Rural Roads and Firm Outcomes in India*

Bharti Nandwani†
Indira Gandhi Institute of Development Research, Mumbai, India

Punarjit Roychowdhury‡
Shiv Nadar University, Delhi NCR, India, GLO & CDES, Monash University

Binay Shankar§
Shiv Nadar University, Delhi NCR, India

August 9, 2025

Abstract

This paper examines the impact of a large-scale rural road construction program—the Pradhan Mantri Gram Sadak Yojana (PMGSY)—on the performance of rural manufacturing firms in India. While these firms provide vital non-farm employment in rural areas, their growth is often thought to be constrained by inadequate infrastructure. Leveraging administrative data and the quasi-random rollout of the program, we estimate effects using a two-way fixed effects framework. We find no evidence that improved road connectivity affects turnover, profits, or employment for formal enterprises. In contrast, informal firms experience significant gains in turnover, expenditure, profits, employment, and wage bills. These effects appear to be driven by reductions in infrastructure-related constraints: treated firms report fewer operational problems and less competition from larger firms, particularly in marketing and distribution. Our findings highlight the heterogeneous effects of rural infrastructure expansion and the greater responsiveness of informal enterprises.

JEL Classification: D22, O12, O18
Keywords: Firms, India, Infrastructure, Roads, Rural

*The authors' names are listed in alphabetical order, and all authors contributed equally to this work. All remaining errors are our own.

†RB1-102, Indira Gandhi Institute of Development Research, Film City Road, Mumbai 400065, India. Email: bharti@igidr.ac.in

‡Corresponding Author. Department of Economics, School of Humanities and Social Sciences, Shiv Nadar University, NH91, Tehsil Dadri, Greater Noida 201314, Uttar Pradesh, India. Email: punarjitroychowdhury@gmail.com.

§Department of Economics, School of Humanities and Social Sciences, Shiv Nadar University, NH91, Tehsil Dadri, Greater Noida 201314, Uttar Pradesh, India. Email:bs378@snu.edu.in.

1 Introduction

Rural manufacturing enterprises play a vital role in advancing structural transformation in low- and middle-income countries (LMICs) by providing non-agricultural employment opportunities and facilitating the movement of labor into higher-productivity sectors. However, these enterprises face persistent constraints, including inadequate infrastructure, frequent electricity outages, limited access to credit, and shortages of skilled labor. These barriers significantly limit their potential contribution to economic development and job creation. In the Indian context—the focus of this study—manufacturing activity was historically concentrated in urban areas. However, over the past two decades, rural areas have experienced a notable shift, with increases in the share of manufacturing output, employment, and the number of enterprises (Ghani, Goswami and Kerr, 2012).[1] According to the Annual Survey of Industries (ASI) 2019–20, approximately 62% of fixed manufacturing capital is now located in rural India, and rural firms account for 44% of total manufacturing employment. Despite this growing presence, the productivity of rural manufacturing firms remains suboptimal, largely due to inadequate infrastructure and other enabling conditions.

This paper investigates how rural infrastructure—specifically, improved road connectivity—shapes the performance of manufacturing firms in rural India. We focus on the Pradhan Mantri Gram Sadak Yojana (PMGSY), a flagship rural road construction program launched by the Government of India in 2000. PMGSY sought to enhance access to markets and services by building all-weather roads that connect previously unconnected villages (population ¿500) to nearby towns or road networks. At the time of its inception, nearly half of India's villages lacked such connectivity. By 2010, the program had made significant strides, and as of December 2024, 95% of the sanctioned Phase I projects had been completed. While a growing body of research has evaluated the impacts of PMGSY on agricultural outcomes, poverty alleviation, human capital, and labor market dynamics, relatively little is known

[1] https://www.ideasforindia.in/topics/productivity-innovation/how-can-india-become-a-manufacturing-powerhouse.html

about its consequences for firm-level outcomes—a key channel through which infrastructure may affect structural transformation. This paper contributes to that gap by providing the first systematic evidence on how rural road infrastructure influences the performance of manufacturing enterprises in rural areas.

A key empirical challenge in studying the impact of transportation infrastructure programs on firm outcomes is the endogenous placement of these investments. Infrastructure projects may be targeted toward underdeveloped areas to stimulate economic activity or, conversely, toward better-performing regions with high concentrations of manufacturing firms that can lobby for improved connectivity. The quasi-random nature of the rollout of PMGSY however allows us to examine the impact of rural roads infrastructure program on firm outcomes in a causal way. The design of PMGSY prioritised connectivity for larger villages, initially targeting those with populations above 1,000, followed by those above 500. As a result, the distribution of road construction across districts depended largely on the demographic composition of villages relative to these population cut-offs. This generates plausibly random variation in rural populations' exposure to paved roads at the district level, allowing us to credibly identify the program's effects.

There are several channels through which rural road networks are expected to influence firm outcomes. First, improved connectivity to market towns can provide access to larger markets that can help these firms scale up—especially relevant in the Indian context, where capacity utilisation averages only around 75% (FICCI survey). Enhanced spatial integration may also improve access to labor and other inputs. Shamdasani (2021) finds that PMGSY increased rural labor market integration, which benefited agricultural establishments. Second, better roads can lower transportation and logistics costs, reducing inventory holdings and increasing firm profitability—consistent with evidence from highway expansions (Datta, 2012). Third, a well developed network of roads can lead to clustering of economic activity leading firms to benefit from agglomeration economies (Hall 2016). However, improved market access may also intensify competition, potentially eroding the profits of incumbent firms

or leading them to reduce labor costs. In some cases, particularly among smaller informal firms, increased competition may result in market exit (Perra, Sanfilippo and Sundaram, 2024). Moreover, greater factor mobility could lead to capital and labor outflows to urban centers, offsetting the benefits of local market access (Banerjee, Duflo and Qian, 2020). Taken together, the net effect of rural road expansion on firm performance is theoretically ambiguous, highlighting the need for empirical investigation—such as the one undertaken in this study.

When studying the impact of the PMGSY program we focus on both formal and informal rural firms. Formal firms are defined as those employing at least 10 workers with power (or 20 without) and registered under the Factories Act of 1948, thereby subject to regulatory oversight and labor laws. In contrast, informal firms are typically smaller, unregistered, and outside the scope of formal regulation. Despite their informality, these enterprises play a crucial role in India's manufacturing landscape as they contribute about 34% to manufacturing Gross Value Added (GVA) and account for nearly 75% of total manufacturing employment.[2]

Our analysis draws on two data sources for rural firms. For informal firms, we use data from the 62nd (2005–06) and 67th (2010–11) rounds of the National Sample Survey (NSS) Enterprise Surveys, conducted by the Ministry of Statistics and Programme Implementation (MoSPI). For formal firms, we use the Annual Survey of Industries (ASI) from 2002–03 to 2009–10. For PMGSY roads construction information, we use the administrative data made available online through Online Monitoring and Management System (OMMS) which provides village-level information on rural roads constructed and upgraded under the program from its inception. While the PMGSY data is at the village-level, the smallest identifiable administrative unit in both the firms dataset is a district. We therefore construct exposure to PMGSY at the district level by computing the a) percentage of district rural population that gets exposed to a new rural road b) total length of the rural roads build in a district under the PMGSY program. We combine these measures with both the firms data to form

[2] https://www.ideasforindia.in/topics/productivity-innovation/formalisation-of-informal-manufacuturing-enterprises-in-india.html

two analytical samples.

Our empirical strategy exploits within-district variation in exposure to PMGSY road construction over time using a two-way fixed effects framework. Identification relies on the assumption that this variation is driven solely by the distribution of villages with populations above the program's eligibility threshold and is therefore plausibly quasi-random. A potential threat to this assumption would arise if road placement systematically deviated from the official population criterion and was instead influenced by other district-level characteristics correlated with firm performance. However, Aggarwal (2018) provide convincing evidence that the population-based eligibility rule was largely followed in the implementation of PMGSY.

Our results suggest that rural informal firms exposed to greater road construction experience significant growth: both turnover and expenditure increase, with the rise in turnover outpacing costs, leading to higher earnings before taxes (i.e., gross profits). We also find that firms in districts with greater road expansion hire more workers—particularly male workers—resulting in a higher overall wage bill. In addition, informal firms in these areas operate for longer hours each day and for more months in a year, suggesting increased business activity. These results are robust to controlling for a wide range of public goods that might influence road placement, as well as to an instrumental variable strategy that uses the PMGSY's population-based eligibility rule to instrument for exposure to rural roads. We also find that the effect of exposure to rural roads is heterogeneous by firm size. Our findings suggest that while small informal firms gain limited benefits from rural road expansion, larger firms (in terms of employment size) gain in terms of increased turnover and profit from improved connectivity.

While we find evidence of improved performance among small rural informal firms, our results do not show a comparable positive impact on turnover, gross profits, or employment for rural formal firms as a whole. These findings likely reflect that smaller and less capital-intensive informal firms often operate below capacity, and enhanced infrastructure can help

them scale up their operations.

To provide suggestive evidence that the documented positive effects on small unregistered firms are indeed attributable to the PMGSY roads, we show that firms with greater exposure to rural roads are significantly less likely to report transportation infrastructure as a major constraint to their enterprise growth. These firms are also less likely to cite competition from larger, well-established firms as a key concern, possibly due to an expanded catchment area enabled by improved connectivity. Importantly, we find no significant changes in the likelihood of firms reporting problems related to fuel shortages, power outages, labor availability, or access to credit. This strengthens the case that the observed improvements in firm outcomes are specifically driven by the rural roads program, rather than by other concurrent government initiatives.

1.1 Literature

Our paper contributes to several strands of literature. It contributes to the growing body of work examining how infrastructure affects firm performance in LMICs. Key contributions in this area include Ghani, Goswami and Kerr (2015); Datta (2012); Kailthya and Kambhampati (2022); Asher and Novosad (2020); Li, Wu and Chen (2017); Wu, Yu and Zhang (2023); Branco, Dohse, dos Santos and Tavares (2023); Perra, Sanfilippo and Sundaram (2024); Chatterjee, Lebesmuehlbacher and Narayanan (2021); Chaurey and Le (2022); Gibson and Olivia (2010) among others. Within this literature, our paper is most closely related to Asher and Novosad (2020), and Chaurey and Le (2022). Using variation from the PMGSY, Asher and Novosad (2020) find that while rural roads increase non-agricultural employment, they do not necessarily lead to firm growth within villages, as workers tend to migrate to external labor markets. In a related context, Chaurey and Le (2022), examine the effects of the Rashtriya Sam Vikas Yojana (RSVY)—a large rural infrastructure initiative in India—and find that improved infrastructure leads to a rise in the number of micro-enterprises and associated employment. However, consistent with our findings, they find no impact of the

program on formal firms. Our paper complements and extends these findings by providing novel evidence of the impact of PMGSY road construction on both informal and formal rural manufacturing enterprises, focusing on performance indicators such as output, employment, and productivity.

Our paper also contributes to the growing literature that evaluates the effects of PMGSY on a range of economic outcomes beyond firm performance. This literature documents mixed results. Shamdasani (2021) shows that improved roads encourage higher adoption of modern inputs and increased crop diversification. Aggarwal (2018) finds that road connectivity enhances market integration and agricultural technology adoption. Adukia, Asher and Novosad (2020) show that improved roads boost middle school enrollment, while Mukherjee (2012) cautions that older children may drop out to join the labor force. Roads have also been linked to better healthcare access (Aggarwal, 2021) and enhanced female mobility and autonomy (Nandwani and Roychowdhury, 2024), though not to increased female labor force participation. On the environmental front, Garg, Jagnani and Pullabhotla (2024) find that labor exits from agriculture due to rural road construction led to a spike in stubble burning, contributing to air pollution. Finally, examining the effects on entrepreneurship, Brahma and Soundararajan (2024) show that while new feeder roads lead to an increase in service-sector enterprises across all caste categories, gains in manufacturing entrepreneurship are concentrated among upper-caste groups. We extend this body of work by demonstrating that PMGSY also influenced the rural manufacturing landscape, especially for small, informal firms—thus highlighting a new and policy-relevant channel through which infrastructure programs can affect structural transformation in rural economies.

Our paper further speaks to the literature examining how market access through infrastructure improvements affects household welfare (Martin Wiegand, 2023; Sedai, Vasudevan, Pena and Miller, 2021; Mensah, Huchet-Bourdon and Latruffe, 2014; Ren Mu, 2011; Lokshin and Yemtsov, 2005; Jacoby, 2000; Jalan and Ravallion, 1998). Hine et al. (2019) provide a systematic review of the economic impact of rural roads across multiple countries, including

India. They find that road investments lead to poverty reduction and increased consumption, but marginal returns diminish in areas with pre-existing road networks. Our study complements this literature by shifting the focus from household-level welfare to firm-level dynamics in rural settings.

Finally, our paper connects to a broader literature examining how external shocks—economic, political, or institutional—affect firm behavior and performance. Del Prete, Di Maio and Rahman (2023); Utar (2024); Klapper, Richmond and Tran (2013) has looked at how conflict affects firms revenues and likelihood of survival, (De Loecker, Goldberg, Khandelwal and Pavcnik, 2016; Brandt, Van Biesebroeck, Wang and Zhang, 2017; Topalova and Khandelwal, 2011; Yu, 2015) study the impact of trade reforms on firm prices, markup and productivity. The impact of global and financial crises has been studied by Jean-Charles Bricongne (2012); Ourania Notta (2014); Claessens, Djankov and Xu (2000); Nisar Ahmad (2023). Görg and Mulyukova (2024) study the impact of special economic zones (SEZs) on firm productivity in India and find that while publicly owned SEZs experienced a decline in their productivity, there was an increase in productivity for privately owned SEZs. While not a shock per se, large-scale infrastructure rollouts such as PMGSY represent transformative policy interventions that may have similarly disruptive or enabling effects. Our paper contributes to this literature by analyzing how such interventions can influence the economic geography and viability of rural manufacturing enterprises.

The rest of the paper is structured as follows. Section 2 gives a background to the PMGSY policy. Section 3 provides details on the data used for analyses. Section 4 describes the empirical methodology. Section 5 discusses the results. The last section concludes.

2 Context

Launched in December 2000, the PMGSY was introduced against the backdrop of nearly 50% of India's 500,000 villages lacking connectivity to market centers. At the time, over

70% of India's population resided in rural areas, implying that approximately 300 million people lacked reliable access to the outside world. The primary objective of PMGSY was to construct all-weather roads connecting approximately 178,000 habitations with populations greater than 500 to the nearest market center, or to another village already linked to such centers. By December 2017, the program had successfully constructed more than 550,000 kilometers of rural roads, connecting over 159,000 habitations.

The rollout of PMGSY was staggered according to village population size, with priority given to larger villages. Villages with populations above 1,000 were targeted first, followed by those with populations above 500. In hilly and tribal states with low population density, villages with populations below 500 were also eligible under the initial phases. In other states, such villages were incorporated into the program in subsequent phases. The third and current phase of PMGSY focuses on last-mile connectivity to the remaining unconnected villages and the maintenance and upgrading of roads built in earlier phases.[3]

As of December 2024, the program had achieved substantial coverage, with 95% and 97% of sanctioned roadworks under Phase I and Phase II, respectively, completed. Currently, around 25,000 villages are undergoing road construction under the ongoing phase. With a total expenditure of USD 27 billion, PMGSY stands as one of the most extensive and expensive centrally sponsored infrastructure programs in India's rural development landscape.

3 Data

The data for the analysis comes from several sources, including administrative records, the Enterprise Survey (ES), the Annual Survey of Industries (ASI), and the Indian Census 2001.
[4]

[3]PMGSY - II was launched in the year 2013 and PMGSY - III was launched in the year 2019. These phases of the program were launched to extend rural roads to all remaining villages and hence population based eligibility rule was not followed for these phases.

[4]Note that while ProwessIQ is an alternative data source for formal firms, around 53.8% of the firms covered in Prowess come from urban cities like Ahmadabad (4.2%), Bangalore Urban (4.5%), Mumbai

3.1 Administrative Data on Road Construction

The administrative data on PMGSY is made publicly available through Online Management, Monitoring and Accounting System (OMMS) that monitors roads constructed and upgraded under PMGSY. The data provides annual village level information on whether the road has been constructed under the program, length of roads constructed, approval, and completion date of the road, upgradation of existing roads in addition to village population and road connectivity status at the start of the program.

3.2 Enterprise Survey

We use data from the 62nd and 67th rounds of the Unorganised Manufacturing Enterprise Survey (ES), conducted in 2005–06 and 2010–11 respectively by the National Sample Survey (NSS). The 62nd round employed Schedule 2.2 and covered 82,897 firms, while the 67th round used an updated Schedule 2.34 and sampled 334,447 firms, both designed to capture detailed information on the activities and performance of unorganised manufacturing firms.[5]

The surveys include comprehensive information on firm characteristics such as operating expenses, receipts, employment, worker compensation, and responses to self-reported questions on operational challenges.[6] We use this data to construct information on the expansion status of enterprise, total expenditure incurred, total receipts which serve as a proxy for firm turnover, gross value addition which is used as a proxy for firm profit, employment generated, for how long the firm is operational, and wage bill paid by the firms.

(18.7%), Pune (4%), Delhi (17.1%) and Chennai (5.3%). Since our purpose is to examine the effect of rural roads on rural firms, we chose not to proceed with this dataset.

[5]The Enterprise Survey conducted by the Ministry of Statistics and Programme Implementation (MoSPI) includes various surveys over time—such as the Informal Non-Agricultural Enterprises Survey, the Survey on Unincorporated Non-Agricultural Enterprises, and the Unorganised Manufacturing Enterprises Survey. Initially, Schedule 2.2 was used until 2005, after which Schedule 2.34 was introduced for the 2010–11 round. The Annual Survey of Unincorporated Sector Enterprises (ASUSE), which came later, builds on this structure with additional modules such as LSU (Household Listing) and ESU (Enterprise Listing). Since our study period spans 2001 to 2010, we rely on data from both Schedule 2.2 and 2.34.

[6]Survey methodologies and documentation are available at NSS 62nd Round and NSS 67th Round.

3.3 Annual Survey of Industries

We use the data from the Annual Survey of Industries (ASI) to examine the effect of rural roads construction program on formal sector manufacturing firms. These are annual surveys that are conducted by the MoSPI to collect information on registered manufacturing establishments in India. [7] The primary unit of enumeration in the survey is a factory in the case of manufacturing industries and the survey covers all factories registered under Sections 2(m)(i) and 2(m)(ii) of the Factories Act, 1948. ASI collects data on various aspects of manufacturing establishments including input, output, value added, employment, assets etc. For defining the manufacturing sector, we follow the MoSPI definition.[8] We obtain a repeated cross-section of the plant-level data from 2002-03 to 2009-10 and restrict ourselves to rural plants for our analysis. This gives us a sample of around 10,000 plants every year. Note that ASI data is reported for a financial year: for example, the 2002-03 year refers to the period April 1, 2002 to March 31, 2003.

3.4 Population Census

We make use of population census of 2001 to compute district level demographic and economic indicators that are used as controls in regressions in the paper. The indicators include the proportion of Scheduled Caste (SC) and Scheduled Tribe (ST) populations, literacy rate,

[7] See the survey methodologies in detail here - ASI Microdata

[8] According to NIC 1998, the manufacturing sector is covered under the codes 15 -manufacture of food products and beverages, 16 - manufacture of tobacco products, 17 - manufacture of textiles, 18 - manufacture of wearing apparel; dressing and dyeing of fur, 19 - tanning and dressing of leather;manufacture of luggage, handbags saddlery, harness and footwear, 20 - manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plating materials, 21 - manufacture of paper and paper products, 22 - publishing, printing and reproduction of recorded media, 23 - manufacture of coke, refined petroleum products and nuclear fuel, 24 - manufacture of chemicals and chemical products, 25 - manufacture of rubber and plastic products, 26 - manufacture of other non-metallic mineral products, 27 - manufacture of basic metals, 28 - manufacture of fabricated metal products, except machinery and equipments, 29 - manufacture of machinery and equipment n.e.c., 30 - manufacture of office, accounting and computing machinery, 31 - manufacture of electrical machinery and apparatus n.e.c., 32 - manufacture of radio, television and communication equipment and apparatus, 33 - manufacture of medical, precision and optical instruments, watches and clocks, 34 - manufacture of motor vehicles, trailers and semi-trailers, 35 - manufacture of other transport equipment, 36 - manufacture of furniture; manufacturing n.e.c., and 37 - recycling.

and work participation rate.

3.5 Analytical Samples

Our final analytical samples are constructed by combining the administrative data on PMGSY roads construction with the unit level data from ES (ES analytical sample from hereon) and ASI (ASI analytical sample from hereon). While the PMGSY data are at the village-year level, the lowest identifiable administrative unit in both ES and ASI is a district. Therefore, similar to Aggarwal (2018) and Nandwani and Roychowdhury (2024), we aggregate the village-level information in PMGSY at the district-level. From both the constructed samples, we exclude union territories and all the North-Eastern states other than Assam. The ES analytical sample consists of 182,127 observations and the ASI analytical sample has 82,775 observations.

Table 2 presents the summary statistics for informal firms. On average, an unregistered firm is operational for 11 months in a year and 7 hours in a day. Around 27% of firms reported to be expanding in the last year, whereas 13% of the firms contracted in the same time period. An average firm makes a gross profit of 12,500 rupees in a month with a turnover of INR 47,100. Note that all these numbers have been corrected for price changes over time using Wholesale Price Index (WPI) data taking 2004-05 as the base year.[9] In the last month, an average firm employed 2.9 workers and incurred an average wage bill of INR 22,700.[10] Table 3 presents the summary statistics for formal firms. Formal firms like informal ones are also operational for 11 months in a year but they are much bigger. An average formal firms earns a gross profit of 130 million and hires 170 workers and pays a wage bill of 11 million in a year. The summary statistics for control variables are reported in Appendix tables A1 and A2.

Note that many studies that evaluate the impact of access to transportation infrastructure

[9]The data source for WPI is MoSPI.

[10]All the financial variables used in our analysis have been winsorized at 1% level to mitigate the influence of outliers.

on firm outcomes use total factor productivity as an indicator of firm performance. While total factor productivity is an important and informative metric to assess firm performance, we are constrained by the lack of panel data for formal and informal firms which precludes us from using the standard approaches proposed by Ackerberg, Caves and Frazer (2015); Levinsohn and Petrin (2003); Olley and Pakes (1992) to calculating total factor productivity.

Table 4 presents the summary statistics for our key explanatory variables, shows that on average 5.8% of rural population in a district got exposed to rural roads in the year 2004 which increased to 15% by 2010. The average exposure to rural roads for our sample period (for unregistered firms) is 11%. The length of the roads constructed under PMGSY in a district was 85.5 km in 2004 which increased to 489.5 km in 2010.

4 Empirical Methodology

We use a two-way fixed effect framework to estimate the impact of rural road expansion under PMGSY on firm-level outcomes such as expenses, receipts, employment, and wages. Specifically, we estimate the following regression equation:

$$Y_{ijdt} = \alpha + \beta ExposedPop_{d(t-1)} + \gamma X_{ijdt} + \delta(Z_d \times t) + \theta_j + \mu_d + \eta_t + \varepsilon_{ijdt} \quad (1)$$

Here Y_{ijdt} is the outcome for firm i in industry j located in district d, in year t. The key independent variable, $ExposedPop_{d(t-1)}$, is the percentage of the (census 2001) rural population[11] in district d exposed to new rural roads constructed under PMGSY upto year $t-1$. In alternative specifications, we compute exposure to PMGSY by calculating the length of rural roads built in a district. X_{ijdt} is a vector of firm-level controls. For informal firms, this includes an indicator for ownership type of the firm,[12] whether the firm engages in multiple

[11]The Population census of 2001 was used to set the population threshold under the PMGSY program which is why we calculate district population exposure to PMGSY roads as a percentage of 2001 district (rural) population.

[12]The categories include Proprietary (Male = 1, Female = 2), Partnerships (Within household = 3, Outside household = 4), Co-operative Society (5) – Available only in the NSS 62nd round, Limited Company (6) –

activities, whether the firm maintains formal financial accounts, whether the firm operates inside the household premises, whether the firm has received any form of external assistance, whether the firm is officially registered, whether the survey respondent is owner/partner, manager, or other worker in the enterprise and finally an indicator for whether the survey respondent was cooperative during the survey. For the regressions conducted for formal firms, the firm levels controls are an indicator for firm's legal structure,[13] indicator for the ownership structure of the firm.[14] Also included are the year in which the firm started its production activities and the number of reporting units for which the data has been collected.

All our regression specifications include industry fixed effects (θ_j) to partial out industry specific characteristics that affect firm outcomes,[15] district fixed effects (μ_d) to partial out district time-invariant socioeconomic characteristics (like level of public goods provision, distance from state capital) that are likely to affect firm outcomes and may be determinants of placements of roads under PMGSY, year fixed effects (η_t) to account for year specific shocks to firm outcomes that are common to all districts. We also include an interaction between district characteristics (population proportion of SC, ST, literacy rate and work participation rate) and survey year ($Z_d \times t$) in all our regression specifications to allow for heterogeneous trends across districts based on their initial conditions.

Our coefficient of interest is β which captures how firm outcomes respond to variation in the share of the population within a district gaining access to roads under the PMGSY program. This variation is primarily driven by differences in the distribution of unconnected

Available only in the NSS 62nd round, Self-Help Group (10) – Available only in the NSS 67th round, Trusts (11) – Available only in the NSS 67th round, and, Others (9)

[13]The categories include individual proprietorship, joint family ownership, partnership, public limited company, private limited company, government departmental enterprise, public corporation by Special Act, Co-operative society, and Others which Includes Trusts, Wakf Boards, etc.

[14]The categories include whether it is a Wholly Central Government, Wholly State and/or Local Government, Central/State/Local Government, Joint Sector Public, Joint Sector Private, and, Wholly Private Ownership enterprise

[15]We use the two-digit National Industrial Classification (NIC2) codes to identify the firm's industry. The 2005-06 round of ES used the 2004 NIC codes, the ES conducted in the year used the 2008 NIC codes. Similarly, the ASI data used the NIC 1998 codes for 2002-03 and 2003-04 rounds, NIC 2004 for the rounds conducted between 2004-05 and 2007-08 and NIC 2008 for the rounds conducted in the years 2008-09 and 2009-10. To ensure consistency in the codes over the years, we use fuzzy matching to group industries with similar descriptions between NIC 1998 and 2008.

village sizes across districts and is plausibly quasi-random.[16] Although eligibility was, in principle, determined solely by village population thresholds, it is conceivable that other factors may have influenced road placement—some of which could also affect firm performance. To address this concern, we note two key points. First, the inclusion of time-varying firm-level covariates helps control for time-varying village-level unobservables, insofar as they are correlated with firm characteristics. Second, as discussed earlier, Aggarwal (2018) finds a sharp discontinuity in the probability of road construction around the population eligibility thresholds of 500 and 1000, and shows that pre-existing public goods provision (as of 2001) is uncorrelated with the likelihood of road construction by 2011. These findings bolster our confidence that the program's population rule was, in fact, implemented as intended. Nonetheless, in Section 5.2, we conduct an extensive set of robustness checks to address potential concerns regarding non-random selection into treatment. Throughout, standard errors are clustered at the district level.

5 Results and Discussion

5.1 Main Results

The results obtained from the estimation of Equation 1 for informal firms are reported in Tables 5-7. In each table, Panel A reports results of the specifications that use population exposed to new rural roads under PMGSY as the measure of PMGSY exposure while panel B reports results of the specifications that use the length of new rural roads built under PMGSY as the measure of PMGSY exposure. The column headings denote the outcome variables used in the regression for which the results are reported in that column.

Table 5 provides evidence of PMGSY's effect on turnover, expenditure and profits for

[16]Since only eligible villages were targeted, the variation in district-level exposure arises from the changing distribution of unconnected villages over time. As long as the PMGSY adhered to its population-based eligibility rule at the village level, district-level variation is driven by village-level implementation and can be considered exogenous to firm outcomes.

informal enterprises. The estimates of the coefficient of interest reported in Panel A, Column 1 suggest that a 10% increase in the population exposed to rural roads in a district increases monthly turnover of informal rural firms by ₹49,000. Quantitatively, this amounts to more than a 100% increase in the turnover as compared to the average monthly turnover of informal firms in our sample. This is also accompanied by an increase in expenses incurred by these firms. However, the increase in turnover is higher than the expenditure increase which increases gross profit of the firms by ₹12,000 in the last month (column 3 in panel A). Given that the monthly profit of an average informal firm in our sample is ₹12,500, this indicates a substantial increase in profits post-PMGSY roads expansion. The estimates presented in Panel B show that our results are robust to considering an alternative way of measuring PMGSY exposure. Specifically, a 10% increase in length of rural roads built in a district improves unregistered firms turnover by ₹1910, expenditure by ₹1390 and by ₹470.

Table 6 presents the results for employment particular for informal sector firms. It shows that the increase in profit is accompanied by increase in employment and wage bill of informal firms. A 10% increase in population exposed to rural roads results in increase in employment by 0.2 which is a 6% increase as compared to the average. Though this coefficient is not statistically significant in Panel A, it becomes significant in all robustness tests that we perform. Further, the coefficient of employment is significant in Panel B where we measure exposure to PMGSY using the length of the roads built under the program. The increase in employment is primarily driven by increase in males being hired by these rural informal firms as we do not see robust increase female employment. This result is consistent with the finding of Nandwani and Roychowdhury (2024) who document that female labor participation is not positively affected by PMGSY. Increase in labor hired also increases the wage bill of these firms—a 10% increase in population exposed to rural roads increases wage bill of rural firms by ₹35,000.

The above results are corroborated by the firm's self-reported opinion on firm's expansion and operational status. Table 7 shows that a 10% exposure to rural roads increases the

number of months a firm reports being operational for by 0.5 months which implies that exposed firms are operational for 11.6 months on average. This result is robust to using alternative way of measuring exposure to PMGSY (Panel B). Columns 3 and 4 in panels A and B suggest that exposure to rural roads increases the likelihood of rural firms reporting that they expanded in the last month and decreases the likelihood of their contraction

Next, we examine the impact of PMGSY rural road expansion on formal manufacturing enterprises. The results, presented in Tables 8 and 9, indicate that unlike their informal counterparts, formal rural firms do not experience meaningful gains in key performance metrics such as turnover, gross profits, or employment. In fact, Panel B of Table 8 shows a decline in total expenditures among formal firms with increased exposure to rural road construction, without any corresponding increase in profitability. This suggests that rural connectivity may not materially improve the operational efficiency or market reach of these firms.

One possible explanation is that formal firms in rural areas may already be operating near capacity and have access to essential infrastructure, limiting the marginal benefits from new road connectivity. Moreover, formal enterprises often face institutional and structural rigidities—such as higher regulatory compliance costs, immobile capital, and more rigid input–output relationships—that constrain their ability to adjust quickly or scale operations in response to improved external conditions. In contrast, informal firms tend to operate below capacity, face more acute first- and last-mile constraints, and have greater flexibility to leverage reductions in transportation and transaction costs brought about by road expansion. Our results are consistent with Chaurey and Le (2022), who document that only informal firms benefit from a rural transportation infrastructure maintenance program in India.

5.2 How Rural Road Expansion Benefits Informal Firms

We use information on self-reported problems faced by small unregistered rural establishments to understand how exposure to PMGSY has affected these firms. The results are

presented in Table 10. Our results indicate that with increase in exposure to the PMGSY roads, firms are less likely to face problems related to transportation infrastructure. This result strengthens our claim that access to market centers via rural roads is aiding the small rural establishment in their growth. We also find that these firms are less likely to report facing difficulties in advertising, customer reach, or product distribution and competition from larger establishments. This could be due to increased access to larger markets after PMGSY. Assuringly, we do not find increased exposure to rural roads to be affecting problems stated by firms that are not expected to be affected by rural roads connectivity. For example, we do not find fuel, electricity, labor or credit related problems to be systematically different in areas with higher rural roads construction.

5.3 Robustness Checks

5.3.1 Inclusion of Controls for Public Goods

While we argue in the introduction that the placement of PMGSY roads was governed by a population-based rule, rendering the exposure to rural roads quasi-random, it is plausible that the actual road placement was influenced by the presence of other public goods in the village, making it potentially endogenous. We address this concern by examining the sensitivity of our coefficient of interest to the inclusion of various public goods present in villages just before the program's introduction.

Specifically, using the village directories from the 2001 Population Census, we calculate the percentage of villages in each district that had commercial banks, post offices, telephone booths, and power supply. Additionally, we compute the average number of rural primary schools, high schools, adult literacy centres, and primary health centres at the district level. We interact these district-level public goods with the program wave and include them as controls in our main specification. Results, presented in Tables A3, A4, and A5, show that although the coefficients decrease slightly in magnitude, they remain statistically and economically significant. This suggests that the results reported in Tables 5, 6, and 7 are

unlikely to be fully driven by road placement influenced by the presence of other public goods.

5.3.2 Oster Bounds

While the stability of coefficients is considered as a suggestive evidence against omitted variables bias, (Oster, 2019) argues that movement in R^2 also has to be considered and that the movement in coefficient alone is insufficient to indicate the extent of omitted variable bias. The paper derives the relationship between coefficient movement, R^2 movement and omitted variable bias under the assumption that selection on unobservables is proportional to selection on observables. The author then calculates the bias adjusted estimate of the coefficient, given a value of proportionality between observables and unobservables (δ) and the value of R^2_{max}. R^2_{max} is the R^2 from a hypothetical regression in which all observables and unobservables have been controlled for. Following (Oster, 2019), we compute the bias adjusted coefficients after considering δ to be 1 (i.e. the selection on observables has to be as important as selection on unobservables) and R^2_{max} to be 1.3 times the R^2 of the controlled regression wherein all observable controls (including all public good controls) have been added. The results, reported in Tables A6, A7 and A8, suggest that bias bias-adjusted coefficient is close to the coefficient from the controlled regression, bolstering our claim that our coefficients are unlikely to be driven by omitted variables bias. Further, the absolute value of δ for which the treatment effect would completely become 0 is well above 1 for most of the outcomes.

5.3.3 Instrumental Variable Estimates

We further allay potential concerns that the construction under PMGSY may have been influenced by factors that could confound our estimates. For instance, if districts with greater lobbying power were able to secure more roads in addition to meeting the population criteria, this could undermine the causal interpretation of our findings. We address this potential

endogeneity by using the population rule of PMGSY roll-out as the instrumental variable. Specifically, the instrument is the percentage of the rural population (as per the 2001 Census, at the start of the program) residing in villages with populations above 500—the PMGSY eligibility threshold. If the rollout adhered to the population rule, we expect this measure to be positively correlated with actual exposure to roads. Moreover, after controlling for district- and firm-level characteristics, this instrument is unlikely to directly influence firm outcomes, plausibly satisfying the exclusion restriction.

Our IV-2SLS results are reported in Tables A9, A10 and A11. Note that since the instrument is time invariant, we interact it with the year of the survey and use this generated variation as the instrument. Table A11 shows turnover, expenditure and profits of unregistered go up with an increase in the exposure to rural roads and the increase is much higher as compared to the increase reported in Table 5. This is plausible as the IV-2SLS results identify the local average treatment effects (LATE). Tables A10 and A11 further confirm the two-way fixed effects results as the IV-2SLS coefficients suggest that informal firms hire more workers, experience an increase in wage bill and are less likely to report to be contracting in the last one year.[17]

5.3.4 Exact Randomization Inference Test

We implement the test of exact randomisation test—a standard placebo exercise—to verify that our estimated effects are not driven by spurious correlations. Specifically, we randomly assign treatment values between 0 and 100 to district-survey year and repeatedly re-estimate regression equation 1. We perform this placebo exercise 1,000 times, recording the estimated coefficient and standard error in each iteration. In an ideal scenario, these placebo coefficients should be statistically insignificant and smaller in magnitude compared to our main estimates in Tables 5 to 7. Figure A1 presents the distribution of these coefficients. Reassuringly, the placebo estimates are tightly centred around zero and exhibit very small magnitudes.

[17]We do not report the IV-2SLS results for formal firms as their coefficients are not significant in the TWFE regressions.

Since the timing of PMGSY rollout coincided with the Golden Quadrilateral (GQ) highway project which began in the year 2001, there could be a concern that our results may be partially driven by rural firms proximity to highways. However, we are confident that highway access is unlikely to drive our results as the highway construction under the GQ project was not based on the population criteria like PMGSY. Additionally, there is no documented evidence of coordination between the PMGSY and the GQ program. See Adukia et al. (2020) for discussion.

We also confirm that our results (not reported) are robust to exclusion district-time varying controls that can potentially be endogenous and thus confound the effect of our treatment.

5.4 Heterogeneity by Firm Size

We have shown that PMGSY significantly improves the performance of rural informal enterprises. However, the magnitude and nature of these effects may vary by firm size. Theoretical and empirical research suggests that smaller informal firms, particularly those with only one or two workers including the owner, often operate as survivalist enterprises rather than growth-oriented ventures (Berner et al., 2012; La Porta and Shleifer, 2014). These microenterprises typically face severe constraints in accessing credit, labor, and markets, and may lack the capacity to leverage improved infrastructure effectively. In contrast, relatively larger informal firms—those employing several non-household workers or operating outside the home—are more likely to be responsive to reductions in transportation and transaction costs brought about by better connectivity.

Consistent with this hypothesis, our analysis—presented in Tables A12, A13, and A14—reveals substantial heterogeneity in treatment effects by firm size. We use three complementary approaches to define firm size. First, in Panels A and B, we divide firms into those with two workers (including the owner) and those with more than two. Second, Panels C and D classify firms based on the employment size distribution, comparing the top 10% to the bottom

90%. Finally, Panels E and F distinguish between firms operating within the owner's household premises and those located externally, with the former expected to be smaller, often family-run units.

Across all specifications, we find that very small firms benefit only marginally from rural road expansion, with modest gains in employment and operating hours. In contrast, larger informal firms exhibit significant improvements across multiple performance dimensions—including turnover, profits, wage bills, male employment, and likelihood of reporting expansion. These results suggest that improved connectivity disproportionately benefits informal firms that are better positioned to scale up, possibly by facilitating access to broader input and output markets, reducing search frictions, or enabling more regular operations.

6 Conclusion

The paper examines how rural manufacturing enterprises that are constrained due to limited enabling infrastructure respond to increased rural roads constructed under the Government of India flagship rural infrastructure PMGSY program. While the existing literature has extensively documented that access to highways improve productivity and output of formal firms, there is limited understanding of how rural informal manufacturing activity, which accounts for 75% of manufacturing employment, is shaped by rural connectivity. We use the PMGSY roll-out rule which prioritised larger villages for roads placement for generating quasi-random variation in exposure of the rural population to roads in a district.

Our empirical results point that improved rural connectivity significantly boosts the turnover, profits, employment, and operational activity of informal firms, more for larger informal than smaller. In contrast, formal firms see no comparable gains, suggesting that capacity-constrained enterprises are better positioned to benefit from enhanced local infrastructure. We provide evidence that these improvements are driven by reductions in infrastructure-related constraints and lower barriers to market access.

The absence of measurable gains for formal enterprises may reflect their pre-existing access to transport infrastructure, greater reliance on national or export markets, and operational rigidities that limit rapid scaling in response to local connectivity improvements. In contrast, informal firms—particularly those serving regional markets—face more acute first- and last-mile constraints, making them more responsive to rural road expansion. These findings suggest that while programs like PMGSY can substantially boost informal sector activity, complementary measures—such as facilitating credit access, easing regulatory constraints, and improving integration with larger logistics networks—may be necessary to enable formal firms to fully capitalize on infrastructure investments.

References

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer, "Identification properties of recent production function estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Adukia, Anjali, Sam Asher, and Paul Novosad, "Educational Investment Responses to Economic Opportunity: Evidence from Indian Road Construction," *American Economic Journal: Applied Economics*, January 2020, *12* (1), 348–76.

Aggarwal, Shilpa, "Do rural roads create pathways out of poverty? Evidence from India," *Journal of Development Economics*, 2018, *133*, 375–395.

_ , "The long road to health: Healthcare utilization impacts of a road pavement policy in rural India," *Journal of Development Economics*, 2021, *151*, 102667.

Asher, Sam and Paul Novosad, "Rural Roads and Local Economic Development," *American Economic Review*, March 2020, *110* (3), 797–823.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian, "On the road: Access to transportation infrastructure and economic growth in China," *Journal of Development Economics*, 2020, *145*, 102442.

Berner, Erhard, Georgina Gomez, and Peter Knorringa, "'Helping a Large Number of People Become a Little Less Poor': The Logic of Survival Entrepreneurs," *European Journal of Development Research*, 2012, *24* (3), 382–396.

Brahma, Ananyo and Vidhya Soundararajan, "Do Rural Roads Promote Inclusive Entrepreneurship?," *Available at SSRN*, 2024.

Branco, Catarina, Dirk C Dohse, João Pereira dos Santos, and José Tavares, "Nobody's gonna slow me down? The effects of a transportation cost shock on firm performance and behavior," *Journal of Urban Economics*, 2023, *136*, 103569.

Brandt, Loren, Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang, "WTO accession and performance of Chinese manufacturing firms," *American Economic Review*, 2017, *107* (9), 2784–2820.

Chatterjee, Santanu, Thomas Lebesmuehlbacher, and Abhinav Narayanan, "How productive is public investment? Evidence from formal and informal production in India," *Journal of Development Economics*, 2021, *151*, 102625.

Chaurey, Ritam and Duong Trung Le, "Infrastructure maintenance and rural economic activity: Evidence from India," *Journal of Public Economics*, 2022, *214*, 104725.

Claessens, Stijn, Simeon Djankov, and Lixin Colln Xu, "Corporate Performance in the East Asian Financial Crisis," *The World Bank Research Observer*, 02 2000, *15* (1), 23–46.

Datta, Saugato, "The impact of improved highways on Indian firms," *Journal of Development Economics*, 2012, *99* (1), 46–57.

Fontagné, Guillaume Gaulier Daria Taglioni Vincent Vicard Jean-Charles Bricongne Lionel, "Firms and the global crisis: French exports in the turmoil," *Journal of International Economics*, 2012, *87* (1), 134–146. Symposium on the Global Dimensions of the Financial Crisis.

Garg, Teevrat, Maulik Jagnani, and Hemant K. Pullabhotla, "Rural Roads, Farm Labor Exits, and Crop Fires," *American Economic Journal: Economic Policy*, August 2024, *16* (3), 420–50.

Ghani, Ejaz, Arti Grover Goswami, and William R Kerr, "Is India's manufacturing sector moving away from cities?," Technical Report, National bureau of economic research 2012.

_ , _ , and William R. Kerr, "Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing," *The Economic Journal*, 04 2015, *126* (591), 317–357.

Gibson, John and Susan Olivia, "The effect of infrastructure access and quality on non-farm enterprises in rural Indonesia," *World Development*, 2010, *38* (5), 717–726.

Görg, Holger and Alina Mulyukova, "Place-based policies and firm performance: Evidence from Special Economic Zones in India," *European Economic Review*, 2024, *165*, 104752.

Hine, John, Manu Sasidharan, Mehran Eskandari Torbaghan, Michael Burrow, and Kristianto Usman, "Evidence of the Impact of Rural Road Investment on Poverty Reduction and Economic Development," 7 2019.

Jacoby, Hanan G., "Access to Markets and the Benefits of Rural Roads," *The Economic Journal*, 2000, *110* (465), 713–737.

Jalan, Jyotsna and Martin Ravallion, "Are there dynamic gains from a poor-area development program?," *Journal of Public Economics*, 1998, *67* (1), 65–85.

Kailthya, Subham and Uma Kambhampati, "Road to productivity: Effects of roads on total factor productivity in Indian manufacturing," *Journal of Comparative Economics*, 2022, *50* (1), 174–195.

Klapper, Leora F, Christine Richmond, and Trang Thu Tran, "Civil conflict and firm performance: evidence from Cote d'Ivoire," *World Bank Policy Research Working Paper*, 2013, (6640).

Koomen, Menno Pradhan Christopher Edmonds Martin Wiegand Eric, "The Impact of Road Development on Household Welfare in Rural Papua New Guinea," *The Journal of Development Studies*, 2023, *59* (6), 933–953.

Levinsohn, James and Amil Petrin, "Estimating production functions using inputs to control for unobservables," *The review of economic studies*, 2003, *70* (2), 317–341.

Li, Zhigang, Mingqin Wu, and Bin R Chen, "Is road infrastructure investment in China excessive? Evidence from productivity of firms," *Regional Science and Urban Economics*, 2017, *65*, 116–126.

Loecker, Jan De, Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik, "Prices, markups, and trade reform," *Econometrica*, 2016, *84* (2), 445–510.

Lokshin, Michael and Ruslan Yemtsov, "Has rural infrastructure rehabilitation in Georgia helped the poor?," *The World Bank Economic Review*, 2005, *19* (2), 311–333.

Mensah, Emmanuel Joseph, Marilyne Huchet-Bourdon, and Laure Latruffe, "Infrastructure Access and Household Welfare in Rural Ghana," *African Development Review*, 2014, *26* (3), 508–519.

Mobarek, Moodhi Raid Nisar Ahmad Asma, "Impact of global financial crisis on firm performance in UK: Moderating role of ESG, corporate governance and firm size," *Cogent Business & Management*, 2023, *10* (1), 2167548.

Mukherjee, Mukta, "Do better roads increase school enrollment? Evidence from a unique road policy in India," *Evidence from a Unique Road Policy in India (August 28, 2012)*, 2012.

Nandwani, Bharti and Punarjit Roychowdhury, "Rural Road Infrastructure and Women's Empowerment in India," *The World Bank Economic Review*, 10 2024, p. lhae048.

Notta, Aspasia Vlachvei Ourania, "The Impact of Financial Crisis on Firm Performance in Case of Greek Food Manufacturing Firms," *Procedia Economics and Finance*, 2014, *14*, 454–460. International Conference on Applied Economics, ICOAE 2014.

Olley, Steven and Ariel Pakes, "The dynamics of productivity in the telecommunications equipment industry," 1992.

Oster, Emily, "Unobservable selection and coefficient stability: Theory and evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

Perra, Elena, Marco Sanfilippo, and Asha Sundaram, "Roads, competition, and the informal sector," *Journal of Development Economics*, 2024, *171*, 103339.

Porta, Rafael La and Andrei Shleifer, "Informality and Development," *Journal of Economic Perspectives*, 2014, *28* (3), 109–126.

Prete, Davide Del, Michele Di Maio, and Aminur Rahman, "Firms amid conflict: Performance, production inputs, and market competition," *Journal of Development Economics*, 2023, *164*, 103143.

Sedai, Ashish Kumar, Ramaa Vasudevan, Anita Alves Pena, and Ray Miller, "Does reliable electrification reduce gender differences? Evidence from India," *Journal of Economic Behavior Organization*, 2021, *185*, 580–601.

Shamdasani, Yogita, "Rural road infrastructure agricultural production: Evidence from India," *Journal of Development Economics*, 2021, *152*, 102686.

Topalova, Petia and Amit Khandelwal, "Trade liberalization and firm productivity: The case of India," *Review of economics and statistics*, 2011, *93* (3), 995–1009.

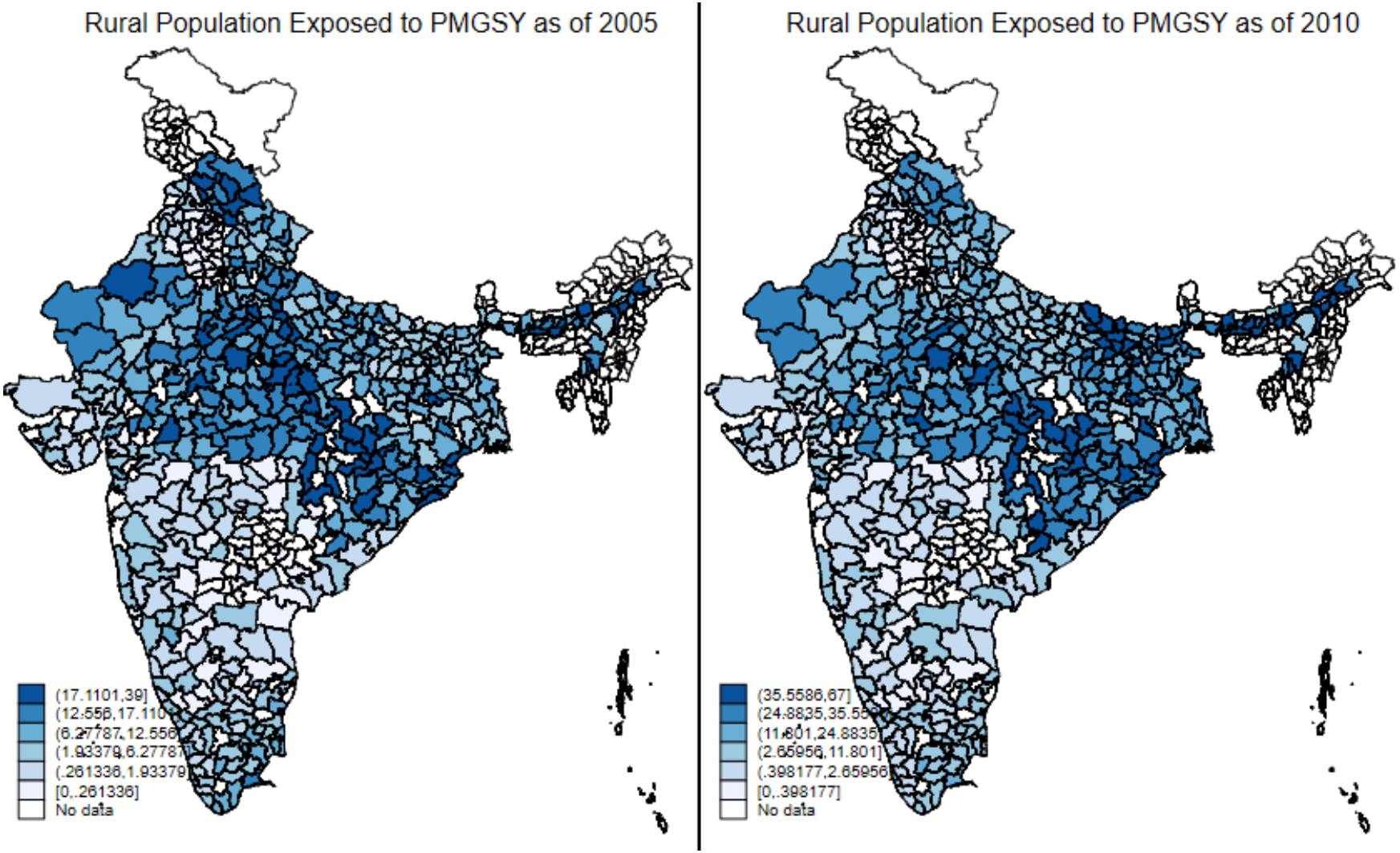
Utar, Hâle, "Firms and labor in times of violence: Evidence from the mexican drug war," *The World Bank Economic Review*, 2024, p. lhae037.

van de Walle Ren Mu, Dominique, "Rural Roads and Local Market Development in Vietnam," *The Journal of Development Studies*, 2011, *47* (5), 709–734.

Wu, Mingqin, Linhui Yu, and Junsen Zhang, "Road expansion, allocative efficiency, and pro-competitive effect of transport infrastructure: Evidence from China," *Journal of Development Economics*, 2023, *162*, 103050.

Yu, Miaojie, "Processing trade, tariff reductions and firm productivity: Evidence from Chinese firms," *The Economic Journal*, 2015, *125* (585), 943–988.

Fig 1: Distribution of Population proportion exposed to PMGSY for 2005 and 2010.



Source: Coordinates from Survey of India and PMGSY exposure from administrative record

Table 1a: Description of outcome variables

Label	Description	Source
Turnover	Turnover is the sum of all the receipts (in INR) that a firm receives. For ES, the reference period is the last 30 days from the survey, and for ASI, the reference period is the accounting year.	ES, ASI
Expenditure	Expenditure is the sum of spending on all the inputs (in INR) that a firm has incurred. For ES, the reference period is the last 30 days from the survey, and for ASI, the reference period is the accounting year.	ES, ASI
Profit	Profit (in INR) is the gross income of the firm. For ES, the reference period is the last 30 days from the survey and for ASI, the reference period is the accounting year.	ES, ASI
Total employment	For ES, it includes working owners, formally hired labor, informally hired labor and helpers during the last 30 days. For ASI, it reports individuals employed in the accounting year including workers employed directly, employed through contracts, supervisory and managerial staff, unpaid family/proprietor members, and others.	ES, ASI
Wage Bill	Wage bill includes all compensation paid (in INR) to all workers. This includes salary, daily payments (for informal sector), contributions to insurance/provident fund/social security, and any bonuses. For ES, the reference period is the last 30 days from the survey, and for ASI, the reference period is the accounting year.	ES, ASI
Hours in a day	Average working hours in a day. The reference period is the last 30 days.	ES
Month in a year	The months during which the firm was operational in the last 365 days for informal firms. "Month" is a continuous period of 30 days, including scheduled holidays.	ES
Status of firm: Expanding	Reports whether the informal firm experienced expansion in the last 365 days.	ES
Status of firm: Contraction	Reports whether the informal firm experienced contraction in the last 365 days.	ES

Table 1b: Description of explanatory variables and controls

Label	Description	Source
	*Explanatory va*riables	
$ExposedPop_{d(t-1)}$	Percentage of rural population (based on Census 2001) in a district that is exposed to new all-weather roads under PMGSY. If a firm is observed in 2005, then the corresponding exposure percentage is considered until 2004.	Administrative data
$Log(KM\ of\ roads_{d(t-1)})$	Log of kilometers of new all-weather roads built under PMGSY. If the firm is observed in 2005, the log of kilometers of roads built until 2004 is considered.	SHRUG 2.1
	Controls	
Informant Type	Indicates whether the respondent is an Owner/Partner, Manager, or Others/Working Staff.	ES
Response Type	Indicates whether the respondent is cooperative and capable, cooperative but not capable, busy but responded, reluctant, and Others.	ES
Mixed activity	Indicates whether the informal firm is engaged in mixed activity. For example, a bakery selling cold drinks, or a rice mill selling sugar.	ES
Ownership Type	Indicates the ownership type of firms. It can be: Proprietary (male), Proprietary (female), Partnership with members of the same household, Partnership with members not all from the same household, Cooperative society, Limited company (outside public sector), Self-help group, Trusts, and Others.	ES
Account maintained	Indicates whether informal firms maintain accounts or not.	ES
Location	Location of the firm: Outside household premises or Within household premises.	ES
Registered under any authority	If the informal firm is registered under any authority (e.g., Silk Board, Municipal Corporation, Panchayat, Shop and Establishment Act, ESI Act, VAT/Sales Tax Act).	ES
Any assistance received	Indicates whether informal firms received any assistance (e.g., subsidy, training, government/non-government schemes).	ES
Organization type	Captures the organization type of formal firms. It can be: Individual Proprietorship, Joint Family (HUF), Partnership, Public Limited Company, Private Limited Company, Government Departmental Enterprise, Public Corporation by Special Act, Cooperative Society, and Others (including Trusts, Wakf Board, etc.).	ASI
Ownership type	Records the type of ownership of formal firms. It can be: Wholly Central Government, Wholly State and/or Local Government, Central Government and State and/or Local, Joint Sector Public, Joint Sector Private, Wholly Private Ownership, and Others.	ASI
Count of units in case of joint returns	Counts the number of firms that filed joint returns.	ASI

Table 2: Summary statistics: Informal sector outcomes

	Pooled Sample			(2005-06)		(2009-10)	
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Obs	Mean
Turnover (in lakhs)	83,306	0.471	2.756	37,004	0.744	46,302	0.254
Expenditure (in lakhs)	82,916	0.336	2.212	36,749	0.564	46,167	0.155
Profit (in lakhs)	83,265	0.125	0.547	36,907	0.167	46,358	0.091
Total employment	83,790	2.99	4.988	37,415	2.82	46,375	3.128
Female employment	35,322	1.961	3.31	16,738	1.659	18,584	2.234
Male employment	66,533	2.812	4.538	29,944	2.618	36,589	2.971
Wage Bill (in lakhs)	24,320	0.227	0.745	9,948	0.335	14,372	0.152
Hours in a day	83,405	7.283	2.591	37,172	6.757	46,233	7.707
Month in a year	83,790	10.955	2.43	37,405	10.65	46,385	11.202
Status of firm: Expanding	83,780	0.269	0.443	37,395	0.195	46,385	0.328
Status of firm: Contraction	83,780	0.129	0.335	37,395	0.164	46,385	0.1

Notes: Calculated by authors using ES data. Variables turnover, expenditure, profit, wage bill, total employment, female employment and male employment are winsorized at 1%, deflated at WPI 2004-2005, and are reported in lakhs.

Table 3: Summary statistics: Formal sector outcomes

	Pooled Sample		
Variable	Obs	Mean	Std. Dev.
Turnover (in lakhs)	71,702	3,203.737	8,733.049
Expenditure (in lakhs)	82,167	1,551.189	3,970.407
Profit (in lakhs)	71,662	1,311.423	4,407.235
Total employment	82,291	169.479	571.615
Female employment	26,024	79.410	280.112
Male employment	74,141	77.951	179.004
Wage bill (in lakhs)	82,291	109.694	262.722

Notes: Calculated by authors using ASI data. Variables turnover, expenditure, profit, wage bill, total employment, female employment and male employment are winsorized at 1%, deflated at WPI 2004-2005, and are reported in lakhs.

Table 4: Summary statistics: Explanatory variables

Year	2004		2005		2008		2009		2010		Pooled used for ES		Pooled used for ASI	
Variables	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean
Population proportion exposed to PMGSY	571	5.831	571	7.708	571	14.871	571	15.264	571	15.411	1,940	11.058	757	9.201
KM of roads constructed under PMGSY	473	85.50046	473	136.7473	473	349.5011	473	427.4505	473	489.5414	1,640	303.101	608	205.95
Log(KM of roads constructed under PMGSY)	473	3.656757	473	4.042853	473	4.814859	473	4.995416	473	5.113034	1,640	4.509	608	4.186

Notes: Calculated by authors using administrative data on rural roads construction and SHRUG 2.1. These statistics are calculated using the new all-weather rural roads constructed under PMGSY.

Table 5: Effect of PMGSY on firm's turnover, expenditure and profit for informal sector

	[1]	[2]	[3]
			[Reported in lakhs]
Reference Period		In the last 30 days	
VARIABLES	Turnover	Expenditure	Profit
Panel A: Population Exposed to PMGSY			
$ExposedPop$	0.049***	0.035***	0.012***
	(0.016)	(0.012)	(0.003)
Observations	83,096	82,707	83,057
R-squared	0.380	0.358	0.409
Panel B: KM of roads under PMGSY			
$Log(KM\ of\ roads)$	0.191***	0.139***	0.047***
	(0.059)	(0.045)	(0.015)
Observations	68,159	67,838	68,144
R-squared	0.391	0.371	0.413
Mean [Outcomes]	0.471	0.336	0.125
Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
DistrictCharactersticsPC2001-Wave trends	Yes	Yes	Yes

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Refer to Table 1 for variable definitions. In both panels, the explanatory variables are lagged by one year. All outcome variables are deflated using the 2004-05 WPI, winsorized at 1%, and reported in lakhs. District-level characteristics from the 2001 Census include SC/ST population share, literacy rate, and work participation rate.

Table 6: Effect of PMGSY on employment particulars for informal sector

	[1]	[2]	[3]	[4]
				[Reported in lakhs]
Reference Period		In the last 30 days		
VARIABLES	Total employment	Female employment	Male employment	Wage Bill
Panel A: Population Exposed to PMGSY				
$ExposedPop$	0.018	0.013	0.029*	0.035***
	(0.017)	(0.016)	(0.016)	(0.012)
Observations	83,573	35,260	66,340	24,257
R-squared	0.501	0.483	0.506	0.413
Panel B: KM of roads under PMGSY				
$Log(KM\ of\ roads)$	0.185**	0.100	0.182**	0.109**
	(0.077)	(0.072)	(0.072)	(0.050)
Observations	68,490	28,437	55,105	19,135
R-squared	0.494	0.409	0.499	0.417
Mean [Outcomes]	2.99	1.961	2.812	0.227
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
DistrictCharactersticsPC2001-Wave trends	Yes	Yes	Yes	Yes

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Refer to Table 1 for variable definitions. In both panels, the explanatory variables are lagged by one year. Outcome variables are winsorized at 1% and variable wage bill is deflated using the 2004-05 WPI and reported in lakhs. District-level characteristics from the 2001 Census include SC/ST population share, literacy rate, and work participation rate.

Table 7: Effect of PMGSY on working status for informal sector

	[1]	[2]	[3]	[4]
Reference Period	In the last 30 days		In the last 365 days	
VARIABLES	Hours in a day	Month in a year	Status of firm: Expanding	Status of firm: Contraction
Panel A: Population Exposed to PMGSY				
$ExposedPop$	0.008	0.051**	0.005*	-0.003
	(0.015)	(0.021)	(0.003)	(0.002)
Observations	83,195	83,575	83,565	83,565
R-squared	0.368	0.180	0.165	0.109
Panel B: KM of roads under PMGSY				
$Log(KM\ of\ roads)$	0.144**	0.163**	0.013	-0.012
	(0.070)	(0.066)	(0.011)	(0.008)
Observations	68,179	68,489	68,476	68,476
R-squared	0.356	0.188	0.170	0.112
Mean [Outcomes]	7.283	10.955	0.269	0.129
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
NIC2 FE	Yes	Yes	Yes	Yes
DistrictCharactersticsPC2001-Wave trends	Yes	Yes	Yes	Yes

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Refer to Table 1 for variable definitions. In both panels, the explanatory variables are lagged by one year. District-level characteristics from the 2001 Census include SC/ST population share, literacy rate, and work participation rate.

Table 8: Effect of PMGSY on turnover, expenditure and profit for formal sector

	[1]	[2]	[3]
			[Reported in lakhs]
VARIABLES	Turnover	Expenditure	Profit
Panel A: Population Exposed to PMGSY			
$ExposedPop$	-11.596	-7.488	-1.158
	(17.634)	(8.557)	(8.074)
Observations	69,910	80,091	69,872
R-squared	0.250	0.261	0.218
Panel B: KM of roads under PMGSY			
$Log(KM\ of\ roads)$	-116.685	-92.271***	-20.485
	(74.522)	(35.379)	(39.397)
Observations	49,011	55,687	48,983
R-squared	0.259	0.270	0.225
Mean [Outcomes]	3,203.737	1,551.189	1,311.423
Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
DistrictCharactersticsPC2001-Wave trends	Yes	Yes	Yes

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Refer to Table 1 for variable definitions. In both panels, the explanatory variables are lagged by one year. District-level characteristics from the 2001 Census include SC/ST population share, literacy rate, and work participation rate.

Table 9: Effect of PMGSY on employment particulars for formal sector

	[1]	[2]	[3]	[4]
				[Reported in lakhs]
VARIABLES	Total employment	Female employment	Male employment	Wage bill
Panel A: Population Exposed to PMGSY				
$ExposedPop$	-1.476	-0.049	-0.723	0.011
	(0.903)	(0.511)	(0.585)	(0.785)
Observations	80,183	25,390	72,386	80,183
R-squared	0.119	0.323	0.232	0.322
Panel B: KM of roads under PMGSY				
$Log(KM\ of\ roads)$	-1.115	2.580	0.630	-4.496
	(5.555)	(1.913)	(2.599)	(3.418)
Observations	55,725	15,007	50,214	55,725
R-squared	0.230	0.204	0.243	0.327
Mean [Outcomes]	169.479	79.410	77.951	109.694
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
DistrictCharactersticsPC2001-Wave trends	Yes	Yes	Yes	Yes

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the district level. Refer to Table 1 for variable definitions. In both panels, the explanatory variables are lagged by one year. District-level characteristics from the 2001 Census include SC/ST population share, literacy rate, and work participation rate.

Table 10: Effect of PMGSY on problems faced by the informal sector

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Reference Period	In last 365 days, Problem faced by the firm:								
VARIABLES	Infrastructure	Competition from large units	Marketing	Local harassment	Fuel related	Electricity related	Labor related	Credit related	Faced at least one
Panel A: Population Exposed to PMGSY									
$ExposedPop$	-0.007***	-0.007**	-0.007*	-0.002	-0.001	0.006*	0.001	-0.001	-0.002
	(0.002)	(0.003)	(0.004)	(0.002)	(0.001)	(0.004)	(0.001)	(0.002)	(0.003)
Observations	37,107	37,107	37,143	37,107	37,107	83,553	83,489	83,489	83,552
R-squared	0.115	0.181	0.241	0.168	0.107	0.291	0.141	0.138	0.319
Panel B: KM of roads under PMGSY									
$Log(KM\ of\ roads)$	-0.022**	-0.038***	-0.029	-0.005	-0.000	0.020	0.005	-0.014**	-0.023*
	(0.010)	(0.013)	(0.018)	(0.007)	(0.007)	(0.015)	(0.004)	(0.007)	(0.013)
Observations	28,720	28,720	28,759	28,720	28,720	68,464	68,406	68,406	68,463
R-squared	0.117	0.172	0.247	0.183	0.124	0.285	0.147	0.135	0.321
Mean [Outcomes]	0.045	0.172	0.232	0.03	0.014	0.183	0.029	0.059	0.571
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DistrictCharactersticsPC2001-Wave trends	No	No	No	No	No	Yes	Yes	Yes	Yes

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Refer to Table 1 for variable definitions. In both panels, the explanatory variables are lagged by one year. District-level characteristics from the 2001 Census include SC/ST population share, literacy rate, and work participation rate.

Appendix

Rural Roads and Firm Outcomes in India

Bharti Nandwani, Punarjit Roychowdhury and Binay Shankar

Table A1: Summary statistics: Controls for informal firms

	Pooled sample			(2005-06)		(2010-11)	
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Obs	Mean
Informant Type	.	.	.	.	.	.	.
Owner/Partner	83,575	0.936	0.244	37,195	0.936	46,380	0.937
Manager	83,575	0.018	0.133	37,195	0.02	46,380	0.016
Others/Working Staff	83,575	0.046	0.209	37,195	0.044	46,380	0.047
Response Type	.	.	.	.	.	.	.
Informant, co-operative and capable	83,575	0.813	0.39	37,195	0.791	46,380	0.831
Informant co-operative but not capable	83,575	0.143	0.35	37,195	0.175	46,380	0.118
Informant busy but responded	83,575	0.028	0.166	37,195	0.013	46,380	0.041
Informant reluctant	83,575	0.013	0.114	37,195	0.019	46,380	0.008
Others	83,575	0.002	0.042	37,195	0.002	46,380	0.001
Mixed activity	.	.	.	.	.	.	.
Yes	83,575	0.969	0.174	37,195	0.962	46,380	0.974
No	83,575	0.031	0.174	37,195	0.038	46,380	0.026
Ownership Type	.	.	.	.	.	.	.
Proprietary (male)	83,575	0.74	0.439	37,195	0.737	46,380	0.743
Proprietary (female)	83,575	0.223	0.417	37,195	0.221	46,380	0.225
Partnership with members of the same household	83,575	0.02	0.14	37,195	0.021	46,380	0.019
Partnership between members, not all from the same household	83,575	0.013	0.114	37,195	0.017	46,380	0.01
Co-operative society	83,575	0.001	0.027	37,195	0.002	46,380	0
Limited company (outside public sector)	83,575	0.001	0.032	37,195	0.002	46,380	0
Self-help Group	83,575	0.001	0.031	37,195	0	46,380	0.002
Trusts	83,575	0.000	0.006	37,195	0	46,380	0.000
Others	83,575	0.001	0.024	37,195	0.001	46,380	0.000
Account maintained	.	.	.	.	.	.	.
Yes	83,575	0.081	0.272	37,195	0.093	46,380	0.071
No	83,575	0.919	0.272	37,195	0.907	46,380	0.929
Location	.	.	.	.	.	.	.
Outside household premises	83,575	0.427	0.495	37,195	0.418	46,380	0.434
Within household premises	83,575	0.573	0.495	37,195	0.582	46,380	0.566
Registered under any authority	.	.	.	.	.	.	.
Yes	83,575	0.801	0.399	37,195	0.826	46,380	0.782
No	83,575	0.199	0.399	37,195	0.174	46,380	0.218
Any assistance received	.	.	.	.	.	.	.
Yes	83,575	0.927	0.261	37,195	0.862	46,380	0.979
No	83,575	0.073	0.261	37,195	0.138	46,380	0.021

Notes: Calculates by authors using the NSS enterprise survey. For ownership type option "Co-operative society" and "Limited company (outside public sector)" were only available in NSS62nd rounds, whereas Self-help group and Trust were available only in the NSS 67th round.

Table A2: Summary statistics: Controls for formal sector

	Pooled sample		
Variable	Obs	Mean	Std. Dev.
Organisation type	.	.	.
Individual Proprietorship	80,183	0.213	0.409
Joint Family (HUF)	80,183	0.012	0.110
Partnership	80,183	0.240	0.427
Public Limited Company	80,183	0.200	0.400
Private Limited Company	80,183	0.291	0.454
Government Departmental Enterprise	80,183	0.002	0.043
Public Corporation by Special Act	80,183	0.005	0.067
Co-operative Society	80,183	0.033	0.178
Others (including Trusts, Wakf Board, etc.)	80,183	0.004	0.067
Ownership type	.	.	.
Wholly Central Government	80,183	0.003	0.054
Wholly State and/or Local Govt.	80,183	0.005	0.073
Central/State/Local Government	80,183	0.002	0.039
Joint Sector Public	80,183	0.011	0.106
Joint Sector Private	80,183	0.006	0.075
Wholly Private Ownership	80,183	0.973	0.162
Others	80,183	0.000	0.009
Count of units in case of joint returns	80,183	1.042	0.236

Notes: Calculated by authors using the Annual survey of Industries.

Table A3: Effect of PMGSY on firm's turnover, expenditure and profit for informal sector with public goods

	[1]	[2]	[3]
			[Reported in lakhs]
Reference Period		In the last 30 days	
VARIABLES	Turnover	Expenditure	Profit
Panel A: Population Exposed to PMGSY			
$ExposedPop$	0.047***	0.035***	0.012***
	(0.016)	(0.013)	(0.004)
Observations	82,453	82,065	82,414
R-squared	0.381	0.359	0.410
Panel B: KM of roads under PMGSY			
$Log(KM\ of\ roads)$	0.175***	0.124***	0.045***
	(0.058)	(0.044)	(0.015)
Observations	67,548	67,228	67,533
R-squared	0.393	0.373	0.413
Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes
NIC2 FE	Yes	Yes	Yes
DistrictPublicGoods-Wave trends	Yes	Yes	Yes

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Refer to Table 1 for variable definitions. These regressions control for district-level public goods. District-level public goods include average number of rural primary schools, high schools, adult literacy centres, primary health centres in a district, and percentage of villages that have commercial banks, post offices, telephone booth, and power supply in a district.

Table A4: Effect of PMGSY on employment particulars for informal sector with public goods

	[1]	[2]	[3]	[4]
				[Reported in lakhs]
Reference Period		In the last 30 days		
VARIABLES	Total employment	Female employment	Male employment	Wage Bill
Panel A: Population Exposed to PMGSY				
$ExposedPop$	0.003	0.011	0.016	0.031**
	(0.018)	(0.016)	(0.017)	(0.012)
Observations	82,930	35,014	65,798	24,084
R-squared	0.501	0.483	0.506	0.416
Panel B: KM of roads under PMGSY				
$Log(KM\ of\ roads)$	0.179**	0.086	0.174**	0.101**
	(0.075)	(0.073)	(0.071)	(0.050)
Observations	67,879	28,202	54,593	18,973
R-squared	0.494	0.409	0.500	0.422
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
NIC2 FE	Yes	Yes	Yes	Yes
DistrictPublicGoods-Wave trends	Yes	Yes	Yes	Yes

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Refer to Table 1 for variable definitions. These regressions control for district-level public goods. District-level public goods include average number of rural primary schools, high schools, adult literacy centres, primary health centres in a district, and percentage of villages that have commercial banks, post offices, telephone booth, and power supply in a district.

Table A5: Effect of PMGSY on working status for informal sector with public goods

	[1]	[2]	[3]	[4]
Reference Period	In the last 30 days		In the last 365 days	
VARIABLES	Hours in a day	Month in a year	Status of firm: Expanding	Status of firm: Contraction
Panel A: Population Exposed to PMGSY				
$ExposedPop$	-0.006	0.043*	0.001	-0.005**
	(0.017)	(0.023)	(0.003)	(0.002)
Observations	82,556	82,932	82,922	82,922
R-squared	0.368	0.179	0.165	0.109
Panel B: KM of roads under PMGSY				
$Log(KM\ of\ roads)$	0.147**	0.161**	0.010	-0.013
	(0.069)	(0.066)	(0.011)	(0.008)
Observations	67,568	67,878	67,865	67,865
R-squared	0.356	0.188	0.170	0.112
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
NIC2 FE	Yes	Yes	Yes	Yes
DistrictPublicGoods-Wave trends	Yes	Yes	Yes	Yes

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Refer to Table 1 for variable definitions. These regressions control for district-level public goods. District-level public goods include average number of rural primary schools, high schools, adult literacy centres, primary health centres in a district, and percentage of villages that have commercial banks, post offices, telephone booth, and power supply in a district.

Table A6: Impact on firm's turnover, expenditure and profit for informal sector: Robustness to omitted variable bias

	[1]	[2]	[3] [Reported in lakhs]
Reference Period		In the last 30 days	
VARIABLES	Turnover	Expenditure	Profit
Coefficient of interest		ExposedPop	
Uncontrolled			
β	-0.009	-0.007	-0.002
R^2	0.001	0.001	0.002
Controlled			
β	0.013	0.011	0.002
R^2	0.354	0.332	0.385
R^2_{max}	0.494	0.465	0.532
β for $\delta = 1$	0.022	0.018	0.004
δ for $\beta = 0$	-1.511	-1.574	-1.381
Coefficient of interest		Log(KM of roads)	
Uncontrolled			
β	-0.004	-0.003	-0.001
R^2	0.000	0.000	0.000
Controlled			
β	0.062	0.046	0.016
R^2	0.368	0.347	0.391
R^2_{max}	0.058	0.482	0.537
β for $\delta = 1$	0.006	0.065	0.022
δ for $\beta = 0$	1.110	-2.387	-2.575

Notes: We report both the coefficient β and R^2 from the controlled and uncontrolled regressions. The controlled regression includes fixed effects for district, NIC-2 industries, and survey waves, as well as interactions between wave and district-level public goods availability from the 2001 population census. It also controls for firm-level characteristics. The uncontrolled regression includes only the key independent variable of interest. Next, we report the bias-adjusted β under the assumption that $\delta = 1$ and an upper bound for R^2_{max}. We use the function $R^2_{max} = \min(\pi . R^2_{controlled}, 1)$, and set $\pi = 1.3$ as suggested by Oster 2019.

Table A7: Impact on employment particulars for informal firms: Robustness to omitted variable bias

	[1]	[2]	[3]	[4] [Reported in lakhs]
Reference Period	In the last 30 days			
VARIABLES	Total employment	Female employment	Male employment	Wage Bill
Coefficient of interest	ExposedPop			
Uncontrolled				
β	-0.024	-0.014	-0.026	-0.004
R^2	0.003	0.002	0.004	0.004
Controlled				
β	-0.004	-0.004	0.000	0.004
R^2	0.482	0.448	0.485	0.375
R^2_{max}	0.651	0.628	0.658	0.537
β for $\delta = 1$	0.003	0.000	0.010	0.008
δ for $\beta = 0$	0.591	0.915	-0.050	-1.154
Coefficient of interest	Log(KM of roads)			
Uncontrolled				
β	0.075	0.033	0.095	-0.003
R^2	0.001	0.000	0.001	0.000
Controlled				
β	0.109	0.075	0.113	0.041
R^2	0.447	0.396	0.433	0.350
R^2_{max}	0.642	0.532	0.649	0.542
β for $\delta = 1$	0.128	0.205	0.103	0.071
δ for $\beta = 0$	-8.566	-3.122	-45.971	-2.383

Notes: We report both the coefficient β and R^2 from the controlled and uncontrolled regressions. The controlled regression includes fixed effects for district, NIC-2 industries, and survey waves, as well as interactions between wave and district-level public goods availability from the 2001 population census. It also controls for firm-level characteristics. The uncontrolled regression includes only the key independent variable of interest. Next, we report the bias-adjusted β under the assumption that $\delta = 1$ and an upper bound for R^2_{max}. We use the function $R^2_{max} = \min(\pi . R^2_{controlled}, 1)$, and set $\pi = 1.3$ as suggested by Oster 2019.

Table A8: Impact on working status for informal firms: Robustness to omitted variable bias

	[1]	[2]	[3]	[4]
Reference Period	In the last 30 days		In the last 365 days	
VARIABLES	Hours in a day	Month in a year	Status of firm: Expanding	Status of firm: Contraction
Coefficient of interest	ExposedPop			
Uncontrolled				
β	0.009	0.008	0.003	-0.001
R^2	0.002	0.001	0.007	0.001
Controlled				
β	0.003	0.007	0.001	0.000
R^2	0.340	0.152	0.126	0.076
R^2_{max}	0.478	0.234	0.215	0.142
β for $\delta = 1$	0.000	0.006	0.000	0.000
δ for $\beta = 0$	1.078	6.954	0.797	0.455
Coefficient of interest	Log(KM of roads)			
Uncontrolled				
β	0.151	0.054	0.020	0.002
R^2	0.011	0.002	0.006	0.000
Controlled				
β	0.068	0.090	0.004	-0.003
R^2	0.331	0.161	0.132	0.079
R^2_{max}	0.463	0.244	0.221	0.146
β for $\delta = 1$	0.033	0.109	-0.008	-0.007
δ for $\beta = 0$	1.970	-4.822	0.333	-0.711

Notes: We report both the coefficient β and R^2 from the controlled and uncontrolled regressions. The controlled regression includes fixed effects for district, NIC-2 industries, and survey waves, as well as interactions between wave and district-level public goods availability from the 2001 population census. It also controls for firm-level characteristics. The uncontrolled regression includes only the key independent variable of interest. Next, we report the bias-adjusted β under the assumption that $\delta = 1$ and an upper bound for R^2_{max}. We use the function $R^2_{max} = \max(\pi . R^2_{controlled}, 1)$, and set $\pi = 1.3$ as suggested by Oster 2019.

Table A9: Effect of PMGSY on firm's turnover, expenditure and profit for informal sectors: Robustness to IV-2SLS method

	[1]	[2]	[3]
			[Reported in lakhs]
Reference Period		In the last 30 days	
VARIABLES	Turnover	Expenditure	Profit
ExposedPop	0.134***	0.097***	0.037***
	(0.036)	(0.028)	(0.008)
Observations	83,096	82,707	83,057
Controls	Yes	Yes	Yes
Dist FE	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
DistrictCharactersticsPC2001-Wave trends	Yes	Yes	Yes
First-stage coefficient	0.010***	0.010***	0.010***
First-stage robust standard errors	[0.001]	[0.001]	[0.001]
First stage F statistic	131.6	131.12	131.54
	[p=0.000]	[p=0.000]	[p=0.000]
Kleibergen Paap rK LM statistic	99.42	99.08	99.33
	[p=0.000]	[p=0.000]	[p=0.000]

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. This table uses percentage of district population that is eligible for PMGSY interacted with year of survey as instrument for exposed population percentage. Refer to Table 1 for variable definitions. In both panels, the explanatory variables are lagged by one year. District-level characteristics from the 2001 Census include SC/ST population share, literacy rate, and work participation rate.

Table A10: Effect of PMGSY on employment particulars for informal sector: Robustness to IV-2SLS method

	[1]	[2]	[3]	[4]
				[Reported in lakhs]
Reference Period		In the last 30 days		
VARIABLES	Total employment	Female employment	Male employment	Wage Bill
$ExposedPop$	0.132**	0.073	0.150**	0.094***
	(0.061)	(0.049)	(0.061)	(0.025)
Observations	83,573	35,260	66,340	24,257
Controls	Yes	Yes	Yes	Yes
Dist FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
DistrictCharactersticsPC2001-Wave trends	Yes	Yes	Yes	Yes
First-stage coefficient	0.010***	0.010***	0.010***	0.010***
First-stage robust standard errors	[0.001]	[0.001]	[0.001]	[0.001]
First stage F statistic	130.93	86.5	152.61	78.44
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
Kleibergen Paap rK LM statistic	99.18	75.89	110.8	58.34
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. This table uses percentage of district population that is eligible for PMGSY interacted with year of survey as instrument for exposed population percentage. Refer to Table 1 for variable definitions. In both panels, the explanatory variables are lagged by one year. District-level characteristics from the 2001 Census include SC/ST population share, literacy rate, and work participation rate.

Table A11: Effect of PMGSY on working status for informal sector: Robustness to IV-2SLS method

	[1]	[2]	[3]	[4]
Reference Period	In the last 30 days		In the last 365 days	
VARIABLES	Hours in a day	Month in a year	Status of firm: Expansion	Status of firm: Contraction
ExposedPop	0.055*	0.083**	0.011*	-0.016***
	(0.032)	(0.036)	(0.006)	(0.005)
Observations	83,195	83,575	83,565	83,565
Controls	Yes	Yes	Yes	Yes
Dist FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
DistrictCharactersticsPC2001-Wave trends	Yes	Yes	Yes	Yes
First-stage coefficient	0.010***	0.010***	0.010***	0.010***
First-stage robust standard errors	[0.001]	[0.001]	[0.001]	[0.001]
First stage F statistic	132.4	130.9	131.35	131.38
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
Kleibergen Paap rK LM statistic	100.24	99.19	99.44	99.44
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. This table uses percentage of district population that is eligible for PMGSY interacted with year of survey as instrument for exposed population percentage. Refer to Table 1 for variable definitions. In both panels, the explanatory variables are lagged by one year. District-level characteristics from the 2001 Census include SC/ST population share, literacy rate, and work participation rate.

Table A12: Heterogeneous effect of PMGSY on firm's turnover, expenditure and profit.

	[1]	[2]	[3]
			[Reported in lakhs]
Reference Period		In the last 30 days	
VARIABLES	Turnover	Expenditure	Profit
Panel A: Firms employing up to 2 workers including owner			
ExposedPop	0.000	0.001	-0.000
	(0.001)	(0.001)	(0.000)
Observations	61,531	61,129	61,495
R-squared	0.166	0.136	0.315
Panel B: Firms Employing more than 2 workers			
ExposedPop	0.270***	0.191***	0.065***
	(0.085)	(0.066)	(0.018)
Observations	21,547	21,564	21,544
R-squared	0.464	0.448	0.472
Panel C: Firm size Below 90%			
ExposedPop	0.002	0.002	0.000
	(0.002)	(0.002)	(0.000)
Observations	72,984	72,567	72,941
R-squared	0.124	0.105	0.204
Panel D: Firm size Above 90%			
ExposedPop	0.709***	0.519***	0.166***
	(0.214)	(0.168)	(0.049)
Observations	10,083	10,112	10,087
R-squared	0.526	0.515	0.512
Panel E: Firms operating within household premises			
ExposedPop	0.002	0.002	0.000
	(0.002)	(0.002)	(0.000)
Observations	47,618	47,182	47,608
R-squared	0.134	0.122	0.179
Panel F: Firms operating outside household premises			
ExposedPop	0.135***	0.095**	0.034***
	(0.048)	(0.037)	(0.010)
Observations	35,477	35,524	35,448
R-squared	0.432	0.410	0.453
Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes
NIC2 FE	Yes	Yes	Yes
DistrictCharactersticsPC2001-Wave trends	Yes	Yes	Yes

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Refer to Table 1 for variable definitions. In all the panels, the explanatory variables are lagged by one year. All outcome variables are deflated using the 2004-05 WPI, winsorized at 1%, and reported in lakhs. District-level characteristics from the 2001 Census include SC/ST population share, literacy rate, and work participation rate.

Table A13: Heterogeneous effect of PMGSY on employment particulars

	[1]	[2]	[3]	[4]
				[Reported in lakhs]
Reference Period		In the last 30 days		
VARIABLES	Total employment	Female employment	Male employment	Wage Bill
Panel A: Firms employing up to 2 workers including owner				
$ExposedPop$	0.004	0.008**	-0.002	-0.000
	(0.003)	(0.003)	(0.003)	(0.000)
Observations	61,943	25,658	45,388	6,738
R-squared	0.216	0.158	0.166	0.424
Panel B: Firms Employing more than 2 workers				
$ExposedPop$	0.050	0.013	0.072	0.051***
	(0.057)	(0.069)	(0.044)	(0.018)
Observations	21,627	9,575	20,948	17,485
R-squared	0.559	0.578	0.569	0.424
Panel C: Firm size Below 90%				
$ExposedPop$	0.012**	0.011**	0.007	-0.000
	(0.005)	(0.005)	(0.005)	(0.000)
Observations	73,436	30,062	56,370	14,517
R-squared	0.265	0.141	0.230	0.177
Panel D: Firm size Above 90%				
$ExposedPop$	0.037	0.083	0.050	0.090***
	(0.152)	(0.183)	(0.118)	(0.032)
Observations	10,108	5,116	9,937	9,697
R-squared	0.587	0.615	0.610	0.450
Panel E: Firms operating within household premises				
$ExposedPop$	0.021*	0.012	0.016	-0.004*
	(0.012)	(0.008)	(0.012)	(0.002)
Observations	47,910	26,995	32,108	5,531
R-squared	0.264	0.217	0.235	0.428
Panel F: Firms operating outside household premises				
$ExposedPop$	0.015	0.038	0.039	0.048***
	(0.034)	(0.062)	(0.030)	(0.016)
Observations	35,662	8,218	34,231	18,662
R-squared	0.581	0.631	0.568	0.422
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
NIC2 FE	Yes	Yes	Yes	Yes
DistrictCharactersticsPC2001-Wave trends	Yes	Yes	Yes	Yes

Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Refer to Table 1 for variable definitions. In all panels, the explanatory variables are lagged by one year. Outcome variables are winsorized at 1% and variable wage bill is deflated using the 2004-05 WPI and reported in lakhs. District-level characteristics from the 2001 Census include SC/ST population share, literacy rate, and work participation rate.

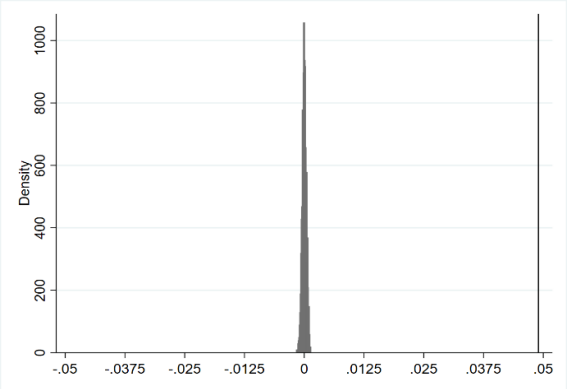
Table A14: Heterogeneous effect of PMGSY on working status

	[1]	[2]	[3]	[4]
Reference Period	In the last 30 days		In the last 365 days	
VARIABLES	Hours in a day	Month in a year	Status of firm: Expansion	Status of firm: Contraction
Panel A: Firms employing up to 2 workers including owner				
$ExposedPop$	0.018	0.031	0.007**	-0.003
	(0.015)	(0.022)	(0.003)	(0.002)
Observations	61,625	61,929	61,921	61,921
R-squared	0.345	0.154	0.174	0.124
Panel B: Firms Employing more than 2 workers				
$ExposedPop$	-0.031	0.084***	-0.000	-0.008
	(0.028)	(0.032)	(0.006)	(0.005)
Observations	21,559	21,623	21,621	21,621
R-squared	0.338	0.344	0.203	0.159
Panel C: Firm size Below 90%				
$ExposedPop$	0.018	0.040*	0.006*	-0.003
	(0.015)	(0.021)	(0.003)	(0.002)
Observations	73,094	73,419	73,410	73,410
R-squared	0.355	0.155	0.170	0.116
Panel D: Firm size Above 90%				
$ExposedPop$	-0.028	0.049	0.006	-0.011
	(0.050)	(0.058)	(0.009)	(0.008)
Observations	10,072	10,127	10,126	10,126
R-squared	0.421	0.470	0.238	0.212
Panel E: Firms operating within household premises				
$ExposedPop$	0.034**	0.033	0.006	-0.002
	(0.017)	(0.029)	(0.004)	(0.003)
Observations	47,658	47,915	47,902	47,902
R-squared	0.349	0.191	0.185	0.141
Panel F: Firms operating outside household premises				
$ExposedPop$	-0.020	0.064**	0.005	-0.008**
	(0.025)	(0.025)	(0.005)	(0.003)
Observations	35,536	35,659	35,662	35,662
R-squared	0.306	0.258	0.185	0.120
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
NIC2 FE	Yes	Yes	Yes	Yes
DistrictCharactersticsPC2001-Wave trends	Yes	Yes	Yes	Yes

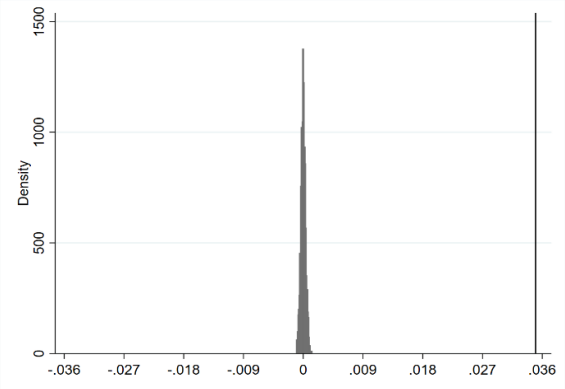
Notes: Robust standard errors clustered at the district level in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Refer to Table 1 for variable definitions. In all panels, the explanatory variables are lagged by one year. District-level characteristics from the 2001 Census include SC/ST population share, literacy rate, and work participation rate.

Fig A1: Falsification graphs for informal sector (1000 randomization of the main explanatory variable – $ExposedPop$)

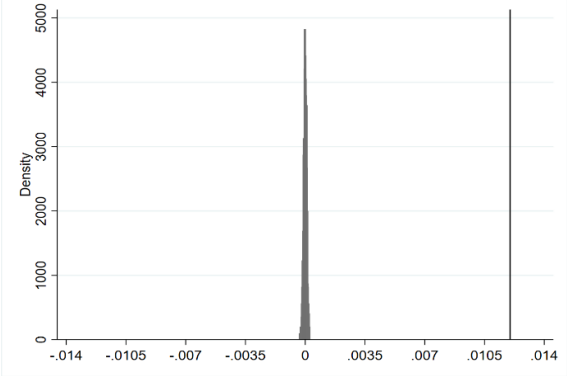
A: Falsification graph for turnover in informal sector



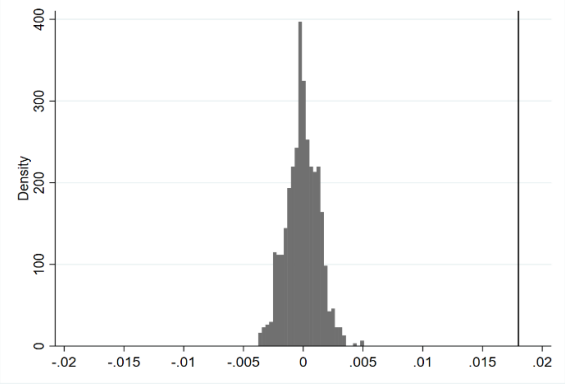
B: Falsification graph for expenditure in informal sector



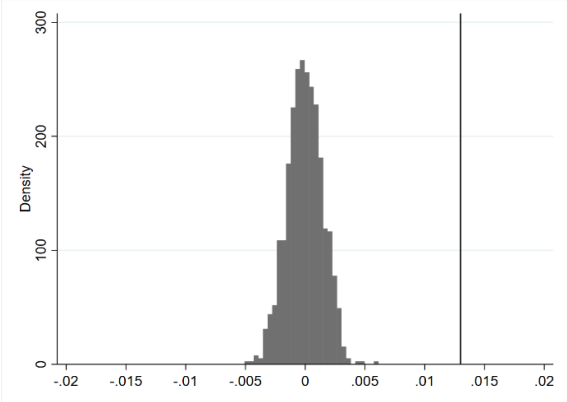
C: Falsification graph for profit in informal sector



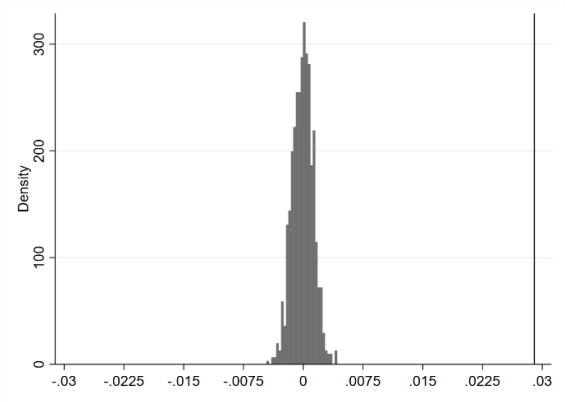
D: Falsification graph for total employment in informal sector



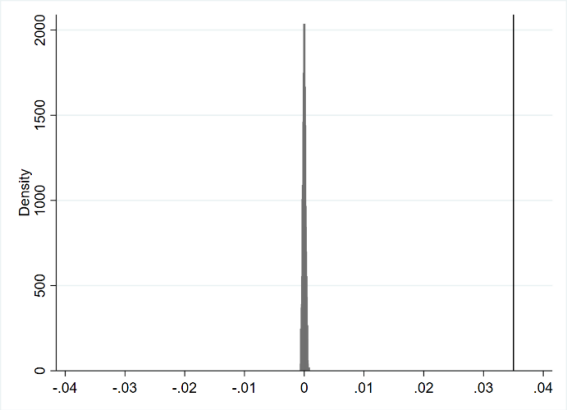
E: Falsification graph for female employment in informal sector



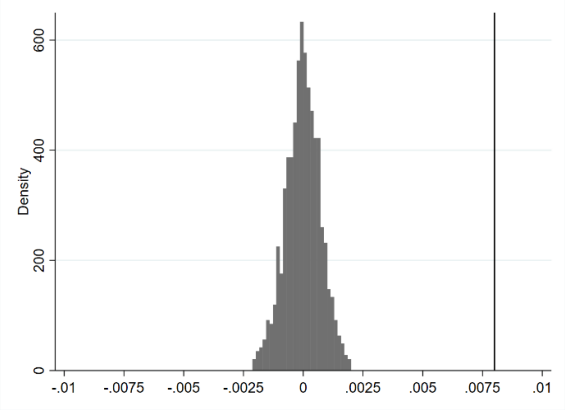
F: Falsification graph for male employment in informal sector



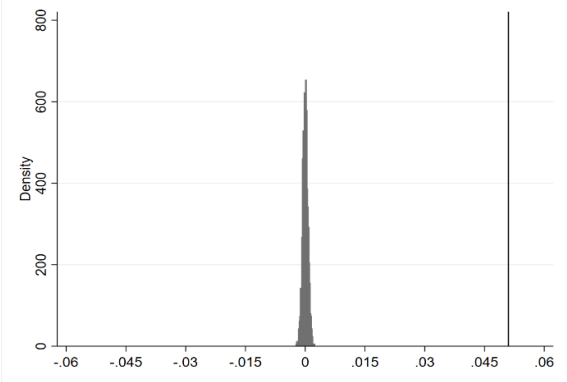
G: Falsification graph for wage bill in informal sector



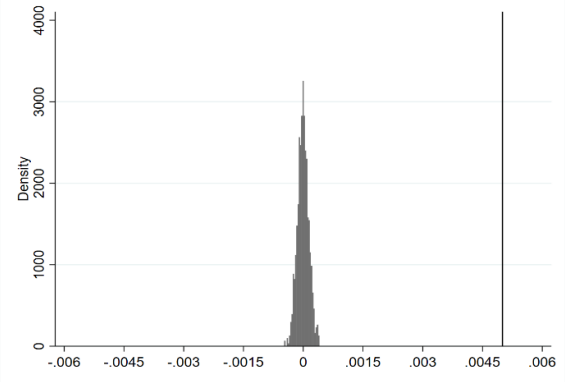
H: Falsification graph for hours in a day in informal sector



I: Falsification graph for month in a year in informal sector



J: Falsification graph for expanding status in informal sector



K: Falsification graph for contraction status in informal sector