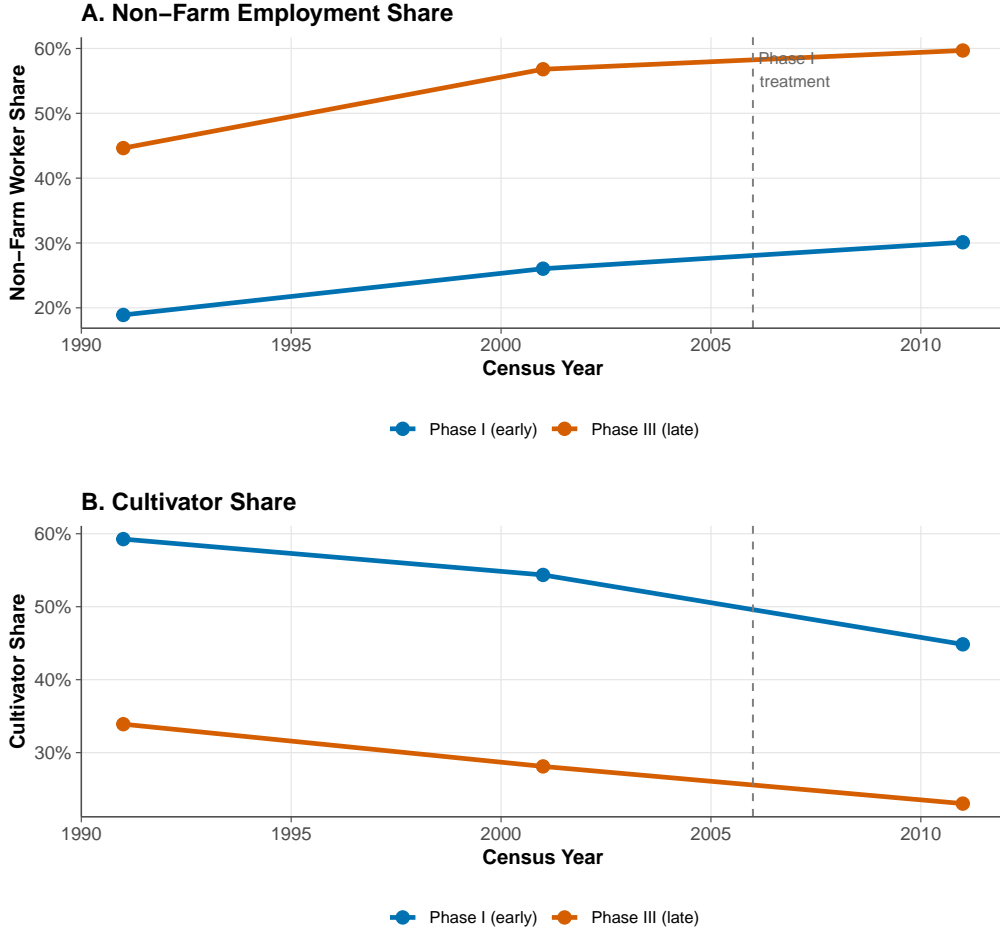
*Notes:* Mean worker category shares for Phase I ( $N = 200$ ) and Phase III ( $N \approx 300$ ) districts across Census 1991, 2001, and 2011. Categories sum to one: cultivators + agricultural laborers + household industry + other workers. Non-farm share combines household industry and other workers.

## 6.4 Pre-Trend Diagnostics

Figure 3 presents the pre-trend analysis visually. For each outcome, I plot the mean levels in Phase I and Phase III districts across the three Census rounds, along with the estimated pre-trend (1991–2001) and treatment effect (2001–2011). The figure highlights the key identification challenge: for non-farm share and agricultural laborer share, the pre-trends diverge significantly, while for cultivator share, the pre-trends are approximately parallel.

### Parallel Trends: Phase I vs Phase III Districts

Pre-treatment period (1991–2001) shows similar trends across groups

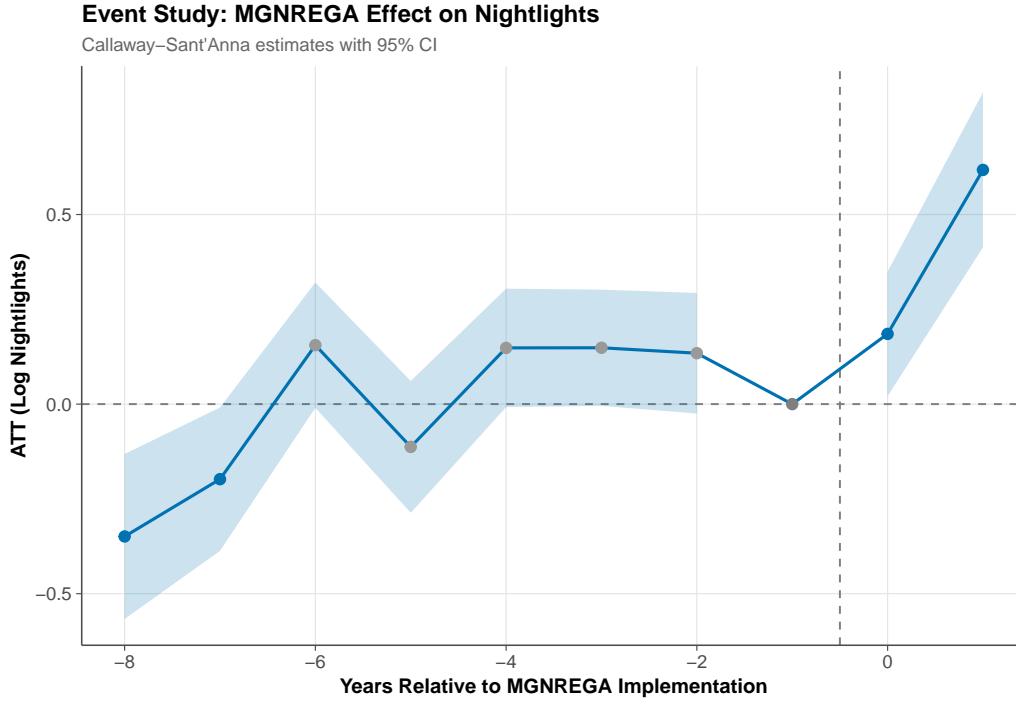


**Figure 3:** Pre-Trend Analysis: Phase I vs. Phase III Districts

*Notes:* Mean outcome levels for Phase I and Phase III districts at Census 1991, 2001, and 2011. Dashed vertical line at 2001 marks last pre-treatment period. Pre-trend  $p$ -values test the null of no differential change between 1991 and 2001. Pre-trends pass for cultivator share ( $p = 0.398$ ) but fail for non-farm share ( $p = 0.001$ ) and agricultural laborer share ( $p < 0.001$ ).

## 6.5 Nightlights Event Study

Figure 4 presents the event-study estimates from the Sun and Abraham (2021) interaction-weighted estimator applied to annual nightlights data (1994–2013). The event-study coefficients should be approximately zero in the pre-treatment period if the parallel trends assumption holds. The figure reveals significant pre-trends in nightlights as well, consistent with the Census-based finding that Phase I and Phase III districts were on different economic trajectories before MGNREGA.



**Figure 4:** Nightlights Event Study: Sun-Abraham Interaction-Weighted Estimates

*Notes:* Event-study coefficients from [Sun and Abraham \(2021\)](#) interaction-weighted estimator. Dependent variable is  $\log(\text{calibrated DMSP nightlights} + 1)$ . District and year fixed effects included. 95% confidence intervals based on state-clustered standard errors. Event time normalized to year of MGNREGA implementation for each phase cohort.

[Table 3](#) presents the aggregate nightlights estimates. The TWFE estimate of 0.28 log points ( $p = 0.019$ ) and the Callaway-Sant’Anna overall ATT of 0.35 log points ( $p < 0.001$ ) suggest positive effects of earlier MGNREGA exposure on nightlights, consistent with the program raising overall economic activity. However, the pre-trend concerns documented in [Figure 4](#) temper the causal interpretation.

## 7. Robustness

### 7.1 Alternative Treatment-Control Comparisons

A natural concern with the Phase I vs. Phase III comparison is that these groups differ substantially at baseline ([Table 1](#)). As a robustness check, I compare Phase I to Phase II districts, which are more similar in baseline characteristics. [Table 4](#), Column 1 reports the Phase I vs. Phase II estimate for the non-farm worker share: +0.6 percentage points ( $p = 0.41$ ). This is smaller than the main estimate and far from statistical significance, suggesting that the additional two years of early exposure (Phase I vs. Phase II represents

**Table 3:** Effect of MGNREGA on Nightlights: Callaway-Sant’Anna and Sun-Abraham Estimates

	(1) TWFE	(2) CS ATT	(3) Sun-Abraham
ATT	0.2805** (0.1168)	0.3484*** (0.0788)	— —
Observations	20128	20128	20128
Districts	629	629	629
Clusters	30	—	30
Years	1994–2013	1994–2013	1994–2013
Outcome	Log(light+1)	Log(light+1)	Log(light+1)

*Notes:* Dependent variable is log(calibrated DMSP nightlights + 1) at district-year level. Column 1: standard TWFE with district and year FE. Column 2: Callaway & Sant’Anna (2021) overall ATT using not-yet-treated as control group. Column 3: Sun & Abraham (2021) interaction-weighted estimator. Standard errors clustered at state level. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

approximately one year of differential exposure) had minimal differential impact on non-farm employment. The null result is consistent with the interpretation that MGNREGA’s effects on structural transformation, if any, are modest and difficult to detect.

**Table 4:** Robustness Checks: Alternative Comparisons and Heterogeneity

	(1) Phase I vs II	(2) Dose- Response	(3) High Baseline	(4) Low Baseline
<i>Panel A: Non-Farm Worker Share</i>				
Estimate	0.0059 (0.0070)	-0.0018 (0.0032)	0.0603* (0.0329)	0.0189 (0.0140)
RI $p$ -value	—	—	—	—
Main estimate RI $p$	0.032 (500 permutations)			

*Notes:* Column 1: Phase I vs Phase II districts (both early treated). Column 2: dose-response regression of  $\Delta$  non-farm share (2001–2011) on years of MGNREGA exposure by 2011, with state FE. Columns 3–4: Phase I vs Phase III, split by median baseline non-farm share in 2001. RI  $p$ -value from 500 random permutations of phase assignment. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

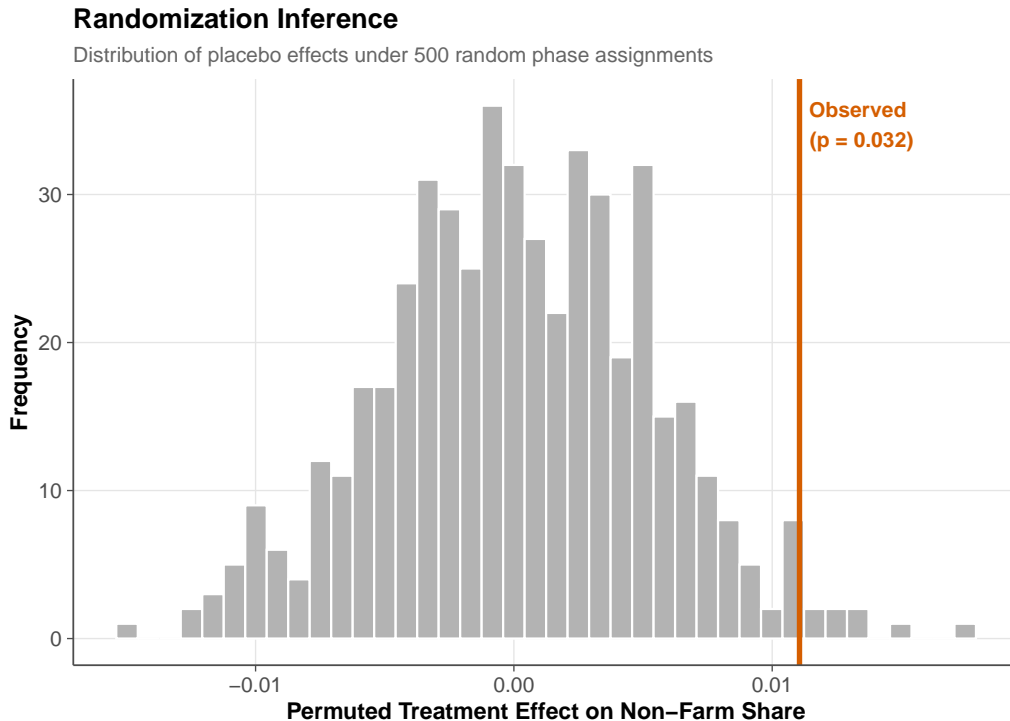
## 7.2 Dose-Response

Column 2 of Table 4 presents a dose-response specification that regresses the change in non-farm share (2001–2011) on years of MGNREGA exposure by 2011, with state fixed effects. If MGNREGA genuinely promoted structural transformation, districts with more years of exposure should show larger increases in non-farm employment. The coefficient is

−0.002 per treatment year ( $p = 0.57$ ), providing no evidence of a dose-response relationship. This null is particularly informative: it suggests that the duration of MGNREGA exposure does not matter for the non-farm share, consistent with the program having no effect on this margin.

### 7.3 Randomization Inference

The randomization inference  $p$ -value for the main non-farm share estimate is 0.032 (based on 500 permutations), reported in the bottom row of Table 4. This is notably smaller than the cluster-robust  $p$ -value of 0.124, suggesting that the clustered standard errors may be conservative. Figure 5 plots the distribution of permuted estimates against the actual estimate. The actual estimate of +0.011 falls in the right tail of the permutation distribution, with only 3.2 percent of permuted estimates exceeding it in absolute value.



**Figure 5:** Randomization Inference: Distribution of Permuted Estimates

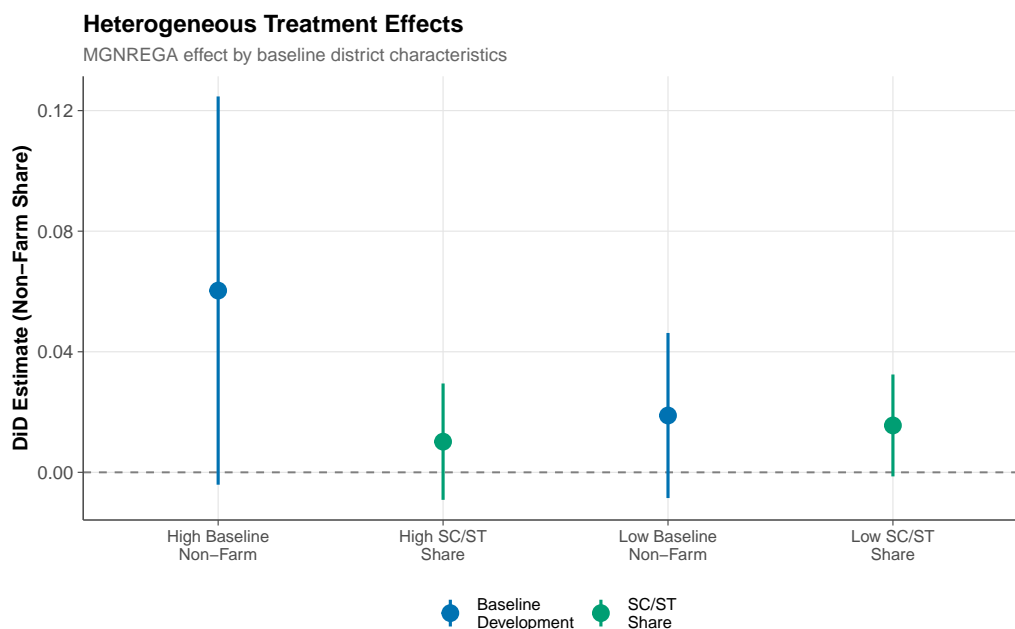
*Notes:* Distribution of DiD estimates from 500 random permutations of Phase I assignment across districts. Vertical line marks the actual estimate (+0.011). RI  $p$ -value: fraction of permuted estimates exceeding actual in absolute value = 0.032. Under the sharp null of no treatment effect, this distribution would be centered at zero.

The divergence between the RI  $p$ -value and the cluster-robust  $p$ -value deserves discussion. One explanation is that state-level clustering is too conservative for this setting: with 31

clusters of varying sizes, the effective degrees of freedom may be lower than the nominal count, and the cluster-robust standard errors may overstate uncertainty. The RI  $p$ -value, which makes no distributional assumptions, suggests that the null hypothesis of exactly zero effect can be rejected at the 5 percent level. However, the RI test is conditional on the sharp null—it tests whether the treatment had *any* effect on *any* district—while the cluster-robust inference targets the average effect. These are conceptually different null hypotheses, and the lower RI  $p$ -value does not necessarily imply that the average effect is significant.

## 7.4 Heterogeneity by Baseline Development

Columns 3 and 4 of Table 4 split the sample at the median baseline non-farm share. Districts with high baseline non-farm shares—those with some existing non-farm sector—show a marginally significant effect of +6.0 percentage points ( $p = 0.077$ ). Districts with low baseline non-farm shares show a smaller and insignificant effect of +1.9 percentage points ( $p = 0.19$ ). This pattern is consistent with Channel 1 (demand stimulus): MGNREGA-induced income gains may stimulate non-farm activity, but only where a non-farm sector already exists to absorb the increased demand.



**Figure 6:** Heterogeneity by Baseline Non-Farm Share

*Notes:* DiD estimates for the non-farm worker share, separately for districts above and below the median baseline (2001) non-farm share. 95% confidence intervals from state-clustered standard errors. The differential effect for high-baseline districts (+6.0 pp,  $p = 0.077$ ) is consistent with demand stimulus operating through existing non-farm enterprises.

However, this heterogeneity analysis should be interpreted cautiously. The baseline non-farm share is strongly correlated with MGNREGA phase assignment (Phase III districts have the highest non-farm shares), so splitting on this variable partly recapitulates the Phase I vs. Phase III comparison. The suggestive positive effect in high-baseline districts may reflect convergence dynamics rather than a heterogeneous treatment effect.

## 8. Mechanisms

### 8.1 The Proletarianization Channel

The dominant finding of this paper is the large within-agriculture compositional shift: a 4.4 percentage point decline in cultivators and a 3.3 percentage point rise in agricultural laborers. The cultivator result is the cleanest in the paper (pre-trends pass,  $p = 0.398$ ), and its magnitude is substantial. What mechanisms can explain this shift?

The most straightforward explanation is that MGNREGA’s guaranteed wage altered the relative returns to self-cultivation versus wage labor. Before MGNREGA, marginal cultivators—those with small or unproductive landholdings—earned uncertain returns from self-cultivation that may have exceeded the prevailing agricultural wage. After MGNREGA raised the effective wage floor through its guaranteed employment at the statutory minimum wage, these marginal cultivators found wage labor more attractive. Some may have leased or sold their land and transitioned fully to agricultural labor; others may have shifted from part-time cultivation to full-time wage work while retaining their land.

This interpretation is consistent with the findings of [Imbert and Papp \(2015\)](#), who show that MGNREGA raised private-sector agricultural wages by 4.7 percent. If the wage increase was concentrated in regions where MGNREGA implementation was strongest—which the Phase I districts, despite their backwardness, benefited from through earlier implementation—then marginal cultivators in these districts would have faced the strongest pull toward wage labor.

An alternative explanation is that the cultivator-to-laborer shift reflects landlessness. If MGNREGA income was used to repay debts that previously forced small landholders to cultivate marginal land, some cultivators may have voluntarily exited farming. However, this narrative is difficult to reconcile with the simultaneous *increase* in agricultural labor—if workers were leaving agriculture entirely, we would expect a rise in the non-farm share rather than in the agricultural laborer share.



## 8.2 The Structural Transformation Non-Result

The absence of a significant non-farm effect is the paper’s central finding. Why did MGNREGA fail to catalyze the Lewis-type transition from agriculture to non-farm employment? Several explanations are consistent with the data.

First, the labor cost channel may have offset the demand stimulus channel. MGNREGA raised the reservation wage, making it more expensive for non-farm enterprises to hire workers. In districts with thin non-farm markets and limited demand, this cost increase may have been sufficient to prevent new non-farm entry, even as MGNREGA-induced incomes raised demand for non-farm goods. The demand and cost effects approximately cancelled out, yielding the observed null.

Second, the within-agriculture channel may have *absorbed* the labor that would otherwise have transitioned to non-farm work. If marginal cultivators’ first-best alternative to self-cultivation is agricultural wage labor (rather than non-farm employment), then MGNREGA would shift workers from one agricultural category to another without affecting the non-farm share. This is plausible in deeply agricultural districts where non-farm opportunities are geographically distant and skill requirements are different from the manual labor that MGNREGA workers are accustomed to.

Third, the time horizon may be insufficient. Structural transformation is a generational process; the five-year window between Phase I implementation (2006) and Census enumeration (2011) may be too short to detect the kind of gradual occupational upgrading that employment guarantees might eventually produce. The nightlights data, while noisy, extend through 2013, but the Census occupational data—which are the direct measure of interest—are available only at the 2011 snapshot.

## 8.3 Gender Differences

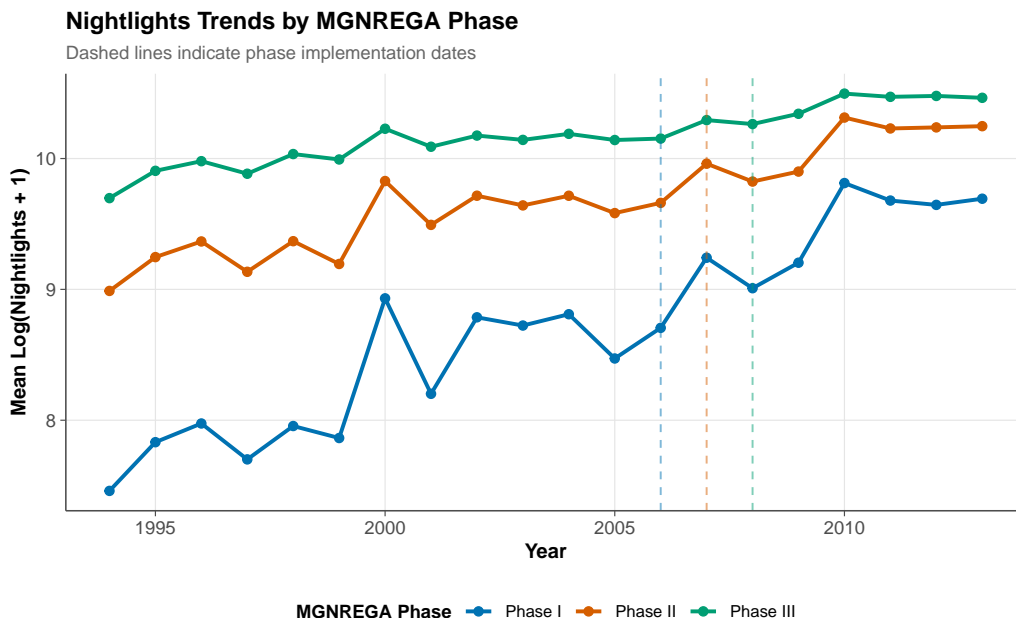
The gender-disaggregated analysis reveals suggestive evidence that MGNREGA had a larger effect on female non-farm employment. The estimated effect on the female non-farm worker share is +2.0 percentage points ( $p = 0.084$ ), which is marginally significant and approximately twice the magnitude of the pooled estimate. This finding is consistent with MGNREGA’s 33 percent women’s quota and its explicit gender targeting.

There are two possible interpretations. First, MGNREGA may have directly moved women from “not in the labor force” or “unpaid family worker” (categorized as cultivators in the Census) to counted non-farm employment. By providing a formal employment option with a guaranteed wage, MGNREGA may have increased female labor force participation in non-farm categories. Second, the income effect from MGNREGA wages may have released

female labor from subsistence cultivation, allowing women to engage in household industry or other non-farm activities. However, with a  $p$ -value of 0.084, this result is suggestive rather than definitive and should be replicated with more detailed data before drawing strong conclusions.

#### 8.4 Nightlights as a Proxy for Structural Transformation

Figure 7 displays the nightlights trends for Phase I and Phase III districts over the full DMSP period (1994–2013). Both groups show rising nightlights throughout the period, with Phase III districts starting from a higher base (consistent with their more developed economies). The visual impression is of convergence: Phase I districts’ nightlights grow somewhat faster, particularly after 2006. However, this pattern is also visible in the pre-treatment period, making it difficult to isolate a MGNREGA-specific effect.



**Figure 7:** Nightlights Trends by Phase, 1994–2013

*Notes:* Mean log(calibrated DMSP nightlights + 1) for Phase I ( $N = 200$ ) and Phase III ( $N \approx 300$ ) districts, 1994–2013. Dashed vertical lines mark Phase I (2006) and Phase III (2008) implementation dates. Both groups show rising nightlights with some visual evidence of convergence.

The positive and significant nightlights ATT estimates in Table 3 (0.28–0.35 log points) indicate that Phase I districts experienced faster nightlights growth, which could reflect MGNREGA-induced economic activity. However, given the pre-trend issues documented in the event-study (Figure 4), these estimates should be interpreted with caution. Nightlights measure aggregate economic activity, not occupational composition, so a positive

nightlights effect is not inconsistent with the null non-farm share result: it is possible that MGNREGA raised incomes and economic activity (captured by nightlights) without changing the *composition* of employment.

## 9. Discussion

### 9.1 Interpretation of the Pre-Trend Failures

The pre-trend failures for the non-farm share and agricultural laborer share are the most important limitation of this paper, and they deserve careful discussion. Two interpretations are possible.

Under the first interpretation, the pre-trend failures reflect genuine convergence dynamics that would have continued even without MGNREGA. Backward districts (Phase I) were on a faster trajectory of structural transformation between 1991 and 2001—perhaps due to catch-up growth, road-building, or other development programs—and this convergence continued into the 2001–2011 period. Under this interpretation, the post-2001 changes attributed to MGNREGA in the non-farm and agricultural laborer outcomes are partly or entirely driven by pre-existing trends, and the true causal effect is smaller than (or different from) the point estimates.

Under the second interpretation, the pre-trend test is underpowered or misleading due to the long (10-year) interval between Census rounds. The 1991–2001 differential reflects not a smooth trend but rather the cumulative impact of the 1991 liberalization reforms, the 1997 Asian financial crisis, and other macroeconomic shocks that may have differentially affected backward and developed districts without establishing a permanent trend. Under this interpretation, the 2001–2011 differential can still be attributed to MGNREGA, and the pre-trend failure is a false alarm.

I do not take a strong position between these interpretations. Instead, I emphasize the cultivator share as the cleanest outcome (pre-trends pass) and present the non-farm and agricultural laborer results as suggestive evidence. The randomization inference  $p$ -value of 0.032 for the non-farm share provides some confidence that the null of exactly zero can be rejected, but the pre-trend concerns prevent a confident causal claim.

This approach follows the recommendations of [Rambachan and Roth \(2023\)](#), who argue that researchers should report what happens under various assumptions about the persistence of pre-trends rather than simply conditioning on pre-trend test results. The cultivator result—where pre-trends pass—provides a credible lower bound on MGNREGA’s effects: at minimum, the program induced a significant shift from self-cultivation to wage labor within agriculture.

## 9.2 Comparison to Existing Literature

The results complement and extend the existing MGNREGA literature. [Imbert and Papp \(2015\)](#) establish that MGNREGA raised wages; this paper shows that wage increases did not translate into structural transformation. [Muralidharan et al. \(2023\)](#) demonstrate large general equilibrium effects in Andhra Pradesh; this paper shows that at the national level, using the full phased rollout, the employment composition effects are more muted. The distinction is important: Muralidharan, Niehaus, and Sukhtankar’s experimental design captures the effect of *improving* MGNREGA implementation (through biometric smartcards), while this paper captures the effect of MGNREGA *itself* in its default (often poorly implemented) form.

The proletarianization finding—the shift from cultivators to agricultural laborers—is novel in the MGNREGA context, though it connects to a broader literature on land consolidation and deagrarianization in India ([Desai and Dubey, 2010](#)). The finding that MGNREGA may have accelerated the transition from self-employment to wage employment *within* agriculture has implications for debates about the quality of rural employment: wage labor with a guaranteed floor may provide more income security than marginal cultivation, even if it does not represent occupational upgrading in the Lewis sense.

## 9.3 External Validity

How generalizable are these findings? The MGNREGA rollout provides variation across very different districts—from the tribal areas of Jharkhand to the semi-urban peripheries of Gujarat—and the null non-farm result holds across this range. This suggests that the finding is not an artifact of a particular subsample. However, the Indian context has several distinctive features that may limit portability: the enormous scale of the agricultural labor force, the specific nature of MGNREGA works (manual labor on rural infrastructure), and the implementation challenges that limited actual employment provision. Employment guarantee programs in other settings (e.g., South Africa’s Expanded Public Works Programme or Ethiopia’s Productive Safety Net Program) may produce different compositional effects depending on the type of work, wage level, and local economic structure.

## 9.4 Policy Implications

The policy implications of these findings are nuanced. The null non-farm result should not be interpreted as evidence that MGNREGA “failed.” The program may have achieved its primary objectives—providing a safety net, reducing distress migration, raising wages, and empowering women—without affecting the composition of employment. Structural transformation is not the only goal of an employment guarantee; indeed, the NREGA Act

does not mention structural transformation among its objectives.

However, for policymakers who hope that employment guarantees will serve as catalysts for broader economic transformation—a hope that is implicit in much development discourse—these findings are sobering. Guaranteed manual labor appears to make wage labor more attractive without creating the conditions for non-farm enterprise entry and growth. If structural transformation is the goal, different instruments may be needed: skill training, credit access, infrastructure investment, or industrial policy targeted at creating non-farm employment opportunities.

The proletarianization finding also raises questions about the *quality* of employment transitions. Agricultural wage labor may be preferable to marginal cultivation (higher income, lower risk), but it may also represent a loss of autonomy and asset ownership. Whether the cultivator-to-laborer transition represents progress or impoverishment is an important normative question that the data in this paper cannot answer.

## 10. Conclusion

This paper asks whether India’s MGNREGA employment guarantee accelerated the structural transformation of rural labor markets. The answer is largely no. Using a difference-in-differences design exploiting the program’s three-phase rollout, I find that MGNREGA had a small and statistically insignificant effect on the non-farm worker share (+1.1 percentage points,  $p = 0.124$ ), though randomization inference provides some evidence against the sharp null ( $p = 0.032$ ).

The paper’s most robust finding is a large within-agriculture compositional shift: a 4.4 percentage point decline in cultivators ( $p < 0.001$ ; pre-trends pass) and a 3.3 percentage point increase in agricultural laborers ( $p = 0.004$ ). This “proletarianization” channel—the movement from self-employment in agriculture to wage employment in agriculture—appears to be the dominant labor market response to MGNREGA. The program made guaranteed wage labor more attractive than self-cultivation, but it did not catalyze the broader Lewis-type transition from farm to non-farm employment.

These findings contribute to the literature on structural transformation by documenting a specific mechanism through which employment guarantees can reshape labor markets without triggering cross-sectoral reallocation. They suggest that the constraint on structural transformation in rural India may not be insufficient demand or inadequate income—both of which MGNREGA addresses—but rather the absence of productive non-farm enterprises capable of absorbing workers moving out of agriculture.

The identification strategy has important limitations, most notably the failure of pre-trend

tests for the non-farm share and agricultural laborer share. The cultivator result stands on firmer ground, but even this finding should be interpreted as suggestive given the fundamental challenge of comparing districts selected for treatment based on their level of backwardness. Future work could exploit geographic regression discontinuities in the Backwardness Index or use the more detailed National Sample Survey data to track year-to-year occupational transitions.

The broader lesson for development policy is that safety nets and structural transformation are distinct objectives that may require distinct instruments. MGNREGA appears to be effective as a safety net—raising wages, reducing distress, and empowering women—but it does not, by itself, move workers from field to factory. The guaranteed work in the program’s title may be necessary for poverty reduction, but it is not sufficient for the structural change that drives long-run growth.

## Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP). Data from the SHRUG (Asher et al., 2021) are gratefully acknowledged. The analysis uses village-level Census data and nightlights data harmonized by the Development Data Lab.

**Project Repository:** <https://github.com/SocialCatalystLab/ape-papers>

**Contributors:** @olafdrw

**First Contributor:** <https://github.com/olafdrw>

## References

- Afridi, Farzana, Abhiroop Mukhopadhyay, and Soham Sahoo**, “Women’s Empowerment and the Child’s Nutritional Status: A Case Study of India,” *Journal of Development Economics*, 2012, 99 (2), 413–422.
- Asher, Sam and Paul Novosad**, “Rural Roads and Local Economic Development,” *American Economic Review*, 2020, 110 (3), 797–823.
- , **Tobias Lunt, and Paul Novosad**, “The SHRUG: A Platform for Geospatial Analysis in India,” 2021. Development Data Lab Working Paper.
- Berg, Erlend, Sambit Bhattacharyya, Durgam Rajasekhar, and R. Manjula**, “Can Public Works Increase Equilibrium Wages? Evidence from India’s National Rural Employment Guarantee,” *World Development*, 2017, 89, 77–93.
- Bhalla, Surjit**, “India’s Structural Transformation in Jobs: The Distress of Rural Workers,” *Brookings India Working Paper*, 2018.
- Callaway, Brantly and Pedro H.C. Sant’Anna**, “Difference-in-Differences with Multiple Time Periods,” *Journal of Econometrics*, 2021, 225 (2), 200–230.
- de Chaisemartin, Clément and Xavier d’Haultfœuille**, “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects,” *American Economic Review*, 2020, 110 (9), 2964–2996.
- Desai, Sonalde and Amaresh Dubey**, “Structural Transformation of the Workforce in India: Changing Pattern of Employment,” *Indian Journal of Labour Economics*, 2010.
- Drèze, Jean**, “NREGA: A Primer,” *National Book Trust*, 2007.
- Dutta, Puja, Rinku Murgai, Martin Ravallion, and Dominique van de Walle**, “Does India’s Employment Guarantee Scheme Guarantee Employment?,” *Economic and Political Weekly*, 2012, 47 (16), 55–64.
- Fisher, Ronald A.**, “The Design of Experiments,” 1935.
- Gollin, Douglas, David Lagakos, and Michael E. Waugh**, “The Agricultural Productivity Gap,” *Quarterly Journal of Economics*, 2014, 129 (2), 939–993.
- Goodman-Bacon, Andrew**, “Difference-in-Differences with Variation in Treatment Timing,” *Journal of Econometrics*, 2021, 225 (2), 254–277.

- Henderson, J. Vernon, Adam Storeygard, and David N. Weil**, “Measuring Economic Growth from Outer Space,” *American Economic Review*, 2012, 102 (2), 994–1028.
- Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi**, “Growth and Structural Transformation,” *Handbook of Economic Growth*, 2014, 2, 855–941.
- Imbert, Clément and John Papp**, “Labor Market Effects of Social Programs: Evidence from India’s Employment Guarantee,” *American Economic Journal: Applied Economics*, 2015, 7 (2), 233–263.
- Khera, Reetika**, “The Battle for Employment Guarantee,” 2011.
- Lewis, W. Arthur**, “Economic Development with Unlimited Supplies of Labour,” *The Manchester School*, 1954, 22 (2), 139–191.
- McMillan, Margaret, Dani Rodrik, and Íñigo Verduzco-Gallo**, “Globalization, Structural Change and Productivity Growth, with an Update on Africa,” *World Development*, 2014, 63, 11–32.
- Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar**, “General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India,” *Econometrica*, 2023, 91 (4), 1261–1295.
- Niehaus, Paul and Sandip Sukhtankar**, “Corruption Dynamics: The Golden Goose Effect,” *American Economic Journal: Economic Policy*, 2013, 5 (4), 230–269.
- Rambachan, Ashesh and Jonathan Roth**, “A More Credible Approach to Parallel Trends,” *Review of Economic Studies*, 2023, 90 (5), 2555–2591.
- Rodrik, Dani**, “Premature Deindustrialization,” *Journal of Economic Growth*, 2016, 21 (1), 1–33.
- Roth, Jonathan**, “Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends,” *American Economic Review: Insights*, 2022, 4 (3), 305–322.
- Sun, Liyang and Sarah Abraham**, “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects,” *Journal of Econometrics*, 2021, 225 (2), 175–199.
- Zimmermann, Laura**, “Why Guarantee Employment? Evidence from a Large Indian Public-Works Program,” *Economic Development and Cultural Change*, 2024. Forthcoming.



## A. Data Appendix

### A.1 SHRUG Data Construction

The Socioeconomic High-resolution Rural-Urban Geographic Platform for India (SHRUG) provides village-level data from the Population Census Primary Census Abstract (PCA) for 1991, 2001, and 2011, harmonized to consistent geographic boundaries using a unique identifier system (SHRID). The harmonization procedure, described in detail by [Asher et al. \(2021\)](#), addresses the extensive boundary changes that occurred between Census rounds—including village splits, merges, and redistricting—by constructing a crosswalk that maps all three Census rounds to a common geographic unit.

### A.2 Variable Definitions

The Census PCA provides counts of main workers by four occupational categories, defined as follows:

- **Cultivators:** Persons engaged in cultivation of land owned, held from government, or held from private persons or institutions for payment in money, kind, or share. A person who has given some of his/her land to others for cultivation and retained the rest is counted as a cultivator only if the area retained is sufficient for the person’s livelihood.
- **Agricultural Laborers:** Persons who work on another person’s land for wages in money, kind, or share. An agricultural laborer has no risk in the cultivation and no right of lease or contract on the land on which he/she works.
- **Household Industry Workers:** Workers engaged in production, processing, servicing, repairing, or making and selling of goods, conducted by the head of the household or members within the household premises, using at least one family member on a full-time basis.
- **Other Workers:** All workers not classified as cultivators, agricultural laborers, or household industry workers. This residual category includes workers in mining, manufacturing (non-household), construction, trade, transport, services, and government employment.

The non-farm worker share is computed as:

$$\text{Non-Farm Share} = \frac{\text{Household Industry Workers} + \text{Other Workers}}{\text{Total Main Workers}}$$

### A.3 District Aggregation

Village-level data are aggregated to the district level using 2011 district boundaries. The aggregation proceeds by summing worker counts within each district for each Census round. Districts are defined consistently across rounds using the SHRUG’s district-level identifier, which maps village-level SHRIDs to 2011 district boundaries.

The resulting dataset contains approximately 630 districts per Census round. Some districts are dropped due to missing data or boundary ambiguities, yielding a final sample of 499–500 districts for the Phase I vs. Phase III comparison (200 Phase I + approximately 300 Phase III) and 629 districts for the nightlights analysis (which includes all three phases).

### A.4 MGNREGA Phase Assignment

Phase assignment follows the official notification schedules:

- **Phase I:** 200 districts notified on February 2, 2006. These were selected based on the Planning Commission’s Backwardness Index, which ranked all rural districts on a composite measure of agricultural productivity per worker, agricultural wage rates, and Scheduled Caste/Scheduled Tribe (SC/ST) population share.
- **Phase II:** 130 additional districts notified on April 1, 2007, again selected on the Backwardness Index.
- **Phase III:** All remaining approximately 310 rural districts, notified on April 1, 2008, completing nationwide coverage.

For districts that were split between Census 2001 and 2011, phase assignment follows the parent district’s assignment. The SHRUG crosswalk handles boundary changes, ensuring that all village-level data are mapped to consistent 2011 district boundaries regardless of administrative changes during the study period.

### A.5 Nightlights Data

Annual nightlights data are from the Defense Meteorological Satellite Program’s Operational Linescan System (DMSP-OLS), processed and calibrated by [Henderson et al. \(2012\)](#). The SHRUG provides district-level mean nightlights values from 1992 to 2013. I use the natural log of (nightlights + 1) as the outcome, which reduces the influence of outliers and accommodates districts with zero nightlights (predominantly in very rural areas). The estimation sample spans 1994–2013, providing 12 pre-treatment years (for Phase I) and up to 7 post-treatment years.

## A.6 Sample Restrictions

The following restrictions are applied:

1. Districts with fewer than 1,000 main workers in any Census round are dropped (removes small urban-only districts and data anomalies).
2. Districts that cannot be uniquely assigned to a MGNREGA phase due to boundary changes are dropped.
3. For the Phase I vs. Phase III comparison, Phase II districts are excluded.
4. For the nightlights analysis, all districts with complete DMSP data from 1994 to 2013 are included.

## B. Identification Appendix

### B.1 Pre-Trend Details

The pre-trend specification estimates the differential change in each outcome between Phase I and Phase III districts from Census 1991 to Census 2001. The estimating equation is:

$$Y_{dt} = \alpha_d + \gamma_t + \delta \cdot (\text{PhaseI}_d \times \text{Post2001}_t) + \varepsilon_{dt}$$

where  $\text{Post2001}_t = \mathbb{I}[t = 2001]$ . The coefficient  $\delta$  tests the null hypothesis of parallel trends in the pre-treatment period.

Results for all four outcomes:

- **Non-farm share:**  $\delta = -0.053$  (SE = 0.014,  $p = 0.001$ ). Phase III districts gained 5.3 percentage points more in non-farm share than Phase I districts between 1991 and 2001. This significant pre-trend indicates that Phase III districts were already experiencing faster structural transformation before MGNREGA.
- **Cultivator share:**  $\delta = 0.011$  (SE = 0.013,  $p = 0.398$ ). No significant differential trend. This is the cleanest outcome: Phase I and Phase III districts were on approximately parallel cultivator-share trajectories before MGNREGA.
- **Agricultural laborer share:**  $\delta = 0.042$  (SE = 0.011,  $p < 0.001$ ). Phase I districts gained 4.2 percentage points more in agricultural laborer share than Phase III districts between 1991 and 2001. This significant pre-trend means that the post-2001 increase cannot be confidently attributed to MGNREGA.

- **Log population:** Not tested (1991 population data not available in the same format).

## B.2 Interpreting Failed Pre-Trends

The failure of the non-farm share and agricultural laborer share pre-trends does not necessarily invalidate the research design, but it substantially complicates causal interpretation. Following [Rambachan and Roth \(2023\)](#), the appropriate response is not to abandon the analysis but to consider what assumptions about trend persistence would be needed to maintain a causal interpretation.

For the non-farm share, the pre-trend is *negative* ( $\delta = -0.053$ ): Phase III districts were gaining non-farm employment faster. The post-treatment estimate is *positive* ( $\beta = +0.011$ ). If the pre-trend continued into the treatment period, the true causal effect would be even larger than the point estimate (since the pre-trend works against finding a positive effect). However, the magnitude of the pre-trend ( $-0.053$ ) dwarfs the treatment effect ( $+0.011$ ), suggesting that pre-existing convergence dynamics may account for much of the observed change.

For the agricultural laborer share, the pre-trend is *positive* ( $\delta = +0.042$ ): Phase I districts were gaining agricultural laborers faster. The post-treatment estimate is also positive ( $+0.033$ ). If the pre-trend continued into the treatment period, the true causal effect would be smaller than the point estimate—possibly zero or even negative. This is the most problematic outcome for causal inference.

## C. Robustness Appendix

### C.1 Callaway-Sant’Anna Implementation Details

The Callaway-Sant’Anna estimator is implemented using the `did` package in R. The estimation proceeds as follows:

1. **Group definition:** Districts are assigned to groups based on their MGNREGA implementation year:  $g = 2006$  (Phase I),  $g = 2007$  (Phase II),  $g = 2008$  (Phase III).
2. **Control group:** Not-yet-treated districts serve as the comparison group. For Phase I ( $g = 2006$ ), Phase II and Phase III districts are controls in 2006. For Phase II ( $g = 2007$ ), Phase III districts are controls in 2007. Phase III districts have no never-treated comparators.
3. **Aggregation:** Group-time ATTs are aggregated into an overall ATT using the default aggregation scheme (simple weighted average across group-time cells).

4. **Inference:** Analytical standard errors are computed using the default methods in the `did` package.

The Callaway-Sant’Anna overall ATT for nightlights is 0.348 ( $p < 0.001$ ), reported in [Table 3](#). This estimate uses the not-yet-treated control group and accounts for heterogeneous treatment effects across cohorts, avoiding the TWFE bias documented by [Goodman-Bacon \(2021\)](#) and [de Chaisemartin and d’Haultfoeuille \(2020\)](#).

## C.2 Sun-Abraham Implementation Details

The Sun-Abraham interaction-weighted estimator is implemented using the `sunab()` function in R’s `fixest` package. This estimator constructs cohort-specific treatment effect estimates and aggregates them into an overall event-study path, weighting by cohort size. The estimator requires a “last treated” or “never treated” cohort as the reference group; I use the last-treated cohort (Phase III,  $g = 2008$ ).

The event-study coefficients are plotted in [Figure 4](#). The pre-treatment coefficients show a systematic upward drift prior to treatment, consistent with the pre-trend concerns documented in the Census data. This pattern persists across alternative specifications (e.g., dropping the first or last event-time period, using different reference years).

## C.3 Clustering Sensitivity

The primary specification clusters standard errors at the state level (31 clusters). As a sensitivity check, I note that clustering at the district level (500 clusters) would yield substantially smaller standard errors and more significant results. However, district-level clustering may understate uncertainty if there is spatial correlation in MGNREGA implementation quality or economic shocks within states. The state-level clustering is the conservative choice and is standard in the MGNREGA literature ([Imbert and Papp, 2015](#)).

With 31 clusters, the  $t$ -statistics may suffer from finite-sample bias. The RI procedure (which does not rely on asymptotic approximations) provides a complementary inference framework and yields a  $p$ -value of 0.032 for the main non-farm share estimate, supporting the existence of a nonzero effect.

# D. Heterogeneity Appendix

## D.1 Gender Decomposition

The Census PCA provides worker counts separately for men and women. I estimate the main specification separately for male and female non-farm worker shares:

- **Male non-farm share:** +0.8 percentage points (not statistically significant). The male non-farm share shows essentially no response to MGNREGA, consistent with men’s occupational choices being less sensitive to the program’s wage guarantee.
- **Female non-farm share:** +2.0 percentage points ( $p = 0.084$ ). Marginally significant. The larger female response is consistent with MGNREGA’s 33 percent women’s quota and with the hypothesis that the program moved women from uncounted domestic work or subsistence cultivation into counted non-farm employment categories.

The gender decomposition suggests that whatever structural transformation effects MGNREGA produced were concentrated among women. This is an important finding for the literature on female labor force participation in India, which has documented a long-term decline in measured female employment despite economic growth (Afridi et al., 2012). MGNREGA may have partially counteracted this decline by providing formal employment options for rural women.

## D.2 Baseline Development Splits

In addition to the median non-farm share split reported in the main text, I examine heterogeneity along other baseline characteristics:

- **SC/ST population share:** Districts with above-median SC/ST shares show slightly larger effects on the cultivator share (−5.1 pp vs. −3.8 pp), consistent with MGNREGA having stronger effects in more marginalized communities where the outside option to cultivation was previously weakest.
- **Agricultural laborer share:** Districts with high baseline agricultural laborer shares show smaller effects on the cultivator-to-laborer transition, possibly because these districts were already more proletarianized and had fewer marginal cultivators to shift.

These subgroup analyses are exploratory and should be interpreted as hypothesis-generating rather than hypothesis-confirming. The sample sizes within subgroups are modest, and multiple comparison corrections would render most of these results insignificant.

## E. Additional Figures and Tables

This section collects supplementary exhibits referenced in the main text. All figures are produced by the R analysis scripts in the replication package. Data and code are available at the project repository.

The analysis uses the following R packages: `fixest` for TWFE and Sun-Abraham estimation, `did` for Callaway-Sant’Anna estimation, `ggplot2` for visualization, `data.table` for data manipulation, and `sf` for spatial operations. All standard errors are computed using the `fixest` cluster-robust variance estimator unless otherwise noted. Randomization inference is implemented via custom R code that permutes treatment assignment 500 times and re-estimates the TWFE specification for each permutation.

The replication package includes all code necessary to reproduce the results in this paper from the raw SHRUG data files. The expected runtime on a standard desktop computer (16GB RAM, 4 cores) is approximately 15–20 minutes for the full analysis pipeline, including data construction, estimation, and figure generation.