Journal of Public Economics 214 (2022) 104725

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Infrastructure maintenance and rural economic activity: Evidence from India ☆

Ritam Chaurey [a,*], Duong Trung Le [b]

[a] Johns Hopkins University, School of Advanced International Studies and IZA, United States
[b] The World Bank, United States

ARTICLE INFO

Article history:
Received 9 May 2021
Revised 27 May 2022
Accepted 2 August 2022
Available online 18 August 2022

JEL Codes:
O12
O18
O25
R11

Keywords:
Infrastructure
Firms
India
Microenterprises
Regression Discontinuity Design

ABSTRACT

We evaluate the impact of a rural infrastructure development and maintenance scheme, directed towards India's most "backward" districts. Importantly, these infrastructure grants were spent on maintenance, improvements, and complementary investments to existing infrastructure. Using a Regression Discontinuity Design and multiple data sets covering the entire firm size distribution, as well as household employment surveys, we show evidence on the effectiveness of the program on local economic activity. In treated districts, the policy appears to have increased overall village-level employment, and the number of microenterprises, but there were no changes for formal firms. These increases were likely driven by a decline in the share of agricultural wage workers. At the individual-level, wages and the number of days worked also increased. There is suggestive evidence that both rural electrification and connectivity were important mechanisms driving our results. We find stronger impacts in electricity and road-intensive industries, and in villages that had paved roads and electricity prior to the program. Overall, our paper suggests that improving rural infrastructure conditions can boost economic activity, especially by stimulating microenterprises.

1. Introduction

There is universal consensus that investments in physical infrastructure (roads, electricity, telecommunications, Internet, dams, irrigation, etc.) are important determinants of economic growth. However, low quality, low operating efficiency, inadequate maintenance, and lack of attention to the needs of users play a role in reducing the development impacts of infrastructure (World Bank, 1994). Such factors may especially be salient in rural and sparsely populated areas, as compared to urban metropolitan areas. A recent World Bank report estimates that infrastructure disruptions due to poor maintenance, and mismanagement have cost households and firms at least USD 390 billion a year in low- and middle-income countries (World Bank, 2021). Although most policymakers agree that investing in infrastructure maintenance would enable growth and well-being of both people and firms (Italian G20 Presidency, 2021), causal evidence on their effectiveness has been sparse.

In this paper, we study the effects of a rural infrastructure grants program, Rashtriya Sam Vikas Yojana (RSVY), launched by the Government of India, that was explicitly meant to make complementary investments to existing infrastructure including improvement and maintenance. In general, analyzing the causal effects of rural infrastructure investments is challenging because they are inherently place-based, often directed either to economically lagging regions to incentivize growth or to potentially fast-growing regions to further accelerate growth. Since RSVY grants were extended to districts using a score-based assignment mechanism, we are able to address this primary concern of non-random placement of infrastructure investments by using a regression

☆ We would like to thank Ama Baafra Abeberese, Sam Asher, Siddharth Hari, Gaurav Khanna, Amit Khandelwal, Solomon Polachek, Martin Rotemberg, David Slichter, Eric Verhoogen, Susan Wolcott, and participants at the Young Economists Symposium (Yale 2017), Northeastern Association of Business, Economics & Technology (Penn State 2017), ACEGD ISI (New Delhi 2018), Eastern Economic Association (Boston 2018), George Mason Schar School of Policy and Government, EPED (Montreal 2018), Ce2 workshop (Warsaw 2019), UNU-WIDER Development Conference (2019), Florida International University, and Georgetown University for helpful comments. This paper subsumes earlier versions, "Infrastructure Grants and the Performance of Microenterprises," as well as "Rural Infrastructure Development and Economic Activity." All remaining errors are our own.

* Corresponding author.
E-mail addresses: rchaurey@jhu.edu (R. Chaurey), dle6@worldbank.org (D.T. Le).

https://doi.org/10.1016/j.jpubeco.2022.104725

discontinuity design. We combine multiple data sets covering the entire firm-size distribution, various household surveys, and data on night-time lights to analyze the effects of RSVY on local economic outcomes and some plausible mechanisms driving these effects.

RSVY was launched in fiscal year 2003–04 with the main goal of facilitating physical infrastructure development in the most economically "backward" districts in India. This program was one of the first direct attempts carried out by the central government to identify and support India's economically lagging districts to reduce regional economic imbalance and speed up development. Under RSVY, each eligible district was entitled to receive 450 million Indian rupees ("Rs.") ($\approx$ 10 million USD (2010 exchange rates)) over three years to address "critical gaps" in physical and social infrastructure. This amounted to around 1.15% of the average eligible district's GDP between 2003–04 and 2005–06. The policy guidelines mandated that the infrastructure gaps should be identified in a decentralized manner at the district level and involve both the community and key stakeholders. This essentially meant that each district, in consultation with various stakeholders, could decide where (and on what) they spent the RSVY funds. Funds were to be used to *improve and maintain* or *make complementary investments to* existing infrastructure rather than on completely new projects. Therefore, RSVY could be considered as a grant explicitly meant for infrastructure maintenance and improvement rather than new construction. For example, for physical infrastructure (i.e., rural connectivity and electrification), RSVY funds could be spent on widening and strengthening of roads, building small bridges, building vital road links to connect to the marketplace, and strengthening rural electricity transmission and distribution infrastructure. Districts had the flexibility to spend the RSVY funds on multiple small infrastructure-related projects. This flexibility in the use of funds for maintenance and improvement differentiates RSVY from other programs that are often focused on building a specific infrastructural facility such as new roads and highways, dams, or electricity grids.

The specific guidelines used by the Government of India to prioritize treated districts makes RSVY an ideal natural experiment. The central government first allocated to each of the 17 major states in the country a pre-specified quota of districts based on the states' poverty headcount ratios. Next, each state government identified the districts deemed fit to receive the grant. However, the central government's guidelines for RSVY specifically requested that the most "backward" districts—based on an official district-level "Backwardness Index"— be prioritized as beneficiaries of RSVY grants.

Our empirical strategy relies on the identification of RSVY-eligible "backward" districts using official documentation from the Planning Commission. We reconstruct the "Backwardness Index" scores for each district and use the distance to the score of the cutoff district in each state as the running variable in a Regression Discontinuity Design (RD) framework. We run RD regressions on various economic outcomes at the village, firm, and individual levels, using multiple data sets.

We find a number of results on the effectiveness of RSVY in the short run. First, the infrastructure grants led to an aggregate increase in overall village employment in treated districts, although the results are not statistically significant for all specifications. This was driven entirely by village employment in microenterprises, with no effect on employment in formal firms. We disentangle the increase in overall village-level employment into extensive and intensive margins. On the extensive margin, we find an increase in the number of firms in villages in RSVY districts, again entirely driven by an increase in the number of microenterprises, with no change in the number of formal firms. However, we find no change on the intensive margin, i.e., the average firm size in villages. We find that RSVY led to greater firm creation and increased total employment for microenterprises, although average firm size remained unchanged. For formal firms, total employment, total number of firms, and average firm size did not change. This is not surprising given that RSVY was a rural infrastructure program, whereas most formal firms in India locate in or near urban areas.[1]

Next, we look at the effects of RSVY on individual- and household-level outcomes. Consistent with the increase in total village-level employment, we find that individual wages, the number of days worked, as well as monthly household consumption expenditure increased. We then check for changes along various relevant margins in the rural labor market that could explain the increase in total employment. We find suggestive evidence for sectoral reallocation (i.e., a reduction in the share of workers in agriculture) and a corresponding increase in the share of non-agricultural workers. However, we do not find any changes in labor force participation, in-migration, or unemployment.

Since RSVY grants could be used across multiple infrastructure projects, it is important to understand the possible channels through which the policy affected the performance of small firms. However, information on the exact projects and locations within the districts is not available. We therefore provide several suggestive pieces of evidence on the channels through which RSVY may have affected local economic activity. First, we find that there was a significant reduction in the likelihood of firms reporting problems with power shortages or access to raw materials, but there was no change in problems *unrelated* to infrastructure. These results suggest quality improvements in rural electrification and connectivity in RSVY districts. To provide additional supporting evidence for these two channels, we focus on firms in electricity- and road-intensive industries. The reductions in the likelihood of a power outage and problems with access to raw materials are concentrated in the most electricity- and road-intensive industries, respectively. We also find that firms in the most electricity- and road-intensive industries had the largest increases in employment, and the proportion of new firms (those established less than 3 years ago) also increased. Furthermore, the purpose of RSVY funds was to fill critical infrastructure gaps and for infrastructure maintenance rather than to provide roads to unconnected areas or develop electricity grids in unelectrified areas. Therefore, the effects of this infrastructure-maintenance policy should have been larger in villages where paved roads and electricity already existed before RSVY, compared to villages without paved roads and electricity. We find evidence consistent with this claim. We show that after the implementation of RSVY, total employment in microenterprises as well as the number of microenterprises increased in villages that had both paved roads and electricity prior to the policy, and we find no effects in villages without paved roads or electricity. These results also lend credibility to the claim that RSVY funds were used on infrastructure maintenance and improvement rather than on building new infrastructure.

Our results withstand multiple placebo and robustness checks. First, using pre-treatment data, we find no effect on any of the main outcome variables before the introduction of the policy. Second, we find no effect of the policy when the eligibility threshold is hypothetically moved to a different point along the distribution of the backwardness-distance scores (running variable). Third, we estimate difference-in-discontinuities regressions and find results similar to our main RD specifications. Finally, our main results are robust to different bandwidth choices.

Although our paper focuses on infrastructure maintenance, we directly contribute to the literature that seeks to establish causal

[1] Only one in five villages in our sample had any formal firm.

links between different types of infrastructure provision and economic outcomes in developing countries, for example, for rural roads (Aggarwal, 2018; Adukia et al., 2020; Asher and Novosad, 2020; Asher et al., 2020; Mannava et al., 2020; Garz and Perova, 2021), highways (Asturias and García-Santana, 2018; Faber, 2014; Ghani et al., 2016), railroads (Donaldson, 2018), bridges (Brooks and Donovan, 2019), electricity (Abeberese, 2017, Aklin et al., 2017, Allcott et al., 2016, Dinkelman, 2011, Rud, 2012a, Lipscomb and Mobarak, 2013, Chakravorty et al., 2014, Burlig and Preonas, 2016, Lee and Miguel, 2020, Hardy and McCasland, 2017, Lenz et al., 2017, Usmani, 2019), dams (Duflo and Pande, 2007), telecommunications and Internet (Jensen, 2007; Aker and Mbiti, 2010; HHjort and Poulsen, 2019), and water and sanitation (Alsan and Goldin, 2019; Devoto et al., 2012). A majority of these papers find positive effects of infrastructure provision on various measures of economic development. However, recent work on rural electrification in Kenya (Lee and Miguel, 2020) and India (Burlig and Preonas, 2016), and rural roads in India (Asher and Novosad, 2020), have questioned the effectiveness of large-scale rural infrastructure programs in benefiting the rural poor.

Our analysis differs from the above papers in several dimensions. First, in contrast to the studies that examine the effects of new electricity connections/grids or new rural roads in places that were previously unelectrified or unconnected, RSVY funds were spent on filling critical gaps in infrastructure, including maintenance and improvement. This is an important distinction because it is more likely that firms would operate in areas with pre-existing infrastructure than in remote areas with no previous electricity. To this extent, our findings that RSVY led to higher employment and number of firms in villages that already had paved roads and electricity prior to the policy, and no effect in unelectrified and unconnected villages, are consistent with these papers. We interpret our findings as suggesting that investments in improving and maintaining existing infrastructure are critical for rural economic development.

Finally, our paper also contributes to the literature on the determinants of microenterprise growth (see Jayachandran, 2020 for a detailed discussion), and since infrastructure programs are directed towards particular locations, to the literature on place-based policies.[2]

Overall, the results of our paper underscore the importance of maintenance and improvements of existing infrastructure. Moreover, investments in maintenance may have important distributional consequences. In our setting, these investments led to large impacts on microenterprises, which employ close to three-quarters of the non-agricultural workforce and are vital to rural economic activity.[3] Especially for resource constrained policymakers in developing countries, such investments may be more cost effective than building new infrastructure in increasing local economic activity.

The rest of the paper proceeds as follows: Section 2 provides a detailed description of RSVY, its objectives, and the assignment algorithm. Section 3 explains the data used for the analysis. We describe our empirical strategy in Section 4. Section 5 presents and discusses the empirical results. Finally, Section 6 concludes.

2. Rashtriya Sam Vikas Yojana (RSVY)

In 2003–04, the Government of India launched the Rashtriya Sam Vikas Yojana (RSVY) to "remove barriers to economic growth, accelerate the development process, and improve the quality of life of the people" (Planning Commission, 2003). The program was one of the first direct attempts carried out by the central government to identify and support India's "backward" districts. RSVY covered a total of 147 backward districts, out of approximately 600 total districts in the country. Under the policy guidelines, each district was entitled to receive rural infrastructure grant amounts of 450,000,000 Rupees (approximately 10 million USD) over fiscal years 2004–05, 2005–06, and 2006–07. The proposed transfer mechanism was equal payments of 150,000,000 Rupees per year. Fig. 1 Panel A shows the details of the 115 districts that were selected using a transparent assignment mechanism (discussed in the next sub-section).

As per the central government's instructions, all RSVY funds were to be used to address "critical gaps" in physical and social infrastructure to alleviate the problems of infrastructure deficits, low agricultural productivity, and excessive unemployment (Planning Commission, 2003). To identify these critical gaps, the policy guidelines mandated a decentralized, district-level, bottom-up planning approach that involved the community and key stakeholders, including (but not limited to) Panchayati Raj Institutions (village-level institutions), community-based organizations, and line departments. This was done in order to ensure that the plan was representative of the needs of the district. District Perspective Plans (DPPs) were then prepared by the District Administration, identifying the project proposals on which the RSVY funds would be spent. According to Planning Commission guidelines, in the physical infrastructure sector (rural connectivity and electrification), RSVY funds could be spent on road upgrades, bridges and culverts, especially vital road links to connect to the marketplace, and strengthening the rural electricity transmission and distribution infrastructure. Furthermore, the guidelines specified that investment in agriculture or irrigation related programs should be accompanied by important forward and backward linkages such as rural connectivity and electrification wherever possible.[4] District Perspective Plans with details on the characteristics of programs undertaken at the district level are not publicly available. However, according to an evaluation study that surveyed 15 districts from 11 states, approximately 77% of the transferred funds were invested in infrastructural interventions, including rural connectivity, electrification, and agricultural and irrigation improvement projects (Program Evaluation Organization, 2010).

In February 2007, the Government of India launched the Backward Regions Grant Fund (BRGF) that subsumed RSVY and extended it to 250 backward districts across 27 states.

2.1. Assignment mechanism

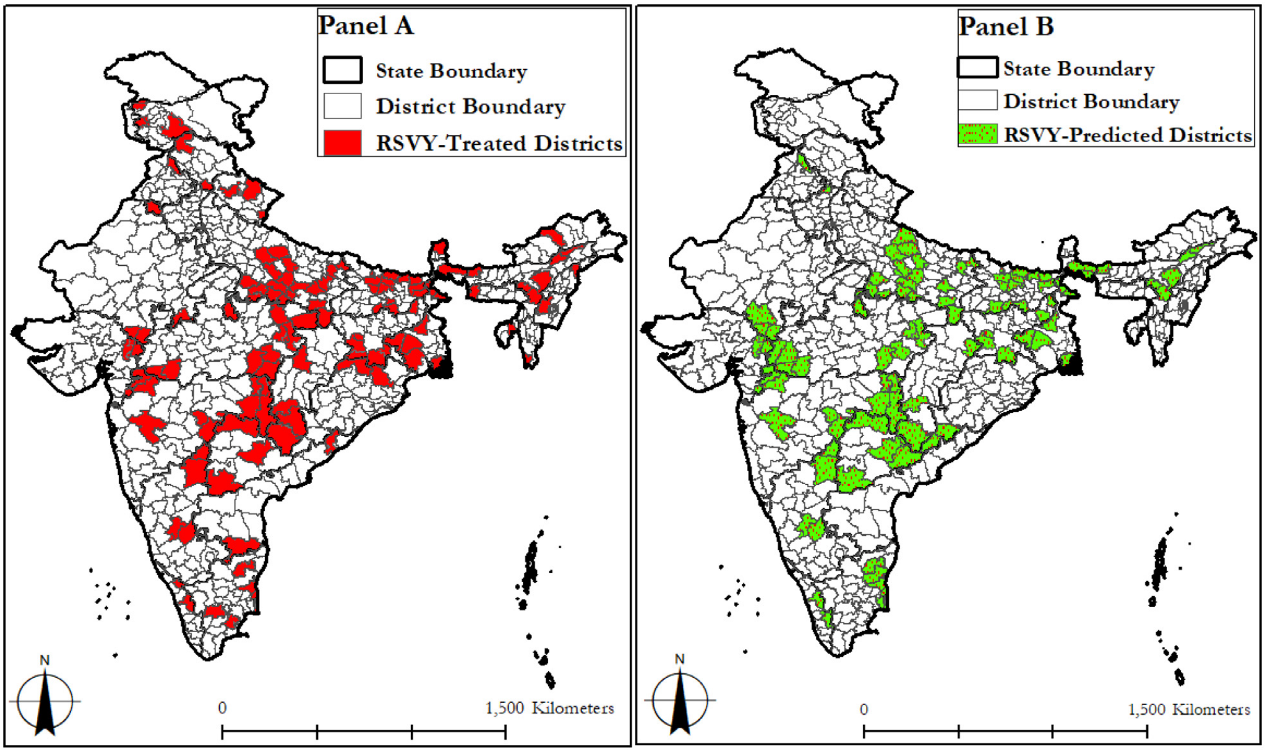
Unlike most place-based programs that are subject to non-random placement, RSVY had a uniquely complete and transparent allocation procedure that was explicitly documented by the Government of India. Following the allocation algorithm, the eligibility of districts under RSVY (i.e., treatment assignment) was based on a two-step process. In the first step, the Central Government determined the number of treatment districts that would be assigned to each of the 17 major Indian states.[5] The quotas were proportional to the state-level prevalence of poverty. In the second step, each state government, in accordance with the assigned quota, chose the specific districts to allocate the RSVY grants. The selection was based on an existing development ranking referred to as the Backwardness Index. This ranking index was public information, and a composite level of districts' economic underdevelopment was constructed from three historical indicators with equal weights:

[2] See Neumark and Simpson, 2015 for a thorough discussion of prior work on place-based policies in developed countries. For work Chinese Special Economic Zones (SEZs), see Wang, 2013; Lu and Wang, 2019; Cheng, 2014; Alder and Shao, 2016. For work on place-based policies in India, see Chaurey, 2017; Shenoy, 2018; Hasan et al., 2017.

[3] Authors' calculation based on the Economic Census of 2005.

[4] See http://www.planningcommission.nic.in/plans/stateplan/guid_rsvy.pdf.

[5] These 17 states are the "non-special category" states that comprised more than 97% of India's population in 2005.



Note: Panel A highlights the districts that received the RSVY grants (in red). Panel B highlights all districts predicted to receive the RSVY grants based on their Backwardness Ranking scores (in green). Selection criteria is discussed in Section 4. Thick black lines represent state boundaries. Thin black lines represent district boundaries.

Fig. 1. Maps of RSVY treated and predicted-treated districts. Note: Panel A highlights the districts that received the RSVY grants (in red). Panel B highlights all districts predicted to receive the RSVY grants based on their Backwardness Ranking scores (in green). Selection criteria is discussed in Section 4. Thick black lines represent state boundaries. Thin black lines represent district boundaries.

(i) value of output per agricultural worker (1990–1993); (ii) agriculture wage rate (1996–1997); and (iii) districts' percentage of low-caste populations - Scheduled Castes/ Scheduled Tribes (1991). The Backwardness Index ranked a total of 447 districts in the 17 major states with available data for all three indicators above. More details on the construction of the index are provided in the Appendix.

In addition to the above algorithm, the government had a separate list of 32 districts that were heavily affected by Maoist/Naxalite violence. These districts were automatically selected into the RSVY program.

3. Data and variable construction

We use several data sources for the analysis. We use the fourth (1998) and fifth (2005) rounds of the Economic Census (EC) for the village-level outcomes on employment, number of firms, and average firm size. For the microenterprise-level outcomes, we use information from rounds 56 (2000–01) and 62 (2005–06) of the National Sample Survey - Manufacturing Enterprises Schedule (NSS - Schedule 2.2). We use NSS – Employment and Unemployment Schedule (NSS - Schedule 10) rounds 55 (1999–2000) and 62 (2005–06) for individual-level outcomes. We control for baseline and time-invariant covariates at the village-level using information from the 2001 Population Census. Since RSVY was a rural infrastructure development and maintenance program, we restrict our analysis to the rural sample across all data sources.[6]

[6] The rural sample definition is based on Census of India and is consistent for all data sets used in our analysis.

3.1. Economic Censuses (EC)

The Economic Census is a complete enumeration of all economic establishments (with the exception of those engaged in crop production, defense, and government administration) conducted by the Ministry of Statistics and Program Implementation (MoSPI). Establishments are defined as any location, commercial or residential, where an economic activity is carried out. Both formal and informal establishments are enumerated, irrespective of firm size, including people working out of their homes. We use the fourth (1998) and fifth (2005) rounds of the Economic Censuses for our analysis. The Economic Census provides information on the number of employees by firm, but does not provide any other information on inputs or output. We restrict the sample to firms engaged in non-agricultural activities. We aggregate the microdata to the village level to get measures of employment, number of firms, and average firm size. We also divide the sample into formal firms and microenterprises. To be consistent with the other data sets (as well as the Factories Act of 1948) we define microenterprises as those with either less than 10 workers and that use electrical power, or less than 20 workers and do not operate with electrical power. The remaining are therefore categorized as formal firms.

3.2. National Sample Surveys (NSS)

The NSS - Schedule 2.2 is a nationally representative survey in India that provides detailed information on manufacturing microenterprises' business activities and performance. The survey also includes questions on firms' subjective perceptions of infrastructure and business environment during the year. Only small, "unorganized" firms with less than 10 workers and that use electrical power, or less than 20 workers and do not operate with electrical power, are included in this survey.[7] Micro firms meeting these employment criteria account for nearly 80% of India's manufacturing employment (Nataraj, 2011).[8] We use rounds 56 and 62 of the NSS manufacturing enterprise surveys. Since RSVY was introduced in June 2004, information from round 62 (2005–06) captures the short-run, post-treatment effects of this policy. Data from round 56 (2000–01) serve as the baseline period and allows us to perform falsification/placebo tests.[9]

We use this data set primarily to test for mechanisms. We use information on employment and whether the microenterprise was established less than three years ago (a proxy for firm age). We also use several self-reported measures on problems faced by firms. First, we use subjective infrastructure-related questions that ask enterprises whether they experienced problems related to power cuts during production in the previous year. The response to this question is a proxy for the availability and/or quality of electricity supply. Next, we use information from another question that asks firms if they experienced problems with availability of raw materials. We interpret the responses to this question as a proxy for rural connectivity. Finally, we combine problems such as lack of capital, marketing-related problems, harassment at the local level, and non-recovery of dues, and interpret these as problems unrelated to infrastructure.

To study the effects of the policy on individual-level outcomes, we use data from the National Sample Survey (Schedule 10) employment-unemployment rounds 55 (1999–2000) and 62 (2005–06).[10] These are nationally representative surveys covering all districts of India. Households in each district are sampled on a rolling basis over the agricultural year (July to June). The survey elicits daily employment and wage information for each household member for the 7 days preceding the interview. We use information on individual-level wages, total number of worker-days, monthly household consumption expenditure (MHCE), labor force participation status of the worker, whether the worker is unemployed, whether the worker is an agricultural wage worker, and whether the worker is employed in the non-agricultural sector. Finally, we also use the National Sample Survey employment-unemployment round 64 (Schedule 10.2, 2007–08) to study migration. The survey elicits information about the last usual place of residence for the household members. We define an internal migrant as one whose last usual place of residence was another district in the same state. An external migrant is defined as a person whose last usual place of residence was another state.[11]

3.3. Population Census (2001), geographic data, and night-time light intensity

We use the Primary Census Abstract (PCA) and Village Directory (VD) for the 2001 Population Census to construct pre-RSVY village-level variables. Specifically, we include baseline controls for all villages in our sample, such as total population, as well as access to representative public goods such as paved roads, electricity, and irrigation facilities.

In addition, we also test for district-level balance using various geographic characteristics. We use Geographic Information System (GIS) software to process the country's shapefiles provided by the Global Administrative Areas organization (www.gadm.org) and use the relevant district's geographic indicators such as area (in square kilometers), boundary perimeter (in kilometers), elevation (in meters),[12] and distance (in kilometers) to the nearest metropolitan cities.

Finally, we also use night-time light intensity measure at village-level. Nightlight luminosity is obtained from satellite imagery of the earth at night, recording light output at the 30 arc-second level, equivalent to approximately 1 square kilometer at the equator. For our empirical analysis, we process the raw GIS digital light composites to obtain a mean night-light radiance for each village from 1999 to 2008, which essentially cover the pre-, during-, and post-RSVY periods.[13] We then transform the village-level mean night-light radiance with an inverse hyperbolic sine transformation.

3.4. Summary statistics

Table 1 presents summary statistics for the important outcome and control variables used in this analysis. Across different data sets, we only focus on the rural sample for our analysis. For illustrative purposes, we employ a common bandwidth that consists of all districts located within 0.03 backward-score distance around the treatment assignment threshold ($|z| \leq 0.03$). According to the 2005 Economic Census, on average, 96.25 people were employed in non-farm activities, and there were 50.59 firms per village. Over 99 percent of the enterprises are categorized as microenterprises. For example, a typical village in our sample had around 50 microenterprises, and only 1 in 5 villages had a formal firm. Furthermore, around 88% of total village non-agricultural employment in our sample is concentrated in microenterprises, with 84.74 people working in microenterprises and 11.51 people working in formal firms. Microenterprises on average involved 1.69 workers, whereas formal firms on average had 54.65 workers.

We focus on microenterprises in Panel B from the Unorganized Manufacturing frame (Schedule 2.2) of the NSS 2000–01. On average, a microenterprise in our sample employs 2.51 workers, with a standard deviation of 3.07. The small scale of microenterprises should be kept in mind while interpreting the results. With respect to infrastructure-related problems, 17 percent of firms in the sample reported having problems related to a power cut in the year of the survey, and 15 percent reported having problems with access to raw materials. 10 percent of microenterprises reported to have been established less than 3 years ago.

An average worker in our sample (Panel C) earns 416 Rs/week, and an average household spends around 2508 Rs/month. Finally, as per the 2001 Demographic Census, 62 percent of villages in an average district in the sample have paved roads, and approximately 77 percent have electricity coverage.

4. Empirical strategy

4.1. Reconstruction of Backwardness Score Index

Since the RSVY selection process followed a transparent, score-based rule, we evaluate the effects of the program using a Regression Discontinuity Design (RD). First, we take the actual

[7] Essentially, small firms meeting these criteria are not required to register with the state governments under India's 1948 Factories Act, hence are often referred to as "unregistered," "unorganized," or "informal" firms.

[8] For two examples of previous papers using the NSS Sch. 2.2., see Nataraj (2011) and Hsieh and Klenow (2014).

[9] The National Sample Survey Organization (NSSO) did not conduct any other similar survey (Schedule 2.2) for microenterprises between rounds 56 and 62.

[10] We do not use NSS employment-unemployment round 61 (Schedule 10) because it was conducted in 2004 and almost overlaps with the first round of RSVY grants. Furthermore, the use of employment-unemployment Round 62 allows us to be consistent with the microenterprise survey that was conducted in the same year (2005–06).

[11] Note that we use NSS round 55 for pre-RSVY migration outcomes.

[12] For topographic information, we use the GTOPO 30 Arc-Second Elevation global raster data set developed and maintained by the U.S. Geological Survey's Center for Earth Resources Observation and Science (EROS).

[13] We collect the population raster dataset named Gridded Population of the World, Version 3 (GPWv3) from the Socioeconomic Data and Applications Center (SEDAC)— a data center in NASA's Earth Observing System Data and Information System (EOSDIS)—hosted by CIESIN at Columbia University.

Table 1
Summary statistics for the main variables (representative bandwidth of $|z| \leq 0.03$).

	Observations	Mean	SD	Source
Panel A: Village employment and firms				
1. Employment				
Total Employment	92,677	96.25	232.73	EC 2005
Formal Employment	92,677	11.51	76.87	EC 2005
Informal Employment	92,677	84.74	194.49	EC 2005
2. Number of firms				
Total Firms	92,677	50.59	111.09	EC 2005
Formal Firms	92,677	0.22	1.32	EC 2005
Informal Firms	92,677	50.37	110.60	EC 2005
3. Average firm size				
Total Firms	92,677	1.92	2.89	EC 2005
Formal Firms	8,942	54.65	77.33	EC 2005
Informal Firms	92,633	1.69	0.75	EC 2005
Panel B: Firms				
1. Microenterprises				
Employment (level)	8,580	2.51	3.07	NSS62 -Sch. 2.2
Problem Related to Power Cut (%)	8,580	0.17	0.38	NSS62 -Sch. 2.2
Problem Related to Accessing Materials (%)	8,580	0.15	0.36	NSS62 -Sch. 2.2
Established Less than Three Years Ago (%)	8,580	0.10	0.30	NSS62 -Sch. 2.2
Panel C: Individual employment, wages, migration				
Labor Force Participation (%)	28,256	0.577	0.494	NSS62 - Sch. 10
Unemployed (%)	28,256	0.017	0.13	NSS62 - Sch. 10
Agricultural Workers (%)	28,256	0.332	0.471	NSS62 - Sch. 10
Non-agricultural Workers (%)	28,256	0.227	0.419	NSS62 - Sch. 10
Workload (days/week)	39,143	6.64	1.04	NSS62 - Sch. 10
Wage (Rs./week)	6,232	415.91	455.58	NSS62 - Sch. 10
Consumption Expenditure (MPCE; Rs./month)	8,246	2,507.94	1,580.05	NSS62 - Sch. 10
Internal (within-state) migration (%)	126,881	0.086	0.281	NSS64 - Sch. 10
External (across-state) migration (%)	126,881	0.028	0.164	NSS64 - Sch. 10
Panel D: Village-level characteristics				
Population	92,677	1684.15	1986.95	DC 2001
Paved Roads (%)	92,677	62.18	48.51	DC 2001
Electricity Coverage (%)	92,677	77.02	42.07	DC 2001
Irrigated Area (%)	92,677	45.69	37.27	DC 2001
Panel E: District-level characteristics				
1. Geographic characteristics:				
Area (km sq.)	115	5,328.36	4,112.39	GADM
Elevation (m)	115	214.80	192.28	GTOPO30
Distance to nearest city (km)	115	116.51	57.13	GADM
2. Socio-demographic characteristics:				
Number of villages per district	115	1,699.41	1,202.35	DC 2001
Share of SC/ST population (% 1991)	115	26.55	12.48	PC 2003
Output per Agricultural Worker (Rs. 1990–93)	115	5,750.84	4,350.64	PC 2003
Agricultural Wage Rate (Rs. 1996–97)	115	32.76	8.42	PC 2003
3. RD running variable				
Backwardness Composite Score	115	0.329	0.069	PC 2003
Distance to Cutoffs (z)	115	0.006	0.013	PC 2003

Note: This table shows summary statistics for the main outcomes and control variables used in the analysis. The sample includes all firms operating in the RD restricted bandwidth of districts with re-centered Backwardness Index Scores (z) within 0.03 points from the cutoff, i.e. $|z| \leq 0.03$. Sources: EC 2005: Economic Census 2005; NSS 62 - Sch. 2.2: National Sample Survey, Round 62 (2005–06) – Unorganized Manufacturing Enterprises; ASI 2005: Annual Survey of Industries (2005); NSS 62 - Sch. 10: National Sample Survey, Round 62 (2005–06) – Employment-Unemployment; DC 2001: Demographic (Population) Census 2001; GADM: Database of Global Administrative Areas; PC 2003: Planning Commission (2003).

number of districts allocated to each of the 17 major states as given. Our main analysis ultimately relies on within-state comparisons of the marginal districts around the state-specific cutoff scores. Therefore, our approach is internally valid when we take the number of districts assigned to each state as-is. We also control for state fixed effects to account for any unobserved variation at the state level that might be jointly correlated with both the outcome variables and the district's treatment status.

Next, we reconstruct the entire selection criteria based on the Backwardness Score Index of districts in each state from the second step of the assignment algorithm.[14] Provided with the allotted number of districts by the central government (from the first step), state governments were supposed to choose the "most backward" districts for selection, based on districts' backwardness scores. This composite index was constructed from three historical parameters with equal weights, including (i) value of output per agricultural worker (1990–1993); (ii) agricultural wage rate (1996–1997); and (iii) districts' percentage of low-caste populations - Scheduled Castes/ Scheduled Tribes (1991) (Planning Commission, 2003). Next, using the composite backwardness score, we construct the running variable employed in our RD setting following the detailed steps below:

1. For each of the 17 states with available backwardness index data, we construct the district's composite index score and denote it as x_{ds}. Subscript d denotes "district" and s denotes "state."
2. Denoting the state's delegated number of RSVY-eligible districts as k_s, we obtain the cutoff score in state s, which is the index score associated with the k_s^{th} district (i.e., the "cutoff" district) in that state in ascending order of x_{ds}. We denote the cutoff score for state s as x_{ds}^k.

[14] See Appendix for more details.

3. We re-center the sequence x_{ds}, so that the cutoff district in the sequence would receive a re-centered distance score of 0. That is:

$$z_{ds} = x_{ds} - x^{k}_{ds}. \tag{1}$$

The district's state-specific re-centered distance score, z_{ds}, serves as the running variable in our subsequent RD regressions. By design, districts to the left of the cutoff (i.e., those with non-positive distance scores) are more backward than the state's cutoff district, and should be RSVY-eligible according to the selection rule.

We obtain the full list of districts that should have been granted RSVY funding had there been perfect compliance with the central government's guidelines. This list includes all districts with state-specific backwardness scores below their state's cutoff. More explicitly, the backwardness index data is available for 447 districts for the 17 major states in India.[15] In our sample, of the 147 districts that actually received the RSVY grants, 32 districts were affected by left-wing extremist violence, and their selection was not based on the backwardness index. Having removed these 32 districts from our sample, we are left with 115 districts that received the grants. These 115 districts are shown in Fig. 1 Panel A. Out of these 115 RSVY districts, 19 (12.9%) belong to states with missing ranking data. To the extent that the actual RSVY assignment to these 19 districts was endogenous (i.e., they were funded without having Backwardness Index information), we remove them from our estimation sample. This leaves us with 96 districts that received RSVY grants and had ranking data available. For these 96 districts, the assignment algorithm had a prediction accuracy of 80.2%, correctly predicting treatment status for 77 of these districts. This implies that some districts that were not supposed to receive the RSVY grants did in fact receive them. These predicted districts are shown in Fig. 1 Panel B. Our prediction accuracy is distinctly different from a random draw of districts from the overall pool (21.48%),[16] and provides credence to our approach. Quantitatively, our estimates should therefore provide a lower-bound of the actual impact of RSVY.

It is worth noting that the same Backwardness Index was used in the implementation of NREGA (National Rural Employment Guarantee Act)—an employment guarantee program implemented in 2006–07, three years after RSVY (Bhargava, 2014; Zimmermann and Khanna, 2017; Zimmermann, 2017; Hari and Raghunathan, 2017). Compared to these studies, our approach differs in one important dimension. Instead of utilizing the state-specific districts' ordinal ranks as the running variable, we adopt the districts' backwardness scores, using the score distance to cutoff as our running variable. From a technical perspective, continuous score distances allow us to deviate from using a discrete running variable in the RD framework. Adopting discrete rankings as a running variable essentially limits the available choices of bandwidth size in estimation, and/or the ability to obtain reliable estimates of the Average-Treatment-Effect (ATE) or the associated standard errors (Lee and Card, 2008; Kolesár and Rothe, 2018).

4.2. Empirical design

Our empirical analysis follows the parametric Regression Discontinuity Design functional form as suggested by Imbens and Lemieux (2008):

$$y_{ids} = \beta_0 + \beta_1 \mathbf{1}\{z_{ds} \leq 0\} + \delta(z_{ds}) + \gamma(X_{ids}) + \pi_s + \varepsilon_{ids} \tag{2}$$

where the subscripts refer to a village- (or firm-/individual-/household-) level observation i, in district d, in state s. Thus, y_{ids} is the village- (or firm-/individual-/household-) level outcome variables of interest. z_{ds} is the constructed re-centered score distance discussed in the previous section, which serves as the running variable in our RD design. Following Gelman and Imbens (2019), $\delta(z_{ds})$ is a polynomial function of the score variable that allows for both linear and quadratic specifications. $\mathbf{1}\{z_{ds} \leq 0\}$ is the predicted treatment indicator that equals one if $z_{ds} \leqslant 0$ (i.e., when the district has a non-positive state-specific score distance to the eligibility cutoff), meaning the district is economically backward enough to receive RSVY grants under the assignment guideline.

In all specifications, we control for state fixed effects π_s since eligibility cutoffs are state-specific. X_{ids} includes covariates that depend on whether the regression is at the village, firm, or individual level. For our village-level regressions using the Economic Census, we include village-level controls from the 2001 Population Census such as total village population, whether the village had paved roads, whether the village had electricity, and the share of area of the village that is irrigated. For the NSS manufacturing firm-level regressions, covariates include the microenterprises' physical operating structure (inside or outside the household, whether with fixed premises or not), and owner's gender and highest education level. In the firm-level regressions we also control for industry fixed effects. For the individual-level regressions we control for age, education, occupation, social group, and religion. Finally, ε_{ids} is the error term clustered at the district level.

The main coefficient of interest is β_1. This coefficient represents the discontinuous changes in outcomes between treated and comparison districts located close to the eligibility cutoff. Under the standard RD identification assumption that marginal districts around the discontinuity are as good as random, β_1 represents the Local Average Intent-to-Treat ("ITT") effect of the policy.

In the Appendix, we also show estimates of the Treatment-on-the-Treated ("TOT") effects. This is to address the concern that there were some differences between the districts that were supposed to receive RSVY grants as per the scores, and those that actually received it. Specifically, we run instrumental variable regressions in the form of a Fuzzy Regression Discontinuity Design, where we instrument the actual RSVY selection with the predicted treatment indicator: $\mathbf{1}\{z_{ds} \leq 0\}$. The first stage of the Fuzzy RD requires that there is a discontinuity in the probability of receiving RSVY at the cutoff. Fig. 2 shows this discontinuity graphically. It plots the probability of receiving RSVY as a function of the running variable (re-centered score distance). It is visually clear that the average probability of receiving RSVY decreases discontinuously to the right of the cutoff.

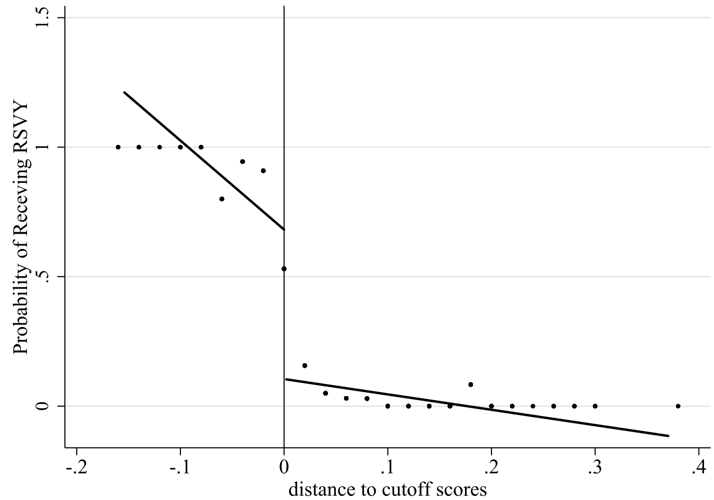
Another key aspect related to regression discontinuity designs is the choice of bandwidth. We test for the sensitivity of our RD estimates on bandwidth selection by reporting the coefficients across three bandwidths, (0.02, 0.025, 0.03), around our RD threshold. Furthermore, we also present a sensitivity analysis associated with all of our main outcomes of interest in the [0.02, 0.03] range, and show our main results using the Calonico et al. (2014) data-driven optimal bandwidth in Table A6.

4.3. Validating the identification assumptions

Treatment assignment at the threshold is only "as good as random" when the polynomial function of the running variable is smooth, or continuous, across the RD threshold. In essence, districts must not be able to manipulate their relative backwardness scores so as to determine their treatment status. This assumption is reasonable because the backwardness score index was constructed using historical development parameters collected in the early 1990s, roughly a decade before the introduction of the RSVY

[15] Data on economic under-development parameters were unavailable for the remaining Indian states classified as "special category" or union territories. Therefore, it is unclear how these state governments selected eligible RSVY districts.

[16] Randomly drawing 96 districts from the pool of 447 districts for which ranking data is available results in a prediction accuracy of 21.48%.



Note: The graph plots the probability of receiving RSVY treatment by district over the RD running variable (district's standardized distance scores from the cutoff). Each point represents a bin-average probability. Scores on the left of the cutoff constitute RSVY-eligible districts.

Fig. 2. Discontinuity in treatment probability (first stage). Note: The graph plots the probability of receiving RSVY treatment by district over the RD running variable (district's standardized distance scores from the cutoff). Each point represents a bin-average probability. Scores on the left of the cutoff constitute RSVY-eligible districts.

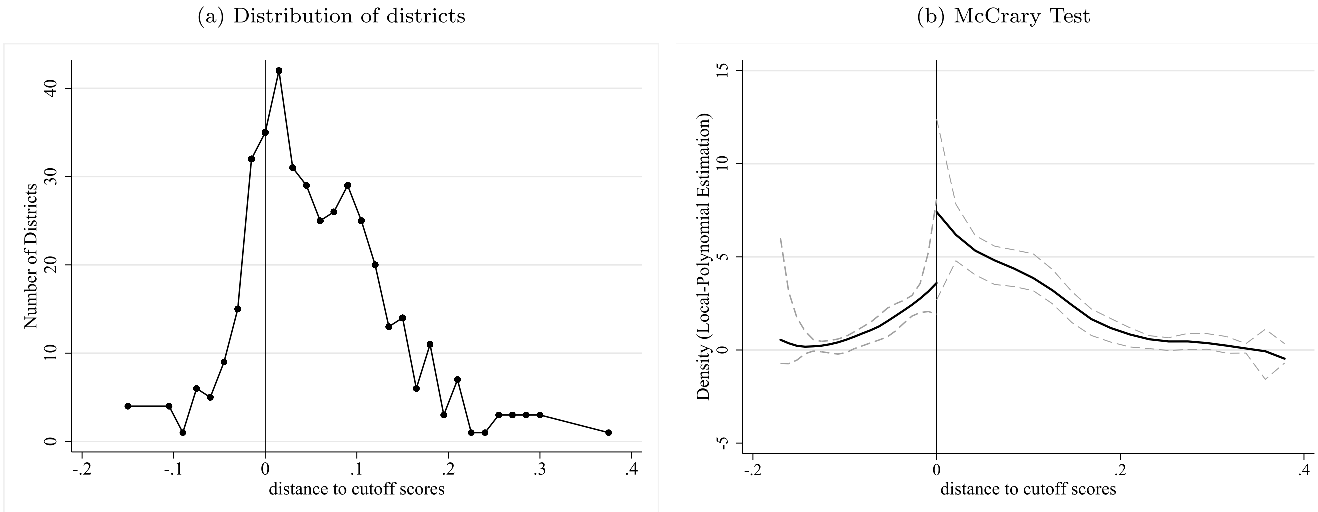
program, thus limiting the possibility of districts strategically misreporting information. Regardless, we visually check for treatment status manipulation in Fig. 3a. This figure plots the distribution of districts over the re-centered distance score measure. We also conduct a McCrary (2008) density test for potential manipulation of the running variable. If there was strategic manipulation, we should see visual evidence of "bunching" in the density of the assignment variable around the treatment cutoff. Fig. 3b shows no such bunching and the kernel density function of the re-centered distance scores is smooth around the threshold. Consistent with the visual evidence, the McCrary (2008) test does not reject the null hypothesis of no discontinuity in the density of districts (McCrary test p-value is 0.2943).

Next, in Table A1 we check whether the various outcome and control variables used in this analysis are balanced around the cutoff. In panel A, we focus on the outcome variables (employment, number of firms, and average firm size) from the EC 1998. In panel B, we look at the baseline firm-level outcomes from NSS 2000–01. Panel C includes the individual- and household-level characteristics, and panel D includes socio-economic, infrastructure, and geographic characteristics. We do not find any statistically significant differences around the cutoff for any of the covariates. The smoothness of these district covariates at baseline around the cutoff provides further support for the identification assumptions needed for a regression discontinuity design.

Another potential threat to identification would be if there were contemporaneous public programs with a similar development focus that were also implemented on the basis of the district backwardness ranking index. To the best of our knowledge, no such district-level program existed during this time. The RSVY program was the first national public infrastructure development initiative that the Government of India introduced, that adopted a transparent assignment formula on the basis of the backwardness index. The other large-scale public/development projects that used the backwardness index to determine eligibility of districts were the Backward Regional Grant Fund (BRGF), and the National Rural Employment Guarantee Act (NREGA). BRGF was introduced in 2007 and in fact subsumed the RSVY program. It extended the total number of eligible districts for infrastructure grants to 250 districts. The first phase of NREGA was implemented in April 2006, covering the 200 most backward districts. Both programs started at least two years after the introduction of RSVY. Hence, these programs do not contaminate our results, at least in the sample considered in our paper.

However, two other village-level infrastructure programs were also introduced by the Government of India in the 2000s—(i) Pradhan Mantri Gram Sadak Yojana (Prime Minister's Village Road Program, or PMGSY) introduced in 2000 and (ii) Rajiv Gandhi Grameen Vidyutikaran Yojana (Prime Minister's Rural Electrification Program or RGGVY) introduced in 2005.[17] Both these programs were implemented based on village-level population cutoffs. For example, PMGSY targeted roads to villages with population exceeding two discrete thresholds (500 and 1,000), and RGGVY targeted electrification to villages with a population larger than 300 people. There are multiple reasons why these two programs do not affect our empirical setting and results. First, the source of identifying variation for RSVY is the district-level distance to cutoff score, which is different than village-level population cutoffs used for PMGSY and RGGVY. Second, in our regressions, we control for village population. Third, for PMGSY and RGGVY to affect our results, it must be the case that the number or share of villages with these population cutoffs (300, 500, 1000) were differentially higher in RSVY treated districts than in the control districts. We check for this in Appendix Table A2 and do not find any such differences with respect to PMGSY- and RGGVY-eligible villages. Finally, in subSection 5.4.2, we find that RSVY led to higher employment and number of firms in villages that already had paved roads and electricity prior to the policy, and had no effect on un-electrified and unconnected villages. This bolsters the view that our results are not driven by PMGSY and RGGVY.

[17] See Asher and Novosad (2020) and Aggarwal (2018) for more detail on PMGSY and Burlig and Preonas (2016) for more detail on RGGVY.



Note: Panel (a) plots the distribution of districts over the RD running variable. Panel (b) plots a non-parametric regression on each side of the distribution following McCrary (2008) and tests for discontinuity in the density around the cutoff.

Fig. 3. Distribution of districts around the cutoff. Note: Panel (a) plots the distribution of districts over the RD running variable. Panel (b) plots a non-parametric regression on each side of the distribution following McCrary (2008) and tests for discontinuity in the density around the cutoff.

5. Results

We begin by presenting results on the effects of RSVY on village-level employment, number of firms, and average firm size from the Economic Census (EC). Then, we look at the effect of RSVY on individual-level wages, hours worked, and household monthly consumption expenditure, followed by outcomes related to labor force participation, migration, unemployment status, and sectoral realignment. After discussing the main findings, we provide evidence for plausible mechanisms driving these effects. Finally, we conduct robustness tests for our results. For all the outcome variables, we present intent-to-treat (ITT) estimates using linear and quadratic functions of the running variable, across three alternate bandwidths – 0.02, 0.025, and 0.03. The tables for the main results also show the mean and standard deviations for the control group, and all results are interpreted with respect to these. In the Appendix, we show the corresponding treatment-on-treated (TOT) estimates. For brevity, we present our TOT results for one representative bandwidth (0.025) across linear and quadratic specifications. Before discussing the results in greater detail, we note that our interpretations are based on the consistent pattern of coefficients across specifications rather than the statistical significance of every coefficient.

5.1. Village-level employment, number of firms, and average firm size

In Table 2 Panel A, we look at the ITT effects of RSVY on employment at the village-level in non-farm activity (total employment, formal employment, and informal employment) from the Economic Census 2005. Note that these are short-term impacts of the policy, measured approximately a year and a half after the introduction of RSVY. We find that total employment in non-farm activity in villages increased by 13.7–16% across columns 1 through 6 in RSVY-treated districts relative to control districts. The results are quantitatively similar across the different bandwidth choices and specifications (linear and quadratic), albeit not statistically significant across all specifications. This is equivalent to 14–16.3 additional jobs in firms per village. To understand whether the effects of RSVY are different across the firm size distribution, we further divide village employment in all non-farm activity into employment in microenterprises and in formal firms. We find that the effect on village-level employment in all firms is almost completely due to employment in microenterprises. Village employment in microenterprises increases in RSVY districts relative to control districts by 15–17.6% across the columns. This is equivalent to 13–15.5 additional jobs in microenterprises per village. However, there is no statistically significant change in village employment in formal firms between treated and control districts. The corresponding TOT results for village-level employment across the firm-size distribution is shown in Appendix Table A3. We find that village-level employment in both all firms and microenterprises increase, but there is no change in employment in formal firms. Figs. 4 (a), (c), and (e) show the graphical representations of the results from Table 2 Panel A. Each scattered point in the graph represents bin-averaged values of village-level employment (IHS-transformed) after partialling out state fixed effects (as shown in Eq. 2). Figs. 4 (b), (d), and (f) show the visual representation of the results using pre-RSVY data from 1998, where we find no discontinuous jumps at the cutoff. This increase in village-level non-farm employment and informal employment in microenterprises may be due to an increase in the number of firms (extensive margin), or an increase in average firm size (intensive margin). We look at these margins in panels B and C, respectively.

In Table 2 Panel B, we look at the extensive margin results for the total number of firms, number of formal firms, and number of informal firms (microenterprises) at the village-level. We find that the total number of firms differentially increased in villages in RSVY districts relative to control districts by 13%-17.6%. Note again that these results are not statistically significant across all specifications. In numbers, these coefficients translate to an increase of 7–9.3 new firms in villages in RSVY districts. Similar to our earlier results, this growth was completely driven by an

Table 2
Village-level employment, number of firms, and average firm size.

		Linear (1)	Quadratic (2)	Linear (3)	Quadratic (4)	Linear (5)	Quadratic (6)
Panel A: Village employment							
1. Total Employment	RD Estimate	0.159*	0.137	0.141	0.150*	0.153*	0.160*
	S.E.	(0.0953)	(0.0922)	(0.0911)	(0.0894)	(0.0881)	(0.0857)
	Observations	73,335	73,335	83,356	83,356	92,677	92,677
	Mean [SD]	100.78 [249.31]		93.68 [237.27]		102 [257.82]	
2. Formal Employment	RD Estimate	0.0578	0.0179	0.0239	0.0358	0.0217	0.0232
	S.E.	(0.0848)	(0.0850)	(0.0781)	(0.0743)	(0.0753)	(0.0747)
	Observations	73,335	73,335	83,356	83,356	92,677	92,677
	Mean [SD]	12.2 [76.25]		11.88 [80.89]		13.63 [89.72]	
3. Informal Employment	RD Estimate	0.169*	0.150*	0.153*	0.161*	0.167*	0.176**
	S.E.	(0.0934)	(0.0898)	(0.0880)	(0.0870)	(0.0862)	(0.0837)
	Observations	73,335	73,335	83,356	83,356	92,677	92,677
	Mean [SD]	88.58 [213.77]		81.80 [199.02]		88.37 [213.83]	
Panel B: Number of firms							
1. All Firms	RD Estimate	0.150	0.130	0.153*	0.156*	0.170**	0.176**
	S.E.	(0.0967)	(0.0915)	(0.0854)	(0.0863)	(0.0856)	(0.0845)
	Observations	73,335	73,335	83,356	83,356	92,677	92,677
	Mean [SD]	53.64 [28.91]		49.44 [119.57]		52.46 [123.78]	
2. Formal Firms	RD Estimate	0.0119	0.00078	0.00216	0.00572	0.00255	0.00283
	S.E.	(0.0243)	(0.0246)	(0.0223)	(0.0212)	(0.0218)	(0.0216)
	Observations	73,335	73,335	83,356	83,356	92,677	92,677
	Mean [SD]	0.24 [1.28]		0.22 [1.27]		0.25 [1.35]	
3. Informal Firms	RD Estimate	0.151	0.133	0.155*	0.158*	0.173**	0.178**
	S.E.	(0.0969)	(0.0917)	(0.0855)	(0.0864)	(0.0858)	(0.0847)
	Observations	73,335	73,335	83,356	83,356	92,677	92,677
	Mean [SD]	53.4 [128.44]		49.22 [119.11]		52.22 [123.25]	
Panel C: Average firm size							
1. All Firms	RD Estimate	0.00666	0.00432	−0.0124	−0.00755	−0.0176	−0.0158
	S.E.	(0.0341)	(0.0349)	(0.0334)	(0.0302)	(0.0302)	(0.0299)
	Obsevations	73,335	73,335	83,356	83,356	92,677	92,677
	Mean [SD]	1.97 [3.66]		1.97 [3.39]		1.97 [3.39]	
2. Formal Firms	RD Estimate	0.0225	0.0323	−0.00580	−0.00631	0.0138	0.0109
	S.E.	(0.0466)	(0.0518)	(0.0474)	(0.0474)	(0.0429)	(0.0418)
	Obsevations	7,100	7,100	7,627	7,627	8,942	8,942
	Mean [SD]	53.38 [73.81]		55.27 [75.38]		55.27 [75.38]	
3. Informal Firms	RD Estimate	0.0146	0.0148	−0.00187	0.00213	−0.00591	−0.00345
	S.E.	(0.0270)	(0.0276)	(0.0271)	(0.0244)	(0.0243)	(0.0236)
	Obsevations	73,302	73,302	83,313	83,313	92,633	92,633
	Mean [SD]	1.71 [0.77]		1.71 [0.76]		1.71 [0.76]	
Bandwidth		0.02		0.025		0.03	
State Fixed Effects & Village Controls		Yes	Yes	Yes	Yes	Yes	Yes

Note: The dependent variables include total employment (panel A), total number of firms (panel B), and average firm size (number of workers per firm; panel C). Inverse hyperbolic sine transformation is applied to all outcome variables. Odd columns show estimates from first-order RD polynomial specifications and even columns from second-order polynomial specifications. Standard errors are clustered at the district level.

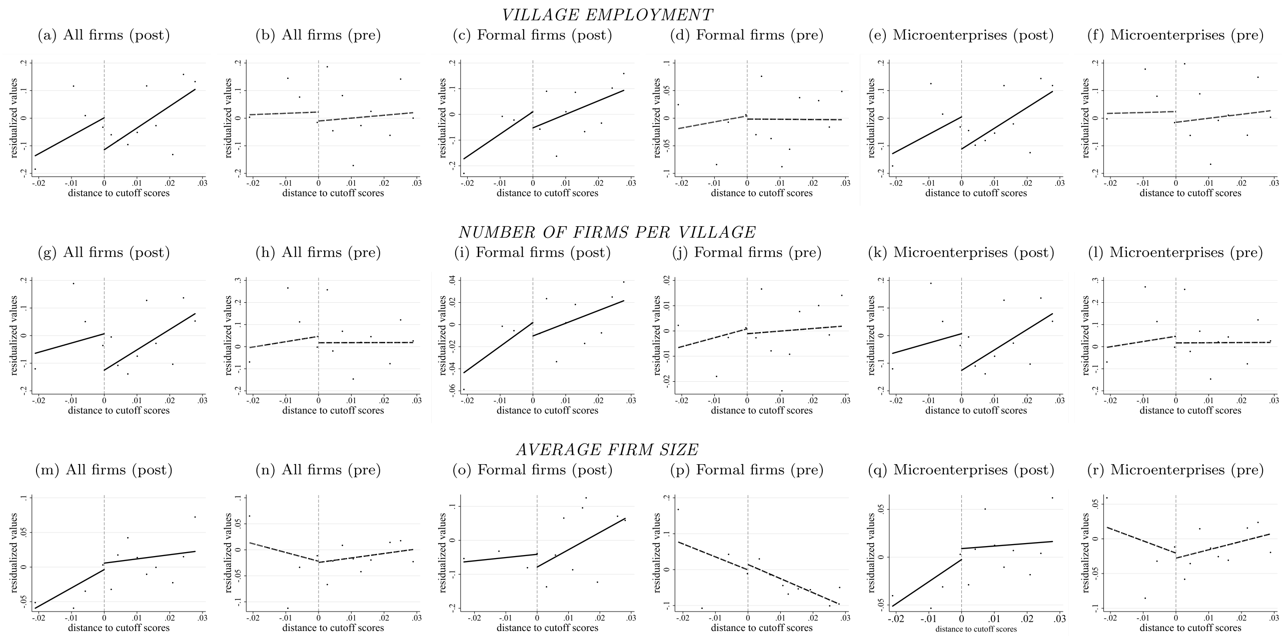
increase in the number of microenterprises. The number of microenterprises in RSVY districts increased by 13.3–17.8%. These correspond to 7–9.3 additional microenterprises in villages in RSVY districts. In contrast, we find no effects of RSVY on the number of formal firms. The graphs corresponding to these results are shown in Figs. 4 (g), (i), and (k). Figs. 4 (h), (j), (l) show the visual representation of the results using pre-RSVY data from 1998, where we find no discontinuous jumps at the cutoff.

Finally, in panel C, we look at the intensive margin results for the average firm size (number of workers per firm) for all firms, formal firms, and informal firms (microenterprises) at the village level. The results are statistically insignificant across all bandwidths and specifications. In essence, we find no change in the average firm size for all firm types between the treated and control districts. The graphs for both the post-RSVY and pre-RSVY periods are shown in Figs. 4 (n) through (r).

Taken together, the sum of our results suggests that overall village employment in RSVY districts increased relative to control districts, mainly driven by increases in village-level employment in microenterprises. Decomposing these overall increases in employment, we find that the number of microenterprises increased, but the number of formal firms did not change. However, the average firm size for all firms, formal firms, or microenterprises did not change as a result of RSVY.[18]

Although we show that RSVY led to increases in the village-level employment in microenterprises and the number of microenterprises, it is important to distinguish between direct effects originating from construction and maintenance work mandated by RSVY relative to the economic effects originating from the presence of infrastructure in equilibrium. To understand this, in Table A8 Panel A, we first check whether the jobs created in firms were involved in doing the infrastructure work mandated under RSVY. We do not find any increases in the number of firms and employment in firms at the village-level in the construction and public administration sector. This suggests that the overall increases were not due to construction and public administration related work for RSVY. In Panel B, we check whether the increases in the number of

[18] This could be due to several reasons. One explanation is that only one in five villages in our sample had a formal firm, and most of them locate in or near urban areas. This makes it unlikely that a rural infrastructure scheme such as RSVY would have an impact on large firms. Another potential explanation is that formal firms had already adjusted to unreliable electricity by investing in generators (Rud, 2012b; Allcott et al., 2016) or had already located near major road or railroad networks.



Note: This figure includes RD graphs using the 4th and 5th rounds of the Economic Census. All graphs with solid fitted lines correspond to post-RSVY outcomes (2005). All graphs with dashed fitted lines correspond to pre-RSVY outcomes (1998; placebo tests). The rows show village-level employment, number of firms per village, and average firm size. Each scatter point represents the bin average of residualized values of the outcome variables.

Fig. 4. RD graphs for village-level outcomes from Economic Censuses. Note: This figure includes RD graphs using the 4th and 5th rounds of the Economic Census. All graphs with solid fitted lines correspond to post-RSVY outcomes (2005). All graphs with dashed fitted lines correspond to pre-RSVY outcomes (1998; placebo tests). The rows show village-level employment, number of firms per village, and average firm size. Each scatter point represents the bin average of residualized values of the outcome variables.

microenterprises as well as the employment in microenterprises were in sectors that were likely to benefit from the improvements in infrastructure. We conduct our analysis on broad sectors - manufacturing, retail, and transportation. We find increases in employment and number of firms for microenterprises in the manufacturing and transportation sector and weak results for the retail sector. These results increase our confidence that the increases in employment and number of firms were due to sectors benefiting from the infrastructure improvements, relative to work directly mandated under RSVY.

5.2. Wages, employment, and household consumption expenditure

The firm- and village-level results show that RSVY led to an increase in the number of microenterprises, which in turn increased overall village-level employment in microenterprises. Next, we look at the effects of RSVY on individual wages, the number of days worked in the last 7 days (intensive margin), and monthly household consumption expenditure in Panels A, B, and C of Table 3. Consistent with our village-level results, we find that wages increased by 11–15.7%, number of days worked increased by 2.3–3.1% for individuals, and monthly household consumption expenditure increased by 4.3–8% in RSVY districts relative to control districts in the short run (approximately two years after the policy). These coefficients translate to an increase of 45–63 Rupees/week in wages, and an increase of 104–194 Rupees/month in household consumption expenditure. These results are also graphically shown in Figs. 5 (a), (c), and (e), and the pre-RSVY results are shown in Figs. 5 (b), (d), and (f).

5.3. Labor force participation and sectoral reallocation

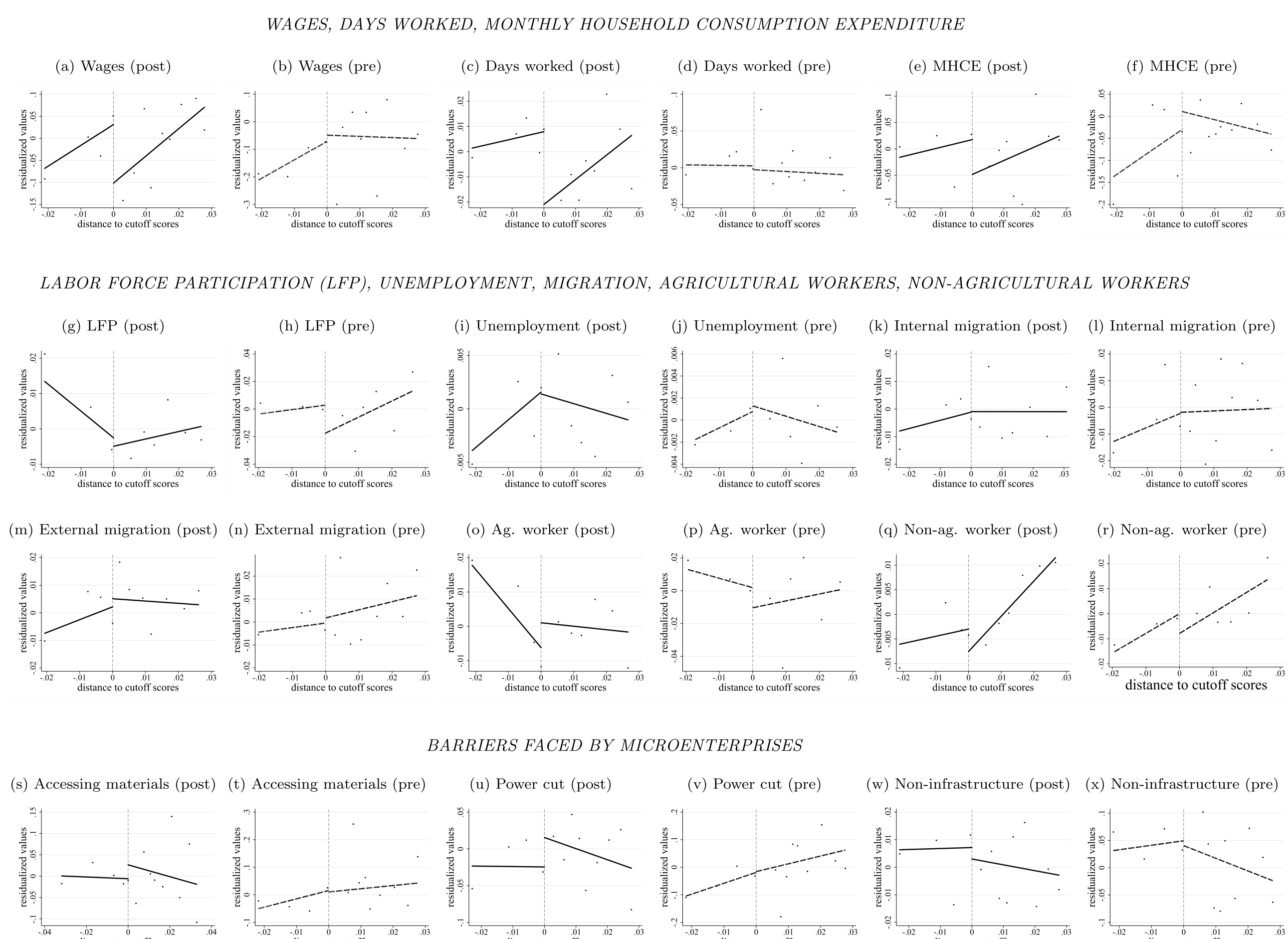
Thus far, we have shown evidence of increases in village-level microenterprise employment and corresponding increases in wages, days worked, and monthly consumption expenditure. However, since the Economic Census only includes non-farm activity, it is unclear whether RSVY led to an overall increase in economic activity or whether there was reallocation of workers from other regions and sectors. First, evidence from Table 2 suggests that there was no reallocation of workers from formal firms to microenterprises in the treated districts. Next, we focus on changes in labor force participation, unemployment, and in-migration between treated and control districts in response to RSVY. In Table 4 Panel A, we look at whether there was a change in the proportion of people that were in the labor force, unemployed, or in-migrants from other districts/states. We define an internal migrant as an individual that moved from within the same state but from a different district, and an external migrant as one who moved from another state. Across bandwidths and specifications, we do not find any statistically significant proportionate difference in labor force participation, unemployment, or in-migration between treated and control districts.[19] The lack of in-migration from other districts and states also suggests a lack of spatial spillovers in employment due to RSVY. Next, in panel B, we focus on whether there was sectoral reallocation of workers from agriculture to non-agricultural activities. We find a decline in the probability of being an agricultural worker of 2.5–4.4% in the treated districts. Interestingly, we find a symmetric increase in the proportion of non-agricultural workers. The proportion of non-agricultural workers in treated districts increased by 2.4–2.7%. The corresponding graphs, both before and after RSVY, are shown in Figs. 5 (g) through (x). These results provide suggestive evidence that agricultural wage workers may have substituted into non-agricultural work in microenterprises. The finding is also consistent with Asher and Novosad (2020), which shows that PMGSY led to a movement of rural workers out of agri-

[19] In general, earlier literature in India has found low migration response to economic factors (Munshi and Rosenzweig, 2016; Topalova, 2010).

Table 3
Wages, days worked, and monthly household consumption expenditure.

	Linear (5)	Quadratic (6)	Linear (3)	Quadratic (4)	Linear (1)	Quadratic (2)
Panel A: Wages						
RD Estimate	0.157**	0.155**	0.149***	0.139**	0.123***	0.110**
S.E.	(0.0607)	(0.0616)	(0.0554)	(0.0560)	(0.0467)	(0.0469)
R-square	0.501	0.435	0.500	0.436	0.499	0.432
Mean [SD]	401.10 [447.30]		404.55 [416.63]		411.02 [456.82]	
Observations	4,914	4,914	5,422	5,422	6,232	6,232
Panel B: Days worked (in the last 7 days)						
RD Estimate	0.0310***	0.0306***	0.0286**	0.0277**	0.0233**	0.0231**
S.E.	(0.0101)	(0.0104)	(0.0111)	(0.0109)	(0.00989)	(0.00995)
R-square	0.074	0.073	0.070	0.070	0.070	0.069
Mean [SD]	6.54 [1.18]		6.57 [1.12]		6.56 [1.14]	
Observations	31,290	31,290	34,818	34,818	39,143	39,143
[Panel C] Monthly household consumption expenditure (MHCE)						
RD Estimate	0.0428	0.0487	0.0803**	0.0776**	0.0765*	0.0761*
S.E.	(0.0388)	(0.0438)	(0.0383)	(0.0390)	(0.0401)	(0.0405)
R-square	0.347	0.347	0.333	0.335	0.320	0.319
Mean [SD]	2424.60 [1655.20]		2418.58 [1414.54]		2457.43 [1711.43]	
Observations	6,602	6,602	7,357	7,357	8,249	8,249
Bandwidth	0.02		0.025		0.03	
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: The dependent variables are an individual's weekly wages (panel A), number of days worked (panel B), and monthly household consumption expenditure (MHCE; panel C). All dependent variables are log-transformed. Odd-numbered columns show RD estimates from linear specifications and even-numbered columns show quadratic specifications. Standard errors are clustered at the district level.



Note: This figure includes RD graphs using the various rounds of National Sample Surveys. All graphs with solid fitted lines correspond to post-RSVY outcomes. All graphs with dashed fitted lines correspond to pre-RSVY outcomes (placebo tests). Each scatter point represents the bin average of residualized values of the outcome variables.

Fig. 5. RD graphs for outcomes from National Sample Surveys. Note: This figure includes RD graphs using the various rounds of National Sample Surveys. All graphs with solid fitted lines correspond to post-RSVY outcomes. All graphs with dashed fitted lines correspond to pre-RSVY outcomes (placebo tests). Each scatter point represents the bin average of residualized values of the outcome variables.

Table 4
Labor force participation, unemployment, migration, agricultural and non-agricultural workers.

	Linear (1)	Quadratic (2)	Linear (3)	Quadratic (4)	Linear (5)	Quadratic (6)
Panel A: Labor force participation, unemployment, migration						
[A1] Labor force participation (%)						
RD Estimate	−0.0160	−0.00934	−0.00709	−0.00807	0.00423	0.00461
S.E.	(0.0210)	(0.0206)	(0.0191)	(0.0185)	(0.0175)	(0.0176)
Mean [SD]	0.45 [0.49]		0.45 [0.49]		0.45 [0.49]	
Observations	22,527	22,527	25,135	25,135	28,256	28,256
[A2] Unemployment (%)						
RD Estimate	0.00719	0.00590	0.00371	0.00425	0.00505	0.00527
S.E.	(0.00591)	(0.00668)	(0.00500)	(0.00495)	(0.00641)	(0.00641)
Mean [SD]	0.01 [0.11]		0.01 [0.11]		0.01	[0.11]
Observations	22,527	22,527	25,135	25,135	28,256	28,256
[A3] Internal (within-state) migration (%) (NSS 64)						
RD Estimate	−0.000176	−0.00582	0.000900	0.00274	−0.00909	−0.00914
S.E.	(0.0226)	(0.0218)	(0.0202)	(0.0197)	(0.0184)	(0.0183)
Mean [SD]	0.10 [0.29]		0.09 [0.29]		0.09 [0.29]	
Observations	102,337	102,337	115,870	115,870	126,881	126,881
[A4] External (across-state) migration (%) (NSS 64)						
RD Estimate	−0.00575	−0.00670	−0.00482	−0.00417	−0.00441	−0.00444
S.E.	(0.0158)	(0.0169)	(0.0144)	(0.0139)	(0.0134)	(0.0134)
Mean [SD]	0.03 [0.18]		0.03 [0.18]		0.03 [0.18]	
Observations	102,337	102,337	115,870	115,870	126,881	126,881
Panel B: Sectoral reallocation						
[B1] Agricultural workers (%)						
RD Estimate	−0.0259*	−0.0253*	−0.0303*	−0.0312**	−0.0442***	−0.0374**
S.E.	(0.0149)	(0.0147)	(0.0167)	(0.0157)	(0.0165)	(0.0161)
Mean [SD]	0.11 [0.32]		0.11 [0.32]		0.11 [0.32]	
Observations	22,527	22,527	25,135	25,135	28,256	28,256
[B2] Non-agricultural workers (%)						
RD Estimate	0.0246*	0.0239*	0.0269*	0.0273*	0.0240*	0.0242*
S.E.	(0.0137)	(0.0140)	(0.0140)	(0.0141)	(0.0124)	(0.0123)
Mean [SD]	0.09 [0.29]		0.10 [0.30]		0.10 [0.29]	
Observations	22,527	22,527	25,135	25,135	28,256	28,256
Bandwidth/z/	0.02		0.025		0.03	
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: The dependent variables are indicators for labor force participation (panel A1), unemployment (panel A2), being an internal (within-state) migrant (panel A3), being an external (across-state) migrant (panel A4), being an agricultural worker (panel B1), and being a non-agricultural worker (panel B2). Odd-numbered columns show estimates from a linear RD specification and even-numbered columns include a second-order polynomial specification. Standard errors are clustered at the district level.

culture. Taken together, we find a decline in the proportion of agricultural workers that is mostly offset by the increase in the proportion of non-agricultural workers, without a corresponding increase in the proportion of people in the labor force. This suggests that RSVY may have led to sectoral reallocation of workers rather than an increase in overall economic activity. These results on the movement of workers out of agriculture to non-agricultural work are similar to Asher and Novosad (2020),Blakeslee et al. (2021).

5.4. Mechanisms

5.4.1. Evidence from microenterprises

One concern with our analysis is that we lack precise information on the projects on which RSVY grants were spent. However, the stated intention for RSVY cash grants was to foster infrastructural improvements in the backward districts. Therefore, we expect that a main channel driving the effects on microenterprises is through direct improvements in the overall infrastructural environment in treated districts. Since the main beneficiaries of RSVY were microenterprises, we focus on them and check whether they benefited from infrastructure improvements. To answer this, we focus on two broad measures from the NSS manufacturing survey in Table 5. In Panel A, we focus on problems related to infrastructure, and in Panel B we focus on problems unrelated to infrastructure. For problems related to infrastructure, we look at two measures. First, we look at firms' reports of problems faced, if any, during the reference year with respect to power cuts. We interpret the responses to this question as a measure of improvements in electrification. Second, we look at whether firms had problems with access to raw materials. Access to raw materials could increase either because there is an increase in firms supplying these raw materials in the same location or because there is better connectivity to suppliers in other locations. We interpret the responses to this question as a measure of improved road connectivity. In Panel A1, we find that in RSVY districts, the probability of a firm experiencing problems related to a power cut decreased by 10.5–17.7%. In Panel A2, we find that there is a 8.8–14.1% decline in the probability that firms had problems accessing raw materials. These coefficients are statistically significant for bandwidths of 0.025 and 0.03. Furthermore, the TOT effects in Appendix Table A3 estimates are also statistically significant. In Panel B, we look at problems unrelated to infrastructure, and find no statistically significant differences across any bandwidths or specifications. This is important to check because it may have been the case that other forms of government assistance to firms improved post-RSVY. In such a case, firm outcomes may have improved regardless of improvements in infrastructure. Reassuringly, null results in Panel B provides credence to our hypothesis that the overall effects were being driven by infrastructural improvements and not through other channels. The results from Table 5, graphically shown in Figs. 5 (m) through (r), provide suggestive evidence that RSVY led to improvements in electrification and rural connectivity that directly impacted microenterprises. However, there were no changes in problems unrelated to infrastructure post-RSVY.

Table 5
Microenterprises – Evidence of impact channels.

	Linear (1)	Quadratic (2)	Linear (3)	Quadratic (4)	Linear (5)	Quadratic (6)
Panel A: Likelihood of encountering infrastructure-related problems (%)						
[A1] Problems related to power cut (electricity)						
RD Estimate	−0.177***	−0.155***	−0.127***	−0.128***	−0.105***	−0.105***
S.E.	(0.0416)	(0.0362)	(0.0371)	(0.0373)	(0.0269)	(0.0268)
R-square	0.191	0.198	0.174	0.179	0.167	0.167
Observations	6,758	6,758	7,614	7,614	8,615	8,615
Mean [SD]	0.19 [0.39]		0.2 [0.4]		0.17 [0.38]	
[A2] Problems related to accessing to materials (road connectivity)						
RD Estimate	−0.108	−0.0880	−0.115*	−0.116*	−0.141**	−0.136**
S.E.	(0.0695)	(0.0594)	(0.0657)	(0.0644)	(0.0679)	(0.0642)
R-square	0.195	0.199	0.178	0.179	0.147	0.150
Observations	6,758	6,758	7,614	7,614	8,615	8,615
Mean [SD]	0.14 [0.35]		0.15 [0.36]		0.15 [0.36]	
Panel B: Likelihood of encountering problems unrelated to infrastructure (%)						
RD Estimate	0.0127	0.0131	0.00853	0.00863	0.00724	0.00724
S.E.	(0.00838)	(0.00844)	(0.00880)	(0.00871)	(0.00813)	(0.00814)
R-square	0.031	0.031	0.029	0.029	0.029	0.029
Observations	6,758	6,758	7,614	7,614	8,615	8,615
Mean [SD]	0.60 [0.49]		0.62 [0.49]		0.59 [0.49]	
Bandwidth	0.02		0.025		0.03	
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: The dependent variables are binary indicators for whether the microenterprise reported experiencing power shortage during production (Panel A1), having no access to raw materials (Panel A2), or experiencing a problem unrelated to infrastructure (Panel B). Odd-numbered columns show estimates from a linear RD specification and even-numbered columns include a second-order polynomial specification. Firm-specific covariates include firm owner's education level and ownership status. Standard errors are clustered at the district level.

If RSVY-induced improvements in rural electrification and connectivity were indeed important channels for microenterprise growth, the effects on employment, firm creation, and infrastructure-related problems may have been differentially larger for firms in electricity- and road-intensive industries. It is possible, however, that firms in the most electricity- and road-intensive industries had already adjusted their production processes to the existing quality of infrastructure. Any subsequent improvements in infrastructure would therefore have only a small effect. We test for these channels in Table 6. We use a measure of electricity intensity from Abeberese (2017)[20] in a 3 digit-industry defined as the average kilowatt-hours of electricity consumed per rupee of output by firms in that industry. For road intensity, we use transportation (traveling and freight) expenses at the 3 digit-industry level. Next, we divide industries by their degree of electricity and road intensity into terciles. We check whether firms in the highest tercile of road and electricity intensity experienced higher employment and whether there was more firm creation in these industries due to RSVY. We also expect firms in the highest tercile of electricity intensity to experience a larger decline in problems related to power cuts[21] and firms in the highest tercile of road intensity to experience a larger decline in problems of having no access to raw materials compared to other industries.

In Table 6, we find evidence consistent with our expectations. In Panel A, we find that for firms in the highest tercile of electricity-intensive industries in RSVY districts, employment increased by 19.3% (Panel A1), the proportion of new firms increased by 11.5% (Panel A2), and problems related to a power cut declined by 21.7% (Panel A3). In Panel B, we find that firms in industries in the highest tercile of road-intensity saw an increase in employment and the proportion of new firms by 18.3% (Panel B1) and 9.7% (Panel B2) respectively, and a decline of 19.0% (Panel B3) with respect to not having access to raw materials. Across all three terciles and across electricity- and road-intensive industries, we find a consistent pattern. Looking across columns (terciles of electricity- and road-intensive industries), we find that the effects on both employment for firms and the proportion of firms established less than 3 years ago are the largest in the highest tercile, followed by the middle tercile, and finally the smallest for the lowest tercile, in response to RSVY. For infrastructure-related problems (problems with a power cut and problems related to accessing raw materials) we find a consistent decline across terciles. This provides further evidence that the effects of RSVY were the largest in electricity- and road-intensive industries and bolsters our claim that RSVY-funded improvements in rural electrification and connectivity were important channels in driving microenterprise growth.

Furthermore, if RSVY grants led to an improvement in roads and electricity in the treated districts, firms may also have reported increases in energy- and transportation-related expenditures. We test for this in Table A9. Consistent with our other results, we find increases in both energy- and transportation-related expenditures for firms in treated districts, although results are not statistically significant across all specifications.

We also look at the effect of RSVY on night-time light luminosity. The use of night-light as a proxy for economic and infrastructure activities has become popular among economists.[22] There is an overwhelming consensus that light intensity and economic activity are closely related. Min (2008) shows that there is a strong association between nightlight luminosity and public-goods provision, especially across low-income countries. Particularly in India, Baskaran et al. (2015) further show that night-time light emission is suitable as a proxy measure for public-service provisions such as electricity, and Burlig and Preonas (2016) use changes in night-time brightness as an indicator for electrification. Hence, night-time lights in our context can be interpreted as both cause (infras-

[20] We thank Ama Baafra Abeberese for sharing this data with us.

[21] Note that our outcome variable is the probability of experiencing a *problem* related to power cuts and not the likelihood of a power cut. A decline in the likelihood of power cuts for all firms would still differentially reduce the *problems* related to a power cut for firms that use electricity intensively.

[22] See Henderson et al. (2012),Alesina et al. (2016), Chen and Nordhaus (2011), Hodler and Raschky (2014), Klomp (2016),Shenoy (2018).

Table 6
Microenterprises – Heterogeneous effects in electricity and road-intensive industries.

	(1)	(2)	(3)
Panel A: Electricity-intensive industries			
[A1] Intensive margin: Microenterprise employment			
RD Estimate	0.193**	0.0565	0.0515
S.E.	(0.0745)	(0.0440)	(0.0956)
Observations	2,394	7,610	1,866
[A2] Extensive margin: Established less than three years ago (%)			
RD Estimate	0.115***	0.0604***	−0.00793
S.E.	(0.0253)	(0.0192)	(0.0298)
Observations	2,394	7,610	1,866
[A3] Channel: problem related to power cut (%)			
RD Estimate	−0.217***	−0.116**	−0.173
S.E.	(0.0754)	(0.0471)	(0.106)
Observations	2,389	7,575	1,858
Degree of electricity dependency (tercile)	>66th	33th to 66th	<33th
Panel B: Road-intensive industries			
[B1] Intensive margin: Microenterprise employment			
RD Estimate	0.183**	0.0376	−0.0473
S.E.	(0.0761)	(0.0562)	(0.0383)
Observations	2,602	2,845	2,167
[B2] Extensive margin: Established less than three years ago (%)			
RD Estimate	0.0973***	0.0352	0.0429
S.E.	(0.0266)	(0.0230)	(0.0359)
Observations	2,602	2,845	2,167
[B3] Channel: problem related to accessing raw materials (%)			
RD Estimate	−0.190**	0.0751	−0.0682
S.E.	(0.0855)	(0.0987)	(0.0866)
Observations	2,599	2,828	2,152
Degree of road dependency (tercile)	>66th	33th to 66th	<33th
State Fixed Effects	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes

Note: Each column reports estimates from a sub-sample of firms in industries with varying degrees of electricity (Panel A) and road (Panel B) dependency. The dependent variables are microenterprises' employment (Panels A1 and B1), an indicator of young firm (being established less than years ago; Panels A2 and B2), firm's likelihood of having a problem related to power cut (Panel A3) and firm's likelihood of having a problem related to accessing raw materials (Panel B3). Firm-specific covariates include firm owner's education level and ownership status. All estimates are reported under the representative bandwidth of $|z| \leq 0.025$. Standard errors are clustered at the district level.

tructure improvements in electricity) and effect (local economic growth). We estimate the RD coefficients for nightlights for each year between 1999 and 2008 separately. The effects of RSVY on nightlights is graphically shown in Appendix Fig. A1, with each point representing the coefficient on the RD estimate for the given year. Before RSVY, we do not find any statistically significant coefficient, although we do find a jump in 2002. We find positive and significant growth in nightlight density in treated districts almost immediately after the introduction of RSVY in 2004. The statistically significant impact lasted for four to five subsequent years and dissipated around 2008. This coincides with the period when RSVY was in effect. The reversal in trends after 2008 is most likely due the introduction of the Backward Region Grants Fund (BRGF), another program with grants for infrastructure, that followed RSVY after 2007 and increased coverage to more districts. BRGF followed an identical selection process as RSVY and essentially converted a majority of the control districts in our analysis into treated districts under the new policy.

Given that we do not have information on the exact locations within the district where RSVY grants were spent, in the next subsection we explore two plausible hypotheses: first, whether the grants led to larger effects in villages that already had paved roads and electricity prior to RSVY; and second, if the effects were larger in villages geographically proximate to the district headquarters.

5.4.2. Village-level heterogeneity

As mentioned earlier, RSVY was an infrastructure-enhancing program in which the grants were supposed to be spent based on District Perspective Plans that would identify critical gaps in infrastructure (including improvement and maintenance) within a district. For example, RSVY funds were meant to strengthen rural electricity transmission and distribution rather than building new grids in un-electrified villages. Similarly, the funds were to be used to build critical road links or improve roads already built rather than to build new roads in villages that had no roads. Hence, if RSVY funds were spent effectively, we would expect that village employment in microenterprises as well as number of microenterprises to be higher in villages that already had paved roads and electricity at baseline (2001) compared to villages that were not endowed with such infrastructure. We check for this in Table 7, and the results match our expectations. In Panel A, we look at village employment in microenterprises, in Panel B at the number of microenterprises in the village, and in Panel C at the average firm size in the village. The sample in columns 1 and 2 is restricted to villages with no paved roads or electricity at the baseline.[23] The sample in columns 3 and 4 includes villages that had paved roads and electricity at the baseline. We find that RSVY did not statistically significantly affect either village employment or the number of microenterprises in villages with no roads or electricity (columns 1 and 2 across Panels A and B).[24] However, RSVY increased village employment in microenterprises by 17.4–18.1% (Panel A) and the number of microenterprises by 18–18.4% (Panel B) in villages endowed with both paved roads and electricity.[25] However, in Panel C, we do not find any statistically significant difference for average firm size across any columns. This provides further credence that RSVY funds were used to fill critical infrastructure gaps (including maintenance and improvement) within the district, and this directly led to growth in employment in microenterprises and number of microenterprises.

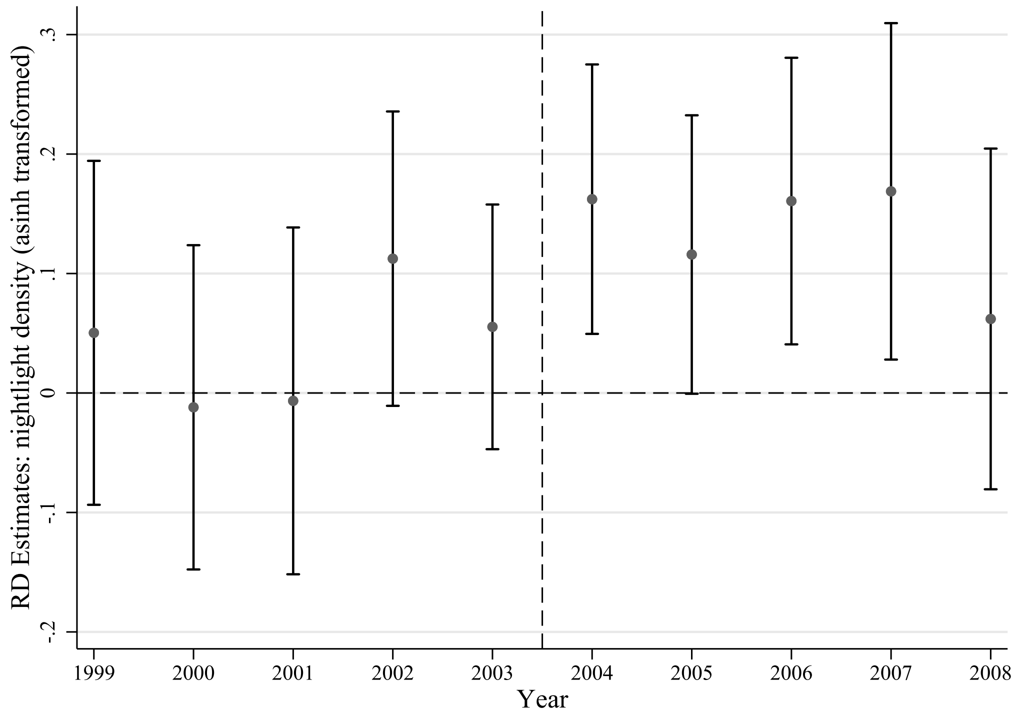
5.5. Robustness tests

To test the robustness of our findings, we conduct a sensitivity analysis across various bandwidths in the range [0.02,0.03]. The coefficients across the bandwidths for our outcomes of interest are shown in Fig. 6. The figure illustrates that our results are not sensitive to the choice of bandwidth. We also explicitly show the results using the Calonico et al. (2014) data-driven optimal bandwidth in Table A6. Having discussed the results in detail, we also perform two falsification tests to show that effect of RSVY becomes statistically indistinguishable from zero under counterfactual events. Specifically, we show that the policy had no effect in districts before RSVY was implemented or in districts that did not receive RSVY grants. First, we run the RD regressions with pre-RSVY data for our various outcomes of interest – the fourth round of the Economic Census (1998), NSS manufacturing survey round 56, schedule 2.2 (2000–2001), and NSS employment-unemployment round 55, schedule 10 (1999–2000). In Appendix Table A4, we show the results for this analysis. We find no statistically significant effects of the policy on all main outcomes before RSVY was implemented. In the second test, we replicate our regressions by adopting a hypothetical cutoff constructed identically to

[23] We use data from the Village Directory for the 2001 Population Census to define baseline infrastructure.

[24] These results are similar in spirit to Asher and Novosad (2020), who find no effects of an Indian rural roads program in villages that did not have paved roads, and Burlig and Preonas (2016), who find that a rural electrification scheme had no effect on unelectrified Indian villages.

[25] This is similar to results in Usmani (2019), who find heterogeneous treatment effects of a rural electrification scheme in India on villages that experienced an exogenous economic shock.



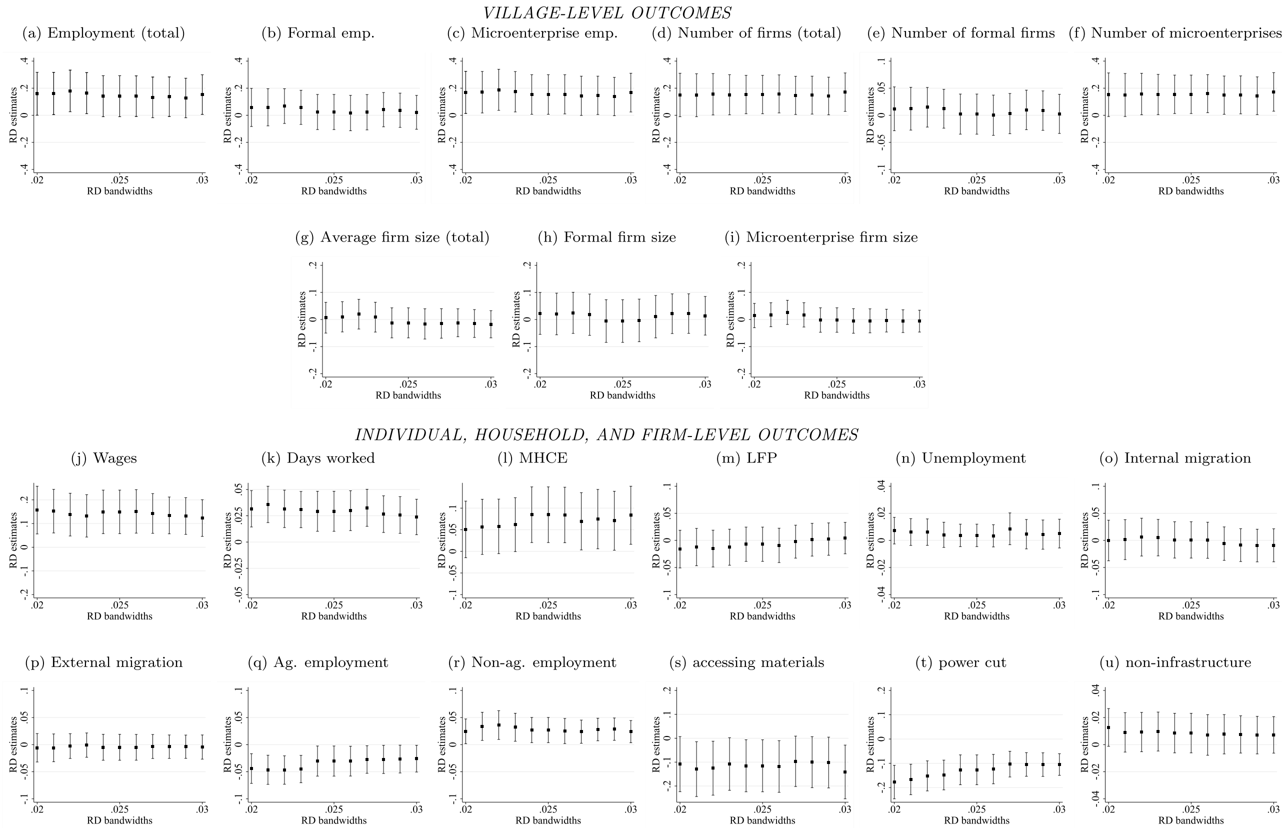
Note: The dependent variable is village-level night-time light intensity measured from satellite imagery. Underlying pixel-level luminosity value ranges from 0 to 63 (top-coded). Points represent yearly RD estimates. Whiskers represent 90-percent confidence intervals.

Fig. A1. Night-time light intensity. Note: The dependent variable is village-level night-time light intensity measured from satellite imagery. Underlying pixel-level luminosity value ranges from 0 to 63 (top-coded). Points represent yearly RD estimates. Whiskers represent 90-percent confidence intervals.

Table 7
Heterogeneity by baseline village infrastructure.

	Linear (1)	Quadratic (2)	Linear (3)	Quadratic (4)
Panel A: Village employment in microenterprises				
RD estimate	0.0663	0.0659	0.174*	0.181**
S.E.	(0.134)	(0.132)	(0.0893)	(0.0871)
R-squared	0.364	0.366	0.583	0.583
Observations	14,419	14,419	50,744	50,744
Sample	No roads or electricity	No roads or electricity	Roads and electricity	Roads and electricity
Panel B: Number of microenterprises in the village				
RD estimate	0.0632	0.0628	0.180**	0.184**
S.E.	(0.131)	(0.129)	(0.0869)	(0.0855)
R-squared	0.382	0.383	0.593	0.593
Observations	14,419	14,419	50,744	50,744
Sample	No roads or electricity	No roads or electricity	Roads and electricity	Roads and electricity
Panel C: Average firm size				
RD estimate	0.00274	0.00265	−0.00819	−0.00516
S.E.	(0.0454)	(0.0457)	(0.0268)	(0.0255)
R-squared	0.038	0.039	0.032	0.034
Observations	14,411	14,411	50,724	50,724
Sample	No roads or electricity	No roads or electricity	Roads and electricity	Roads and electricity
State Fixed Effects	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes

Note: The dependent variables are village-level microenterprises employment (Panel A), number of microenterprises (Panel B), and average employment per microenterprise (Panel C). Inverse hyperbolic sine transformation is applied to the outcome variables. Odd columns show estimates from first-order RD polynomial specifications and even columns from second-order polynomial specifications. Village covariates include baseline total population (log), share of irrigated area, paved road coverage, and electricity coverage. All estimates are reported under a restricted sample using a representative bandwidth of $|z| \leq 0.03$. Standard errors are clustered at the district level.



Note: Figures show the RD point estimates and confidence intervals of all main outcome variables from the Economic Census (2005) and National Sample Surveys (2005-06) under varying bandwidths.

Fig. 6. Sensitivity tests. Note: Figures show the RD point estimates and confidence intervals of all main outcome variables from the Economic Census (2005) and National Sample Surveys (2005–06) under varying bandwidths.

our baseline specifications, but after removing all the treated districts from the sample. Essentially, in this exercise we test whether RSVY grants had any effect on districts that did not actually receive the grants. Overall, in Appendix Table A5, we find no statistically significant effects at these hypothetical cutoffs.[26]

As an additional robustness test, we estimate a difference-in-discontinuities design for the main village-level results. Our empirical analysis follows (Grembi et al., 2016):

$$y_{idst} = \beta_0 + \beta_1[\mathbf{1}\{z_{ds} \leq 0\} \times year_{2005}] + \beta_2[\delta(z_{ds}) \times year_{2005}] + \gamma_{2005} + \pi_i + \varepsilon_{idst}. \quad (3)$$

All subscripts remain the same as in Eq. 2. The new terms in Eq. 3 include the post-treatment dummy $year_{2005}$, which is interacted with the predicted treatment status $\mathbf{1}\{z_{ds} \leq 0\}$ and functional forms of the running variable $\delta(z_{ds})$. The main coefficient of interest, β_1, provides the difference-in-discontinuities estimate of the effects of RSVY. The village-panel structure allows us to control for both time fixed effects γ_{2005} and village fixed effects π_i. Village fixed effects essentially absorb all district-specific and village-specific predetermined covariates that are included in the main RDD regression. Both the 1998 (pre-RSVY) and 2005 (post-RSVY) rounds of the Economic Census are included. The regression sample for the largest RD bandwidth ($z \leqslant 0.03$) consists of 183,340 village-level observations that appear in both EC rounds and contain a complete set of regression covariates.[27] The results from this estimation are shown in Table A7. The results can be interpreted as the changes in village outcomes after RSVY in treated districts compared to control districts (around the cutoff), relative to the pre-RSVY period. Similar to our main results, we find that overall village employment and number of firms increases after RSVY, mainly driven by increases in village employment in informal firms. The results are not statistically significant for all bandwidths. However, we again find no changes in village employment or number of firms for formal firms. We also find no changes in average firm size for all firms, formal firms, or informal firms. These tests provide credibility to our claim that the main effects are indeed caused by the RSVY grants.

6. Conclusion

Infrastructure investments have often been misallocated in developing countries, with too much spent on building new infrastructure and not enough on maintenance and improvements. Inadequate maintenance has been a central reason for the failure of infrastructure in developing countries (World Bank, 1994). In this

[26] Note that in these regressions, after removing our original treated districts, we are essentially comparing outcomes in districts that are not similar in terms of observables. In fact, districts to the right of the hypothetical cutoff are better off than districts on the left of the cutoff. Hence, as expected, for most outcomes we find a negative coefficient.

[27] Out of the 92,677 villages in EC 2005 that have non-missing village-level covariates used in the main RD regressions in Table 2, we are able to exactly match 91,670 with data from EC 1998 when constructing the village-level panel to estimate the difference-in-discontinuities regression. The mismatch (approximately 1% of sample) is the result of village observations appearing in one round and not the other, potentially due to administrative merging and/or splitting of villages during the interim period.

paper we study the effects of a rural infrastructure grants program, Rashtriya Sam Vikas Yojana (RSVY), in India's backward districts, where spending was mandated on improvements and maintenance of existing infrastructure. We exploit the transparent treatment-assignment mechanism of RSVY that allows us to reconstruct state-specific score cutoffs. We estimate the effects of the policy using a regression discontinuity design across the entire firm-size distribution. We find that RSVY led to increases in village employment, entirely driven by employment in microenterprises. We find corresponding increases in the number of firms in villages, again driven by increases in the number of microenterprises. Average firm size remained unchanged. There were no changes for formal firms or on the proportion of people in the labor force, unemployed, or who were in-migrants. However, there was a reduction in the share of agricultural wage workers and a corresponding increase in the share of non-agricultural workers, suggesting a movement out of agriculture into microenterprises.

One limitation of our analysis is that we lack information on the actual projects on which the infrastructure grants were spent in the district. Therefore, we remain cautious in overgeneralizing across contexts. Our empirical results, however, shed light on potential mechanisms underlying the effects on firm outcomes. We show that RSVY cash grants directly led to firms reporting significantly lower likelihoods of encountering problems related to power cuts and access to raw materials. Furthermore, our microenterprise-level results are stronger for firms in electricity- and road-intensive industries. We also find that employment in microenterprises and the number of enterprises are higher in villages that had roads and electricity prior to RSVY. Overall, the results provide suggestive evidence that both RSVY-funded improvements in rural electrification and connectivity led to microenterprise growth.

A caveat of our analysis is that we are only able to look at the short-term effects of infrastructure grants. Studying the long-run effects of infrastructure investments is critical, especially for developing countries with large infrastructure gaps. Firms and households could gain from public investments in *new* roads, electrification, dams, better telecommunication (Internet, mobile telephone networks), and maintenance and improvements of already existing infrastructure. Maintenance and improvement may have differing distributional effects compared to new construction. For example, investments in maintenance may disproportionately benefit smaller firms and vulnerable households who are unable to adapt to infrastructure disruptions. New infrastructure provision on the other hand may have more equitable impacts across the distribution of firms and households. Among the plethora of options available to policymakers, which ones should be prioritized is a very important question. These are promising avenues for future research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

Tables A1–A9

Table A1
Balance in district characteristics around the threshold.

	Linear			Quadratic		
	RD Estimate	S.E.	R-Square	RD Estimate	S.E.	R-Square
Panel A: Village-level outcomes: Economic Census 1998						
1. Employment						
Total Employment	0.308	(0.226)	0.671	0.296	(0.228)	0.672
Formal Employment	0.226	(0.387)	0.500	0.198	(0.394)	0.503
Informal Employment	0.341	(0.220)	0.679	0.330	(0.221)	0.680
2. Number of firms						
All Firms	0.325	(0.222)	0.658	0.322	(0.223)	0.658
Formal Firms	0.297	(0.321)	0.550	0.264	(0.323)	0.555
Informal Firms	0.326	(0.222)	0.658	0.324	(0.223)	0.658
3. Average firm size						
All Firms	−0.016	(0.049)	0.298	−0.025	(0.050)	0.313
Formal Firms	−0.070	(0.112)	0.213	−0.066	(0.113)	0.214
Informal Firms	0.013	(0.027)	0.312	0.005	(0.030)	0.346
Panel B: Micro and formal enterprises (NSS; ASI)						
1. Microenterprises; NSS 2000–01						
Employment (log)	0.039	(0.054)	0.460	0.039	(0.053)	0.460
Problem Related to Power Cut (%)	−0.030	(0.053)	0.293	−0.029	(0.053)	0.295
Problem Related to Accessing Materials (%)	−0.036	(0.062)	0.395	−0.033	(0.061)	0.401
Problem Unrelated-to-infrastructure (%)	0.019	(0.067)	0.497	0.022	(0.066)	0.502
Panel C: Wage, days worked, MHCE, LFP, unemployment, migration, and sectoral employment (NSS 1999–2000)						
Wages (Rs./week; log)	0.037	(0.168)	0.142	0.035	(0.169)	0.142
Days worked (days/week; log)	0.011	(0.012)	0.493	0.010	(0.011)	0.507
MHCE (Rs./month; log)	0.011	(0.051)	0.710	0.014	(0.050)	0.723
Labor force participation (LFP) (%)	0.024	(0.015)	0.417	0.024	0.015	0.422
Unemployed (%)	−0.001	(0.004)	0.461	−0.002	(0.004)	0.461
Agriculture workers (%)	−0.016	(0.017)	0.695	−0.016	(0.017)	0.696
Non-agriculture workers (%)	0.003	(0.006)	0.569	0.004	(0.006)	0.580
Panel D: Socio-economic, infrastructure, and geographic characteristics						
1. Under-development (backwardness) indicators						
Share of SC/ST population (% 1991)	−1.679	(1.118)	0.939	−1.567	(1.131)	0.939
Output per Agricultural Worker (Rs. 1990–93)	−271.8	(491.8)	0.905	−268.3	(486.3)	0.905
Agricultural Wage Rate (Rs. 1996–97)	−1.506	(1.196)	0.883	−1.386	(1.213)	0.884

Table A1 (*continued*)

	Linear			Quadratic		
	RD Estimate	S.E.	R-Square	RD Estimate	S.E.	R-Square
2. Baseline population and infrastructure indicators (Population Census 2001)						
Population (log)	0.271	(0.181)	0.525	0.252	(0.179)	0.531
Paved Roads (%)	0.015	(0.045)	0.829	0.022	(0.044)	0.833
Electricity Coverage (%)	−0.041	(0.043)	0.876	−0.039	(0.043)	0.876
Irrigation Area (%)	0.004	(0.059)	0.699	0.008	(0.058)	0.700
3. Geographic characteristics						
Distance to nearest city (km)	9.381	(14.15)	0.411	5.147	(14.38)	0.444
Area (km sq.)	982.0	(747.9)	0.775	934.3	(781.7)	0.776
Elevation (m)	−15.52	(41.50)	0.744	−15.63	(41.30)	0.744

Note: This table presents district-level RD regressions with the main outcome variables pre-RSVY used in our analysis (Panels A to C), and other time-invariant socio-economic, infrastructure, and geographic characteristics (Panel D). Estimates are obtained with linear and quadratic RD polynomial specifications in the representative bandwidth of $|z| \leq 0.025$.

Table A2
Balance on different village population cutoffs (sample $|z| \leq 0.03$).

	Treatment districts	Control districts	Treatment–control	p-value on difference
Number of villages				
All	877.25	829.30	47.95	0.67
Population > 300	800.35	740.73	59.62	0.52
300 < Population < 450	68.25	77.08	−8.83	0.63
Population > 500	708.12	636.49	71.62	0.35
500 < Population < 750	114.92	120.57	−5.65	0.78
Population > 1000	493.88	420.35	73.53	0.15
Share of villages				
Population > 300	0.92	0.92	−0.00	0.95
300 < Population < 450	0.070	0.075	−0.005	0.55
Population > 500	0.83	0.82	0.01	0.81
500 < Population < 750	0.12	0.13	−0.01	0.56
Population > 1000	0.59	0.59	0.00	0.88
Number of districts	52	63		

Note: This table presents the means across treatment and control districts for the number and share of villages by population cutoffs important for the eligibility criteria of PMGSY and RGGVY.

Table A3
Estimates for Treatment-on-the-Treated (TOT).

	Linear		Quadratic		
Dependent variable	RD Estimate	S.E.	RD Estimate	S.E.	Observations
Panel A: Village employment, number of firms, average firm size (Economic Census 2005)					
[A1] Village employment					
Total Employment	0.326	(0.213)	0.344*	(0.211)	83,356
Formal Employment	0.0463	(0.161)	0.0497	(0.161)	83,356
Informal Employment	0.357*	(0.214)	0.377*	(0.212)	83,356
[A2] Number of firms					
All Firms	0.364*	(0.213)	0.377*	(0.213)	83,356
Formal Firms	0.00544	(0.0464)	0.00607	(0.0464)	83,356
Informal Firms	0.368*	(0.214)	0.382*	(0.215)	83,356
[A3] Average firm size					
All Firms	−0.0327	(0.0882)	−0.0199	(0.0795)	83,356
Formal Firms	−0.0118	(0.0946)	−0.0128	(0.0944)	7,627
Informal Firms	−0.00491	(0.0711)	0.00562	(0.0640)	83,313
Panel B: Wage, work hours, consumption, migration, sectoral employment (NSS Sch. 10; 2005–06)					
Wage	0.407*	(0.236)	0.400*	(0.229)	5,424
Days worked (in last 7 days)	0.0765*	(0.0431)	0.0746*	(0.0417)	35,695
MHCE	0.185*	(0.110)	0.185	(0.114)	7,680
Labor force participation (%)	−0.0598	(0.0630)	0.0114	(0.0180)	33,862
Unemployment (%)	0.00300	(0.00473)	0.00144	(0.00406)	33,862
Internal in-migration (%)	0.00241	(0.0544)	−0.0227	(0.0170)	115,870
External in-migration (%)	−0.0129	(0.0367)	−0.0110	(0.0141)	115,870
Agricultural Workers (%)	−0.0636*	(0.0384)	−0.00211	(0.0176)	33,862
Non-agricultural workers (%)	0.0411	(0.0258)	0.00148	(0.0105)	33,862
Panel C: Microenterprises' impact channels (NSS Sch. 2.2; 2005–06)					
Power Cut (%)	−0.281*	(0.157)	−0.292*	(0.163)	7,614
Accessing Materials (%)	−0.378**	(0.170)	−0.377**	(0.180)	7,614
Unrelated-to-infrastructure (%)	0.0303	(0.0301)	0.0302	(0.0329)	7,614

Note: This table presents Treatment-on-the-Treated effects on all outcome variables presented previously in the analysis. Estimates are reported for the representative bandwidth of $|z| \leq 0.025$. All else remains unchanged from the main exercises.

Table A4
Pre-treatment impacts (placebo test 1).

	Linear		Quadratic		
Dependent variable	RD Estimate	S.E.	RD Estimate	S.E.	Observations
Panel A: Village employment, number of firms, avereage firm size (Economic Census 1998)					
[A1] Village employment					
Total Employment	0.101	(0.0970)	0.112	(0.0958)	83,695
Formal Employment	0.0363	(0.0875)	0.0285	(0.0942)	83,695
Informal Employment	0.111	(0.0939)	0.123	(0.0919)	83,695
[A2] Number of firms					
All Firms	0.0983	(0.0851)	0.116	(0.0804)	83,695
Formal Firms	0.0126	(0.0253)	0.0102	(0.0271)	83,695
Informal Firms	0.100	(0.0855)	0.118	(0.0809)	83,695
[A3] Average firm size					
All Firms	−0.000268	(0.0427)	−0.00616	(0.0460)	83,695
Formal Firms	−0.0673	(0.0521)	−0.0669	(0.0530)	7,546
Informal Firms	0.00799	(0.0336)	0.00313	(0.0365)	83,591
Panel B: Wage, work hours, comsumption, migration, sectoral employment (NSS Sch. 10; 2000–01)					
Wage	0.0473	(0.0524)	0.0461	(0.0521)	15,838
Days worked (in last 7 days)	0.0435	(0.0300)	0.0432	(0.0299)	38,366
MHCE	−0.000755	(0.0556)	−0.00595	(0.0547)	6,286
Labor force participation (%)	−0.00151	(0.0136)	−0.00180	(0.0130)	99,831
Unemployment (%)	−0.00107	(0.00239)	−0.00104	(0.00230)	99,831
Internal in-migration (%)	0.0137	(0.0166)	0.0141	(0.0165)	134,769
External in-migration (%)	−0.00907	(0.0123)	−0.00889	(0.0121)	134,769
Agricultural Workers (%)	−0.00776	(0.0155)	−0.00812	(0.0139)	99,831
Non-agricultural workers (%)	0.0115	(0.00718)	0.0115	(0.00716)	99,831
Panel C: Microenterprises' impact channels (NSS Sch. 2.2; 2000–01)					
Power Cut (%)	−0.0440	(0.0443)	−0.0447	(0.0438)	17,839
Accessing Materials (%)	−0.0892	(0.0547)	−0.0894	(0.0545)	17,839
Unrelated-to-infrastructure (%)	−0.0477	(0.0536)	−0.0478	(0.0523)	17,839

Note: This table shows placebo effects using pre-RSVY data sets. Estimates are reported for the representative bandwidth of $|z| \leq 0.025$. All else remains unchanged from the main exercises.

Table A5
Hypothetical eligibility threshold (placebo test 2).

	Linear		Quadratic		
Dependent Variable	RD Estimate	S.E.	RD Estimate	S.E.	Observations
Panel A: Village employment, number of firms, avereage firm size (Economic Census 2005)					
[A1] Village employment					
Total Employment	−0.236	(0.170)	−0.235	(0.168)	84,665
Formal Employment	0.115	(0.0921)	0.114	(0.0917)	84,665
Informal Employment	−0.243	(0.168)	−0.242	(0.167)	84,665
[A2] Number of firms					
All Firms	−0.191	(0.145)	−0.190	(0.144)	84,665
Formal Firms	0.0353	(0.0270)	0.0351	(0.0270)	84,665
Informal Firms	−0.191	(0.146)	−0.190	(0.145)	84,665
[A3] Average firm size					
All Firms	−0.0393	(0.0364)	−0.0393	(0.0361)	84,665
Formal Firms	−0.0269	(0.102)	−0.0242	(0.0995)	7,851
Informal Firms	−0.0463	(0.0340)	−0.0460	(0.0334)	84,647
Panel B: Wage, work hours, comsumption, migration, sectoral employment (NSS Sch. 10; 2005–06)					
Wage	0.0502	(0.0948)	0.0636	(0.106)	1,850
Days worked (in last 7 days)	−0.0382	(0.0239)	−0.0370*	(0.0208)	9,334
MHCE	−0.0293	(0.0712)	−0.0816	(0.0602)	2,108
Labor force participation (%)	−0.00642	(0.0276)	−0.0106	(0.0314)	9,798
Unemployment (%)	0.00802*	(0.00476)	0.00746	(0.00554)	13,157
Internal in-migration (%)	0.000900	(0.0202)	0.00274	(0.0197)	115,870
External in-migration (%)	−0.00482	(0.0144)	−0.00417	(0.0139)	115,870
Agricultural Workers (%)	0.0547	(0.0403)	0.0425	(0.0447)	13,157
Non-agricultural workers (%)	−0.000175	(0.0204)	−0.00429	(0.0201)	13,157
Panel C: Microenterprises' impact channels (NSS Sch. 2.2; 2005–06)					
Power Cut (%)	−0.00224	(0.133)	0.0880	(0.123)	2,854
Accessing Materials (%)	0.238*	(0.125)	0.237	(0.145)	2,854
Unrelated-to-infrastructure (%)	−0.0867	(0.0616)	−0.0733	(0.0635)	2,854

Note: Placebo effects using a hypothetical threshold based on the backwardness index scores (e.g., cutoffs unrelated to RSVY). Estimates are reported for the representative bandwidth of $|z| \leq 0.025$. All else remains unchanged from previous exercises.

Table A6
Robustness with data-driven Calonico et al. (2014) bandwidth.

	Linear		Quadratic		
Dependent variable	RD Estimate	S.E.	RD Estimate	S.E.	Observations
Panel A: Village employment, number of firms, average firm size (Economic Census 2005)					
[A1] Village employment					
Total Employment	0.161*	(0.0938)	0.139	(0.0918)	74,459
Formal Employment	0.0530	(0.0872)	−0.0005	(0.0881)	72,303
Informal Employment	0.170*	(0.0921)	0.151*	(0.0894)	74,459
[A2] Number of firms					
All Firms	0.142*	(0.0838)	0.152*	(0.0844)	89,978
Formal Firms	−0.00437	(0.0315)	−0.0101	(0.0331)	59,395
Informal Firms	0.144*	(0.0840)	0.153*	(0.0846)	89,978
[A3] Average firm size					
All Firms	−0.0206	(0.0281)	−0.0199	(0.0284)	104,046
Formal Firms	0.0218	(0.0440)	0.0253	(0.0428)	8,494
Informal Firms	−0.0165	(0.0228)	−0.0174	(0.0224)	110,830
Panel B: Wage, work hours, comsumption, migration, sectoral employment (NSS Sch. 10; 2005–06)					
Wage	0.117*	(0.0633)	0.0701	(0.0712)	3,831
Days worked (in last 7 days)	0.0230*	(0.0119)	0.0186	(0.0140)	21,189
MHCE	0.158*	(0.0904)	0.127	(0.0967)	3,714
Labor Force Participation (%)	0.0061	(0.0241)	0.0062	(0.0238)	19,398
Unemployment (%)	0.00103	(0.00752)	−0.00129	(0.00795)	18,302
Internal in-migration (%)	−0.00825	(0.0180)	−0.00777	(0.0179)	130,170
External in-migration (%)	−0.00591	(0.0155)	−0.00666	(0.0161)	104,609
Agricultural Workers (%)	−0.0666***	(0.0178)	−0.0609***	(0.0206)	19,726
Non-agricultural workers (%)	0.0366**	(0.0172)	0.0366*	(0.0193)	14,663
Panel C: Microenterprises' impact channels (NSS Sch. 2.2; 2005–06)					
Power Cut (NSS; %)	−0.152***	(0.0373)	−0.142***	(0.0329)	7,144
Accessing Materials (NSS; %)	−0.141**	(0.0679)	−0.136**	(0.0642)	8,615
Unrelated-to-infrastructure (NSS; %)	0.0174	(0.0112)	0.0174	(0.0111)	7,614

Note: This table presents RSVY estimates on all outcome variables presented previously in the analysis under the data-driven optimal bandwidth selected according to the Calonico et al. (2014) criteria. All else remains unchanged from the main exercises.

Table A7
Difference-in-discontinuities estimation.

		Linear	Quadratic	Linear	Quadratic	Linear	Quadratic
		(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Village employment							
1. Total Employment	RSVY X Post	0.207**	0.205**	0.135*	0.138*	0.106	0.105
	S.E.	(0.0848)	(0.0818)	(0.0748)	(0.0778)	(0.0690)	(0.0719)
	Observations	144,834	144,834	164,864	164,864	183,340	183,340
2. Formal Employment	RSVY X Post	0.0457	0.0294	−0.0442	−0.0408	−0.0437	−0.0431
	S.E.	(0.0790)	(0.0724)	(0.0682)	(0.0699)	(0.0717)	(0.0701)
	Observations	144,834	144,834	164,864	164,864	183,340	183,340
3. Informal Employment	RSVY X Post	0.206**	0.205**	0.140*	0.143*	0.111	0.109
	S.E.	(0.0848)	(0.0830)	(0.0762)	(0.0789)	(0.0704)	(0.0735)
	Observations	144,834	144,834	164,864	164,864	183,340	183,340
Panel B: Number of firms							
1. All Firms	RSVY X Post	0.162*	0.189**	0.127	0.127	0.112	0.111
	S.E.	(0.0840)	(0.0821)	(0.0783)	(0.0770)	(0.0739)	(0.0751)
	Observations	144,834	144,834	164,864	164,864	183,340	183,340
2. Formal Firms	RSVY X Post	0.0136	0.00701	−0.0114	−0.0104	−0.0115	−0.0114
	S.E.	(0.0241)	(0.0199)	(0.0193)	(0.0201)	(0.0192)	(0.0188)
	Observations	144,834	144,834	164,864	164,864	183,340	183,340
3. Informal Firms	RSVY X Post	0.162*	0.189**	0.127	0.127	0.112	0.112
	S.E.	(0.0847)	(0.0828)	(0.0790)	(0.0777)	(0.0745)	(0.0758)
	Observations	144,834	144,834	164,864	164,864	183,340	183,340
Panel C: Average firm size (number of employees)							
1. All Firms	RSVY X Post	0.0365	0.0116	0.00367	0.00657	−0.00837	−0.00899
	S.E.	(0.0459)	(0.0355)	(0.0382)	(0.0358)	(0.0331)	(0.0332)
	Observations	144,834	144,834	164,864	164,864	183,340	183,340
2. Formal Firms	RSVY X Post	0.0240	0.0963	0.0305	0.0432	0.0262	0.0318
	S.E.	(0.103)	(0.0967)	(0.0960)	(0.0915)	(0.0886)	(0.0870)
	Observations	5,634	5,634	5,990	5,990	6,910	6,910
3. Informal Firms	RSVY X Post	0.0359	0.0115	0.00792	0.0107	−0.00398	−0.00481
	S.E.	(0.0408)	(0.0304)	(0.0338)	(0.0310)	(0.0290)	(0.0288)
	Observations	144,616	144,616	164,592	164,592	183,056	183,056
Bandwidth		0.02		0.025		0.03	
Village & Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the estimated coefficients using a difference-in-discontinuities approach (Eq. 3) as a robustness exercise corresponding to the main village-level results presented in Table 2. The outcome variables include average employment (Panel A), average number of firms (Panel B), and average firm size (number of workers per firm; Panel C). Odd columns show estimates from first-order RD polynomial specifications and even columns from second-order polynomial specifications. All regressions control for year and village fixed effects. Standard errors are clustered at the district level.

Table A8
Results for transportation and energy expenditures (NSS 62 - Sch. 10).

	Linear (1)	Quadratic (2)	Linear (3)	Quadratic (4)	Linear (5)	Quadratic (6)
Panel A: Transportation expenditure (NSS 62 Schedule 2.2)						
RD Estimate	0.295	0.346	0.422	0.427	0.469*	0.470*
S.E.	(0.323)	(0.331)	(0.308)	(0.307)	(0.280)	(0.281)
R-square	0.158	0.158	0.161	0.161	0.167	0.167
Observations	6,758	6,758	7,614	7,614	8,615	8,615
Panel B: Energy expenditure (NSS 62 Schedule 2.2)						
RD Estimate	0.652*	0.499*	0.556*	0.576*	0.421	0.454
S.E.	(0.348)	(0.271)	(0.312)	(0.316)	(0.383)	(0.337)
R-square	0.337	0.340	0.311	0.315	0.302	0.303
Observations	6,758	6,758	7,614	7,614	8,615	8,615
Bandwidth	0.02		0.025		0.03	
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: Estimates are reported using a linear RD polynomial specification and a restricted sample with a representative bandwidth of $|z| \leq 0.025$. The outcome variables include village-level number of firms (first three columns), village employment (middle three columns), and average employment per firm (last three columns) by employment sector. All else remains the same as in the main tables.

Table A9
Results for transportation and energy expenditures (NSS 62 - Sch. 10).

	Linear (1)	Quadratic (2)	Linear (3)	Quadratic (4)	Linear (5)	Quadratic (6)
Panel A: Transportation expenditure (NSS 62 Schedule 2.2)						
RD Estimate	0.295	0.346	0.422	0.427	0.469*	0.470*
S.E.	(0.323)	(0.331)	(0.308)	(0.307)	(0.280)	(0.281)
R-square	0.158	0.158	0.161	0.161	0.167	0.167
Observations	6,758	6,758	7,614	7,614	8,615	8,615
Panel B: Energy expenditure (NSS 62 Schedule 2.2)						
RD Estimate	0.652*	0.499*	0.556*	0.576*	0.421	0.454
S.E.	(0.348)	(0.271)	(0.312)	(0.316)	(0.383)	(0.337)
R-square	0.337	0.340	0.311	0.315	0.302	0.303
Observations	6,758	6,758	7,614	7,614	8,615	8,615
Bandwidth	0.02		0.025		0.03	
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table shows RD estimation results with transportation expenditure (panel A) and energy expenditure (panel B) as the outcome variables (NSS 62 - Sch. 10). All specifications otherwise follow the main regression.

Appendix B. Detailed construction of the backwardness index

The backwardness index is constructed by equally weighting three historical parameters: (i) value of output per agricultural worker (1990–1993); (ii) agriculture wage rate (1996–1997); and (iii) districts' percentage of low-caste populations - Scheduled Castes/ Scheduled Tribes (1991 Census). This backwardness index ranks a total of 447 districts in 17 major states with available data for the parameters above. Data on agricultural productivity per worker was available for only 17 states. As a result, the state of Goa and all special category states (with the exception of Assam) were excluded from the analysis. There is available information for 482 districts in these 17 states. In addition, the Task Force Department further decided to exclude districts with urban agglomerations of over one million population as per the 2001 census. State capitals were also excluded. The reason for these exclusions is that urban centers would almost always generate economic activates that would obviate the need for public works programs. Consequently, 35 additional districts were further excluded from the analysis. Therefore, the backwardness index can be calculated for the remaining 447 districts across 17 states.

It should also be noted that in most states the number of districts has increased since 1991 due to division of old districts. In those cases, the Scheduled Caste and Scheduled Tribe (SC/ST) population proportion for the original district in 1991 would be applied to the new districts created by the division of the district. This imputation process is done similarly for agricultural wages and agricultural productivity per worker.

B.1. Ranking Computation

The index was computed for each variable. For agricultural productivity per worker and agricultural wages, the index was computed as follows:

$$\frac{(ActualValue - MinimumValue)}{(MaximumValue - MinimumValue)}$$

The lower the index value, the more backward the district. In the case of the parameter for SC/ST population, it is assumed *a priori* that districts with higher proportion of SC/ST population are more backward. To ensure that the index values in the three variables moved in the same direction, the index for SC/ST population was calculated as follows:

$$\frac{(MaximumValue - ActualValue)}{(MaximumValue - MinimumValue)}$$

The districts with higher percentage of SC/ST population would have a lower value for the index.

Next, the three sub-indices were aggregated with equal weights, resulting in a composite index. The Planning Commission used the composite index as the final product to rank districts on their level of backwardness. The districts with low wages, low productivity, and high SC/ST population were ranked as more backward on the index, (i.e. getting a lower rank value). The discrete ranking, thus, ranges from 1 for the most backward district, to 447 for the least backward, subject to data availability.

References

Abeberese, A.B., 2017. Electricity Cost and Firm Performance: Evidence from India. Rev. Econ. Stat. 99, 839–852.
Adukia, A., Asher, S., Novosad, P., 2020. Educational Investment Responses to Economic Opportunity: Evidence from Indian Road Construction. Am. Econ. J.: Appl. Econ. 12, 348–376.
Aggarwal, S., 2018. Do Rural Roads Create Pathways out of Poverty? Evidence from India. J. Dev. Econ. 133, 375–395.
Aker, J.C., Mbiti, I.M., 2010. Mobile Phones and Economic Development in Africa. Journal of Economic Perspectives 24, 207–232.
Aklin, M., Bayer, P., Harish, S., Urpelainen, J., 2017. Does Basic Energy Access Generate Socioeconomic Benefits? A Field Experiment with Off-grid Solar Power in India. Science Advances 3, e1602153.
Alder, S., Shao, L., Zilibotti, F., 2016. Economic Reforms and Industrial Policy in a Panel of Chinese Cities. J. Econ. Growth 21, 305–349.
Alesina, A., Michalopoulos, S., Papaioannou, E., 2016. Ethnic Inequality. Journal of Political Economy 124, 428–488.
Allcott, H., Collard-Wexler, A., O'Connell, S., 2016. How Do Electricity Shortages Affect Industry? Evidence from India, American Economic Review 106, 587–624.
Alsan, M., Goldin, C., 2019. Watersheds in Child Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880–1920. Journal of Political Economy 127, 586–638.
Asher, S., Garg, T., Novosad, P., 2020. The Ecological Impact of Transportation Infrastructure. Econ. J. 130, 1173–1199.
Asher, S., Novosad, P., 2020. Rural Roads and Local Economic Development. American Economic Review 110, 797–823.
Asturias, J., García-Santana, M., Ramos, R., 2018. Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India. Journal of the European Economic Association 17, 1881–1940.
Baskaran, T., Min, B., Uppal, Y., 2015. Election Cycles and Electricity Provision: Evidence from a Quasi-Experiment with Indian Special Elections. Journal of Public Economics 126, 64–73.
Bhargava, A.K. (2014): "The Impact of India's Rural Employment Guarantee on Demand for Agricultural Technology, IFPRI Discussion Paper 01381.
Blakeslee, D., Chaurey, R., Fishman, R., Malik, S., 2021. Land Rezoning and Structural Transformation in Rural India: Evidence from the Industrial Areas Program. The World Bank Economic Review 36, 488–513.
Brooks, W. and K. Donovan (2019): "Eliminating Uncertainty in Market Access: The Impact of New Bridges in Rural Nicaragua*. Working Paper.
Burlig, F. and L. Preonas (2016): "Out of the Darkness and Into the Light? Development Effects of Rural Electrification, Energy Institute at Haas Working Paper.
Calonico, S., Cattaneo, M.D., Titiunik, R., 2014. Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. Econometrica 82, 2295–2326.
Chakravorty, U., Pellib, M., Marchand, B.U., 2014. Does the Quality of Electricity Matter? Evidence from Rural India. Journal of Economic Behavior & Organization 107, 228–247.
Chaurey, R., 2017. Location-Based Tax Incentives: Evidence from India. Journal of Public Economics 156, 101–120.
Chen, X., Nordhaus, W.D., 2011. Using Luminosity Data as a Proxy for Economic Statistics. Proc. Nat. Acad. Sci. 108, 8589–8594.
Cheng, Y. (2014): "Place-based Policies in a Development Context - Evidence from China, Working Paper, UC Berkeley.
Devoto, F., Duflo, E., Dupas, P., Parienté, W., Pons, V., 2012. Happiness on Tap: Piped Water Adoption in Urban Morocco. American Economic Journal: Economic Policy 4, 68–99.
Dinkelman, T., 2011. The Effects of Rural Electrification on Employment: New Evidence from South Africa. American Economic Review 101, 3078–3108.
Donaldson, D., 2018. Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. American Economic Review 108, 899–934.
Duflo, E., Pande, R., 2007. Dams. Q. J. Econ. 122, 601–646.
Faber, B., 2014. Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. The Review of Economic Studies 81, 1046–1070.
Garz, S., Perova, E., 2021. Prioritising job creation without undermining public works construction among road improvement projects in rural Nicaragua. Journal of Development Effectiveness 13, 224–245.
Gelman, A., Imbens, G., 2019. Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. Journal of Business & Economic Statistics 37, 447–456.
Ghani, E., Goswami, A.G., Kerr, W.R., 2016. Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing. Econ. J. 126, 317–357.
Grembi, V., Nannicini, T., Troiano, U., 2016. Do Fiscal Rules Matter? Am. Econ. J.: Appl. Econ. 8, 1–30.
Hardy, M. and J. McCasland (2017): "Light off, Light on: The Effects of Electricity Shortages on Small Firms, Working Paper.
Hari, S. and K. Raghunathan (2017): "Providing more than just Employment? Evidence from NREGA in India, Working Paper.
Hasan, R., Y. Jiang, and R.M. Rafols (2017): "Place-Based Preferential Tax Policy and Its Spatial Effects: Evidence from India's Program on Industrially Backward Districts, Working Paper.
Henderson, V., Storeygard, A., Weil, D.N., 2012. Measuring Economic Growth from Outer Space. American Economic Review 102, 2012.
Hjort, J., Poulsen, J., 2019. The Arrival of Fast Internet and Employment in Africa. American Economic Review 109, 1032–1079.
Hodler, R., Raschky, P., 2014. Regional Favoritism. Quart. J. Econ. 129, 995–1033.
Hsieh, C.-T., Klenow, P.J., 2014. The Life Cycle of Plants in India and Mexico. Quart. J. Econ. 129, 1035–1084.
Imbens, G.W., Lemieux, T., 2008. Regression discontinuity designs: A guide to practice. Journal of Econometrics 142, 615–635.
Italian G20 Presidency (2021): "G20 Policy Agenda on Infrastructure Maintenance, Tech. rep., G20.
Jayachandran, S. (2020): "Microentrepreneurship in Developing Countries, NBER Working Paper No. 26661.
Jensen, R., 2007. The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. Q. J. Econ. 122, 879–924.
Klomp, J., 2016. Economic Development and Natural Disasters: A Satellite Data Analysis. Global Environmental Change 36, 67–88.
Kolesár, M., Rothe, C., 2018. Inference in Regression Discontinuity Designs with a Discrete Running Variable. American Economic Review 108, 2277–2304.
Lee, D.S., Card, D., 2008. Regression Discontinuity Inference with Specification Error. Journal of Econometrics, 655–674.
Lee, K., Miguel, E., Wolfram, C., 2020. Experimental Evidence on the Economics of Rural Electrification. Journal of Political Economy.
Lenz, L., A. Munyehirwe, J. Peters, and M. Sievert (2017): "Does Large-Scale Infrastructure Investment Alleviate Poverty? Impacts of Rwanda's Electricity Access Roll-Out Program, World Development, 89, 88 – 110.
Lipscomb, M., Mobarak, A.M., Barham, T., 2013. Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil. Am. Econ. J.: Appl. Econ. 5, 200–231.
Lu, Y., Wang, J., Zhu, L., 2019. Place-Based Policies, Creation, and Agglomeration Economies: Evidence from China's Economic Zone Program, American Economic Journal. Economic Policy 11, 325–360.
Mannava, A., E. Perova, and P.T.M. Tran (2020): "Who Benefits from Better Roads and Why? Mixed Methods Analysis of the Gender-Disaggregated Impacts of a Rural Roads Project in Vietnam. World Bank Policy Research Working Paper.
McCrary, J., 2008. Manipulation of the running variable in the regression discontinuity design: A density test. Journal of Econometrics 142, 698–714.
Min, B. (2008): "Democracy and Light: Electoral Accountability and the Provision of Public Goods, Working Paper, mimeo UCLA.
Munshi, K., Rosenzweig, M., 2016. Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap. American Economic Review 106, 46–98.
Nataraj, S., 2011. The Impact of Trade Liberalization on Productivity: Evidence from India's Formal and Informal Manufacturing Sectors. Journal of International Economics 85, 292–301.
Neumark, D. and H. Simpson (2015): "Chapter 18 - Place-Based Policies, in Handbook of Regional and Urban Economics, ed. by G. Duranton, J.V. Henderson, and W.C. Strange, Elsevier, vol. 5 of Handbook of Regional and Urban Economics, 1197 – 1287.
Planning Commission (2003): "Backward Districts Initiative - Rashtriya Sam Vikas Yojana - The Scheme and Guidelines for Preparation of District Plans, Tech. rep., Government of India.
Program Evaluation Organization (2010): "Evaluation Study on Rashtriya Sam Vikas Yojana (RSVY), Tech. rep., Government of India.
Rud, J.P., 2012a. Electricity Provision and industrial Development: Evidence from India. J. Dev. Econ. 97, 352–367.
Rud, J.P., 2012b. Infrastructure regulation and reallocations within industry: Theory and evidence from Indian firms. J. Dev. Econ. 99, 116–127.
Shenoy, A., 2018. Regional Development through Place-Based Policies: Evidence from a Spatial Discontinuity. J. Dev. Econ. 130, 173–189.
Topalova, P., 2010. Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India. Am. Econ. J.: Appl. Econ. 2, 1–41.
Usmani, F. and T. Fetter (2019): "Fracking, Farmers, and Rural Electrification in India, Working Paper.
Wang, J., 2013. The Economic Impact of Special Economic Zones: Evidence from Chinese Municipalities. J. Dev. Econ. 101, 133–147.
World Bank (1994): "World Development Report 1994: Infrastructure for Development, Tech. rep., World Bank.
World Bank (2021): "Well Maintained: Economic Benefits from More Reliable and Resilient Infrastructure, Tech. rep., World Bank.
Zimmermann, L. (2017): "Why Guarantee Employment? Evidence from a Large Indian Public-Works Program, Working Paper.
Zimmermann, L., Khanna, G., 2017. Guns and Butter? Fighting Violence with the Promise of Development. J. Dev. Econ. 124, 120–141.