Rural roads, small business creation, and the gender-gap in entrepreneurship

Sumit Agarwal

Pulak Ghosh

Abhiroop Mukherjee

S. Lakshmi Naaraayanan

November 2023 *

Abstract

We study the causal effect of new roads on small business creation. Our identification strategy uses a program that prioritized road-building for Indian villages with populations above specific thresholds. Findings reveal a doubling of new micro-enterprises in these villages after the program, compared to a 9% increase in their below-threshold counterparts. Our evidence indicates two channels behind this growth. First, paved roads allow year-round accessibility to external markets, which especially matters in flood-prone areas. Second, many new food/beverage-related businesses start up, benefiting from the ability to transport perishable goods more quickly and easily. Notably, the increase in new businesses is largely driven by a four-fold increase in female entrepreneurship, leading to a substantial reduction in the pre-existing gender gap. Two mechanisms, in turn, contribute to this reduction: an erosion of patriarchal social norms resulting from men migrating to towns for work; and a reduction in the time women have to spend on childcare, due to connectivity-induced improvements in local schools.

JEL classification: O2, O18, O43, L26

Keywords: Entrepreneurship, Rural Roads, Small Businesses, Infrastructure and Growth

*Agarwal: National University of Singapore; Ghosh: Indian Institute of Management, Bangalore; Mukherjee: Hong Kong University of Science and Technology, and Naaraayanan: London Business School. We are grateful to Britta Augsburg, Giorgia Barboni, Isaac Hacamo, Rebecca De Simone, Xavier Giné, Sean Higgins, Rajkamal Iyer, Yan Ji, Yatang Lin, Song Ma, Lukas Menkhoff, Ramana Nanda, Kasper Nielsen, Jacopo Ponticelli, Francisco Queiró, Rik Sen, Rui Silva, Manpreet Singh, Daniel Wolfenzon, Constantine Yannelis, Alminas Žaldokas, Nicolas Ziebarth and conference/seminar participants at the DIW (Berlin) Finance and Development Workshop, Nova SBE Entrepreneurship and Innovation Symposium, Junior Entrepreneurial Finance and Innovation Workshop, and Workshop in Finance and Development for helpful comments. We thank Chiman Cheung and Benas Žurauskas for excellent research assistance. This research project was partially funded by the Strategic Public Policy Research (SPPR) Scheme of the Central Policy Unit of the Hong Kong Government, Project No. SPPR17RG01. Naaraayanan thanks both the Wheeler Institute for Business and Development and the Research and Materials Development Grant at London Business School for financial support. Naaraayanan thanks Columbia University for hosting him for a part of the time during which this research was conducted.

1 Introduction

Rural development – aimed at improving living conditions for the 3.5 billion people who still live in villages – is increasingly being linked to rural entrepreneurship. As Petrin and Gannon (1997) from the Food and Agricultural Organization of the United Nations point out,

> *"Institutions and individuals seem to agree on the urgent need to promote rural enterprises: development agencies see rural entrepreneurship as an enormous employment potential; politicians see it as the key strategy to prevent rural unrest; farmers see it as an instrument for improving farm earnings; and women see it as an employment possibility near their homes which provides autonomy, independence and a reduced need for social support. To all these groups, however, entrepreneurship stands as a vehicle to improve the quality of life for individuals, families and communities and to sustain a healthy economy and environment."*

Researchers at many such international organizations argue that insufficient infrastructure, especially poor rural road access, hampers the growth of rural entrepreneurship (e.g., Asian Development Bank Operations Evaluation Department (2002), World Bank (2022)). Villagers starting small businesses, such as bakeries, handloom weaving, or dairies, need easy and quick access to resources and markets. Given the small scale of village economies, these resources and markets are frequently located away from the village, making all-weather road access key for business success.

For instance, the ILO shares the story of Cecilia Baptista from Dare village, about 15 kilometers from Dili, Timor-Leste's capital. Previously, the poor condition of the rural road through her *suco* made it impassable during the rainy season, forcing her 15-member family to rely on subsistence farming. But when a government program built an all-weather road passing through her village, Cecilia's family started a bakery: "With the improved road, it is now possible to rent a truck and transport flour from Dili at affordable prices". Cecilia initially baked bread for her neighbors and relatives. As demand for her bread grew, she expanded her business, again using the new road to reach larger markets nearby.

Hoping to create more rural entrepreneurs like Cecilia, many developing nations have begun constructing vast networks of rural roads over the past two decades. For example, between 2000 and 2016, India alone built 1.96 million kilometers of rural roads, granting many villagers year-round access to the outside world for the first time. But does such improved accessibility through new rural roads meaningfully impact rural entrepreneurship?

We use population-based discontinuities in road construction to identify the causal effect of road access on the propensity to start small rural businesses. Our findings indicate that twice as

many new rural manufacturing businesses are started in villages just above the road construction thresholds compared to those barely below. Notably, an equal number of these new businesses are started by women and men. Considering that before our road program, women started one business for every seven started by men, this constitutes a significant decrease in the entrepreneurship gender gap.

Understanding the link between small-scale entrepreneurship and new infrastructure is challenging because it is difficult to determine a suitable counterfactual — a scenario which could show what would have happened to newly connected villages if they had remained unconnected. An ideal experiment to create this counterfactual would randomly provide some villages with roads while leaving others without. Unfortunately, such an experiment is rarely feasible in the real world.

We find a way to progress by utilizing a policy directive surrounding India's *Pradhan Mantri Gram Sadak Yojna*, (henceforth PMGSY), a program designed to build all-weather roads for previously unconnected villages. Program rules explicitly listed building new roads to connect villages above pre-specified population thresholds. This resulted in discontinuities in the probability of treatment at these population thresholds, which we exploit to identify our effects. For example, villages with populations just above a round figure, say 500, were to be prioritized under the program. Under the assumption that villages with populations just below the threshold are very similar to those above, particularly if they are geographically close, the resulting variation in roads is quasi-random. We first confirm that earlier findings, which showed that these population thresholds predict road construction, also apply to our sample: a village just above the threshold is 51% more likely to get a road than one just below it. Further, our first-stage F-statistics range from 18.6 to 69.7, indicating that thresholds are a strong and valid instrument for road construction.

Our data on rural entrepreneurship comes from an enterprise-level survey dataset made available through the All India Census of Micro, Small and Medium Enterprises (henceforth, MSME). Given that we examine businesses in small villages within a narrow bandwidth of PMGSY cutoffs, these businesses are mostly micro enterprises: the average size of business started, based on investment in plant and machinery, is Rs. 24,000 (around 530 USD in 2006), and their average number of employees is 1.4.

Our analysis is based on a differences-in-discontinuity specification. In essence, this is similar to a difference-in-differences specification. The time series difference comes from the change in the number of new businesses started in a village post-PMGSY vs. pre-PMGSY. The cross-sectional element comes from a discontinuity: it involves comparing the time-series change in entrepreneurship across two similar villages – one above, and the other below the population-thresholds used by PMGSY.

Our baseline results show significant effects of road connectivity: there were 4 new businesses

started per 1000 villagers before the PMGSY program in an above-cutoff village. This increases to an average of 8.4 new businesses in the same village after PMGSY. The corresponding change in the number of new businesses started in below-cutoff villages is from 4 to 4.36. These results are robust: they are not sensitive to the choice of bandwidths, or to functional form assumptions about the relationship between entrepreneurship and village population.

Next, we study the two main channels through which policy makers expect roads to affect entrepreneurship, and find suggestive evidence in support for both. First, we find that connectivity throughout the seasons, including during monsoons when mud paths can flood, indeed seems to be important. This is reflected by the fact that new businesses respond more to roads in areas prone to drastic monsoonal flooding. Second, we also find that about one-fourth of the aggregate increase in entrepreneurship is driven by new food and beverage-related enterprises. These enterprises rely on their ability to reach input and/or output markets quickly and easily with perishables, which is now made possible by new roads.

Interestingly, we see that a significant share of these new businesses are started by women. Our most conservative estimates suggest that there are two more new businesses per 1000 villagers started by women in villages above thresholds after PMGSY, relative to below-threshold villages. Compared to a baseline rate of 0.5 female-led enterprises per 100 villagers in the pre-PMGSY period, this is a notable increase. We also see two more businesses started by men in these villages after PMGSY, but (i) this is not statistically significant, and (ii) compared to a much larger base-rate – 3.4 new male-led businesses used to be started in the same villages in the pre-PMGSY period – it reflects a much smaller percentage change.

To understand why roads might be helping reduce the gender gap in business-starting, we examine three channels: (1) the availability of formal financing in connected villages , which can be less biased against women (Agarwal, Mukherjee, and Naaraayanan (2022); for evidence on the effectiveness of financing, see Beck et al. (2014), King and Levine (1993a,b), Levine (1997, 2005, 2008), Demirguc-Kunt and Levine (2008a,b), Beck (2012), Demirguc-Kunt et al. (2013)). (2) significant outmigration of men from connected villages for work in towns (Asher and Novosad, 2017, 2020), which might erode gender-biased social norms that have been shown to be important in this context, by e.g., Goldin and Olivetti (2013), Field et al. (2010), and Naaraayanan (2022). and (3) better schools, due to roads allowing teachers to commute from nearby towns (SCIO), which can reduce the time women spend looking after children (Kimmel (1998), Gelbach (2002)). We find evidence consistent with the latter two channels – more women start small businesses in villages with more male outmigration and those with better schools. However, there is no evidence in favor of the financing channel, possibly due to the fact these microenterprises are less reliant on formal financing.

Finally, businesses that start in above-cutoff villages hire as many people as those in below-cutoff villages. However, we find that this masks significant heterogeneity in the gender composition in employment. Relative to new businesses started in villages below PMGSY thresholds, entrepreneurs in villages above these thresholds are likely to employ 40% more women. This is consistent with the idea that women entrepreneurs in socially conservative rural India are more likely to employ other women (Ghani et al. (2016)), with rural roads enabling female labor force participation.

We contribute mainly to the literature on small business creation in developing countries. Closest to our paper, Chaurey and Le (2022) study infrastructure grants targeted towards the least developed districts in India for maintenance, improvements, and complementary investments to existing infrastructure, and find that these grants stimulated village-level employment, as well as microenterpreneurship. In particular, they find stronger effects in electricity and road-intensive industries, and in villages that had paved roads and electricity prior to the program. We study different programs using a cleaner identification strategy, leading to a few key differences. (1) We focus on new all-weather roads, not maintaining or improving existing infrastructure. (2) By focusing specifically on roads, we can take a closer look at mechanisms behind any impact, often key to policy design. (3) Most importantly, we examine how the gender gap in entrepreneurship responds to connectivity, an issue of much recent policy interest. In another related paper, Ghani et al. (2016) also examine entrepreneurship in the context of another road network in India (the Golden Quadrilateral), but focus on typically larger manufacturing units which are usually located in more urban areas; so the ways roads might be helping there are likely different from our study.

More generally, rural entrepreneurship has been getting significant attention in policy circles (see, e.g., Blattman et al. (2013), Drabenstott et al. (2003), Mukherjee and Zhang (2007), McElwee and Smith (2014)). Jayachandran (2020) provides a recent review of this literature. However, the causal effect of new rural roads on small business creation has not been studied in the literature, despite being a popular narrative in policy circles. Our paper takes a modest first step towards this goal.

Our evidence on a substantial response of female microentrepreneurship to roads in rural India is important in light of the earlier findings from multiple countries in Rijkers and Costa (2012), who show that women are less likely to be non-farm entrepreneurs than men are. Taken together, our results suggest that one reason behind this difference is the lack of road access in small villages. Field et al. (2010), Bernhardt et al. (2017) and Naaraayanan (2022) document social norms as a constraint to female entrepreneurship. We find a similarly encouraging picture for women's entrepreneurship in India in response to roads as Ghani et al. (2014); the main difference here being that road-building was not a policy targeted towards gender equality, unlike the political reservations they examine. Roads, then, seem another determinant of small business creation by

women in developing countries (see, e.g., Revenga and Dooley (2020) for a survey of determinants).

The PMGSY program we study, of course, has also been used to show that new roads led to higher school enrolment and retention (Mukherjee (2011), Adukia, Asher, and Novosad (2020)), productivity improvements for affected villages (Shamdasani (2021) and Aggarwal (2018)), more financing for productive uses (Agarwal, Mukherjee, and Naaraayanan, 2022), and a reallocation of villagers from agriculture to wage labor (Asher and Novosad, 2017, 2020). Our research is different because we are looking at how roads might affect small businesses in rural areas, a topic that hasn't been studied before despite a lot of policy discussion. Furthermore, our results qualify and extend those in Asher and Novosad (2020): while they show that in an average village roads have little impact on most economic outcomes at the village level, we show that it might be possible to see some benefits in rural entrepreneurship, especially in reducing the entrepreneurship gender gap. Our results agree in that such new small businesses are unlikely to play any transformative role for overall village-level economy. Still, our evidence suggests that a broader analysis of these small businesses should consider their impact on the gender gap in starting businesses, and their ability to provide more jobs for women. Both of these are important goals in rural development.

Our evidence also adds more generally to the literature estimating effects of public infrastructure in low- and middle-income countries. This literature generally finds economically meaningful effects of such projects on a wide range of outcomes. Specifically, transportation infrastructure has been shown to raise the value of agricultural land (Donaldson and Hornbeck (2016)), increase agricultural trade and income (Donaldson (2018)), reduce the risk of famine (Burgess and Donaldson (2012)), increase migration (Morten and Oliveira (2014)), and accelerate urban decentralization (Baum-Snow et al. (2017)). In addition, there is mixed evidence that transportation costs can increase (Ghani et al. (2016, 2017); Khanna, 2014; Storeygard (2016); Das et al. (2019)), decrease (Faber (2014)), or leave unchanged (Banerjee et al. (2020)) growth rates in local economic activity.

2 Data

The main dataset we use is the Fourth All India Census of Micro, Small and Medium Enterprises 2006-07. The data was collected until 2009, the results of which were published in 2011-12. This dataset contains detailed information on registered as well as unregistered businesses across various locations in India, along with information on first production dates, original value of plant and machinery, and the gender and social caste of the entrepreneur. The census adopted different methodologies for registered and unregistered sectors. While complete enumeration of enterprises was adopted for the registered sector, a sample survey was resorted to for the unregistered sector, which includes enterprises eligible for registration as a small-scale industry (SSI). All enterprises

in our village sample turn out to be unregistered, not surprising given that narrow bandwidths around PMGSY population-cutoffs restrict our analysis to small villages only.

We use the MSME data for two reasons: 1) it is the only dataset on entrepreneurship in villages; and 2) it is the only dataset that provides information on start up dates for rural enterprises (Naaraayanan (2022)).[1] Moreover, in most of our analyses, we restrict our attention to unregistered enterprises in the manufacturing sector (which constitutes 74% of all unregistered enterprises in the six states we consider). This is because service sector enterprises, which are relatively rare in village India, do not seem to respond to our roads program (Internet Appendix Table IA1 presents these results for completeness).

We obtain data on rural road construction in India from the website of *Pradhan Mantri Gram Sadak Yojna* (PMGSY), the road-building program in India we study. The data includes detailed information on road sanction and completion dates, which we scrape.[2] The PMGSY data is structured to consist of information both at the habitation-level and at the road-level. The correspondence between habitation and roads is many-to-many, as several roads serve multiple habitations, and habitations may be connected to multiple roads. We define our treatment at the village-level wherein a village is considered as "treated" under PMGSY if at least one habitation in the village – which was previously unconnected to the paved "all-weather" road network – received a (completed) road during our sample period. We conduct our analysis at the village-level because all of our villages consist of one habitation.

We hand-match the administrative road data to the location of an unregistered enterprise in the MSME dataset. We successfully match over 90% of habitations listed on the PMGSY website to their corresponding census villages. Overall, we are able to match 983 villages in 6 states across both datasets. We also use data on demographics and village-level amenities (such as electricity, roads, schools etc) from the 2001 Population Census and the previously listed PMGSY webpage. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of villages that did not have paved roads at the start of our sample.

We supplement this dataset with data on school enrolment in each village from District Information System for Education (DISE), village-level outcomes from the Economic Census (EC) conducted by the Government of India, district-level rainfall data from the Indian Meteorological Department (IMD) and the Indian Space Research Organisation (ISRO), natural disasters from

[1] Widely-used datasets like the Economic Census do not contain startup dates, while others such as CMIE Prowess and Annual Survey of Industries do not have geographic information at the village-level.

[2] We scrape the data as on January 2018.

the EM-DAT database, and aggregate district-level lending data from the Reserve Bank of India (RBI).

3 PMGSY and empirical strategy

3.1 The PMGSY program

The main challenge in identifying the impact of infrastructure investments on financing is endogeneity of such infrastructure. Factors such as political favoritism or local economic conditions could be correlated directly with both road placement as well as the outcomes of interest, which can render OLS estimates biased (Beck (2008)). In this section, we describe the empirical strategy we use to make some progress in identification.

Our identification strategy is based on guidelines set forth by a national road building program called *Pradhan Mantri Gram Sadak Yojana* (PMGSY). This program was launched by the central government in December of 2000 to provide access to "all-weather" roads to 74% of India's population that still lived in villages. PMGSY proved to be one of the largest rural road programs the world has ever seen, with 480,000 kilometers of rural roads built under it by 2016, doubling the size of India's rural road network.

The program mainly focused on hitherto unconnected villages, defined as those without any pre-existing all-weather road within 500 meters of its boundaries, and the aim was to construct roads to connect them to the closest town or market center, identified by the program as the "center of activities for marketing agricultural produce and inputs, servicing of agricultural implements, health, higher education, postal, banking services etc."

Program guidelines prioritized villages to receive new roads based on population thresholds. At the time most of these roads were constructed, the last nationwide official population record was from the 2001 census. The instructions required state officials to target villages in the following order: (i) villages with population greater than 1000; (ii) villages with population greater than 500; and (iii) villages with population greater than 250. Our identifying assumption is therefore that even if selection into road connectivity could be determined by many factors in general, these factors are not likely to change discontinuously at these population thresholds. Hence, if these rules were followed by the officials in charge – which we can test – we can estimate the effect of road connectivity on financing outcomes using a discontinuity design.

Earlier papers, for example, Asher and Novosad (2020) and Agarwal, Mukherjee, and Naaraayanan (2022), have used PMGSY-based discontinuity and shown its validity/strength as an instrument.

We show that instrument strength/validity extends to our sample of states, and is robust to various non-linear specifications and bandwidths, in the following sections.

Note that while other programs, such as the district-level program studied by Chaurey and Le (2022), were also undertaken within our sample period to improve rural economies, concerns for contamination are greatly attenuated by the fact that our identification strategy is based on within-district, village-population-level discontinuities unique to PMGSY.

3.2 Empirical strategy

We first test for threshold manipulation under the PMGSY program. This is important to understand whether, for example, a powerful politician was getting local officials to systematically classify some villages with populations below the threshold as being above it, so that these villages get roads. This can be problematic for identification, since then we would not know whether any effect we identify in villages that get roads is indeed attributable to the road connectivity, or to the same politician's simultaneous influence on rural entrepreneurship.

To make sure that our estimates are not confounded by such issues, we use population figures from 2001 Population Census, which was conducted before PMGSY policy cutoffs were finalized. While this may produce noise in estimates if the road-building authorities used more updated figures, it ensures validity.[3] Still, we check for any indication of manipulation using tests for discontinuities in the density of our running variable, village-level population (McCrary (2008)).

In Figure 1, where we plot the histogram of villages in our six states by population as recorded in 2001 Census, we can see that there are no discrete jumps in population around the PMGSY thresholds of 500 and 1000, indicating no manipulation for these thresholds. Following McCrary (2008), we also test for discontinuity in our pooled sample, combining the thresholds into one cutoff, and normalizing the village population by subtracting the relevant thresholds, failing to reject the null hypothesis of no discontinuity.

In Figure 2, we examine a simple scatter plot of the proportion of villagers in each population bin with access to a road. Our evidence indicates that there is a significant jump in the probability of receiving a road just above the population cutoff relative to just below it. The figure also shows two quadratic polynomials, fit separately above and below the cutoff. To check for the statistical stability of this jump, we estimate the following regression specification as a first step, where we

[3] As we show in Section 4.1, our estimates still retain enough power to ensure that our cutoffs are not weak instruments.

examine the effect of population cutoffs on actual PMGSY road construction:

$$\begin{aligned} Road_{i,v} = \beta_0 + \beta_1 1[Population\_above\_threshold] \\ + \beta_2 1[500 - h \leq pop_{i,v} < 500 + h] * f(pop_{i,v} - 500) \\ + \beta_3 1[1000 - h \leq pop_{i,v} < 1000 + h] + \beta_4 1[1000 - h \leq pop_{i,v} < 1000 + h] * f(pop_{i,v} - 1000) + \epsilon_{i,v} \end{aligned} \tag{1}$$

where $Road_{i,v}$ measures whether the village v – unconnected as of 2001 – received a PMGSY road by 2009, and $pop_{i,v}$ is the baseline village population as recorded in 2001 census, and $f(pop_{i,v})$ is a polynomial (quadratic or cubic) function of (scaled) population. We restrict our sample to villages with population within a certain bandwidth around the threshold, such that $pop_{i,v} \in$ [c-h, c+h], where h is the population bandwidth around threshold c.

Further, our identifying assumption is that crossing the population threshold discontinuously affects only the probability of receiving a road under PMGSY – but is not discontinuously related to other things at the village level. Under this assumption, there should be no jumps in other village characteristics (baseline covariates) at the population thresholds (Imbens and Lemieux, 2008) used for road building. In Figure 3, we examine a simple scatter plot of means of various village characteristics by different population bins (each with size of 25) around the threshold, to check for discontinuities in baseline village-level covariates, and find no such evidence. The characteristics we examine include the presence of schools, health centers, electricity, presence of a telegraph office, distance from the nearest town, share of scheduled castes in population, and land irrigated.

Table 1 examines this evidence for statistically discernible differences, using these characteristics one by one as dependent variables in the regression discontinuity setting described in Equation 1. Panel A of that table shows that in the overall Population Census sample, above-cutoff villages tend to have no significant differences with those below, with the exception of the proportion of land irrigated (which is 8% higher in above-cutoff villages). When we examine our merged MSME sample that we use for small businesses (Panel B of Table 1), none of these characteristics are statistically different across cutoffs.

Note that our identification comes from two sources: (1) *discontinuous threshold* effects based on population-based rules, e.g., villages with population 550 vs 450, and (2) the *timing of new business starts* – before these threshold rules came into effect with PMGSY (so in the year 2000 in our sample), vs. after. In essence, we compare entrepreneurial activity in villages above vs. below thresholds after PMGSY against a benchmark 'baseline' difference between the same above

vs. below villages before PMGSY. As a matter of deliberate choice we do not exploit the variation across villages in the timing of actual road construction. While there is indeed some additional variation in *when* individual villages receive roads, this is largely *endogenous*. While the government rules specify that a village with a population of, say, 1080 should get a road before a village with a population of 920, they do not specify whether a village with a population of 1080 should get a road before or after one with a population of 1170.

Next, we turn to the sample of unregistered MSME businesses to test our key hypotheses. The merge between the MSME dataset and PMGSY sample yields 130 villages within a population bandwidth of ± 100. We choose 100 as the main bandwidth in our tests, although results are robust to higher bandwidths as well. Our results remain qualitatively similar if we restrict the bandwidth further but the number of villages falls rapidly, and the resultant decline in statistical power makes our estimates lose significance. Given the low number of villages in our sample, we do not estimate the thresholds of 500 and 1000 separately due to statistical power considerations. In addition, in our main specifications we allow outcome variables to be related to population differently in villages with populations around 500 and 1000, with different slopes, quadratic (and cubic) terms, in addition to different intercepts.

In this merged sample, we employ a specification similar to a difference-in-differences specification, where the time difference comes from *within* village changes in the number of entrepreneurs per capita who started businesses after the PMGSY program, relative to those who started before the PMGSY program. The cross-sectional difference comes from the presence of a road, instrumented by the population-based thresholds – villages above the thresholds being more likely to be treated, as we found from the first stage. Our results are identified under the assumption that the only meaningful difference between these villages across thresholds, which we verified above, is in their likelihood of getting a new road.

Figure 4 examines rural entrepreneurship, analogous to Figure 2. In this scatter-plot, we residualize changes in entrepreneurship by adding district fixed effects to account for time-invariant district characteristics. As before, we also fit a quadratic polynomial on either sides of the cutoff. This figure shows a jump in small business creation at the normalized cutoff for road-building, despite usual small-sample noise.

To keep our interpretation simple for more formal tests, we focus on reduced form estimation. Note that this means that our estimates provide an 'intent-to-treat' effect rather than an average treatment effect of receiving a new rural road on our outcomes. The baseline specification we use to estimate the effect of rural roads on rural entrepreneurship is a long-differenced reduced form specification, where we use the size of the jump at the aforementioned population-based threshold.

Specifically, we estimate

$$\Delta Y_{i,v} = \gamma_0 + \gamma_1 1[Population\_above\_threshold] + \gamma_2 1[500 - h \leq pop_{i,v} < 500 + h] * f(pop_{i,v} - 500) + \gamma_3 1[1000 - h \leq pop_{i,v} < 1000 + h] + \gamma_4 1[1000 - h \leq pop_{i,v} < 1000 + h] * f(pop_{i,v} - 1000) + \epsilon_{i,v} \quad (2)$$

where $\Delta Y_{i,v}$ are the long differenced outcomes of interest, e.g., the number of new businesses per capita started in a village v in 2007 minus the same quantity in 2000. We illustrate the timing convention as follows using the example of our baseline entrepreneurship rate variable, but our other tables also follow the same convention. First, we differentiate all enterprises surveyed in the data into those that were started before PMGSY (when there is no difference in the probability of receiving a road between villages that are above and below the respective population cutoff) and those started after (when the cutoffs determined the propensity to get a new rural road). Then we take the difference (Δ) between the total number of new unregistered MSME businesses started after 2000 and those started before 2000, and scale the change by village population as recorded in the 2001 Population Census. We use heteroskedasticity-robust standard errors throughout.

4 Results

In this section, we describe and discuss our main findings.

4.1 Do population cutoffs predict road construction?

Table 2 formalizes the visual evidence in Figure 2 by presenting first stage estimates from Equation 1 in the census sample of all PMGSY villages without a paved road. Here, our unit of analysis is a village. The estimates from column 1 imply a 3 percentage point increase in the probability of treatment around the cutoff. This jump is statistically significant, with F-statistic above 40, implying that we are not subject to a weak instrument problem. Importantly, the economic magnitude of this jump is large: the unconditional probability of getting a road is about 5.8%, so this is about a 51% jump.

Note that our evidence suggests that the economic magnitudes of our estimates do not differ significantly across specifications or bandwidths ranging between 80 to 120. Overall, our results confirm that there is a significant increase in the probability of getting a road around the population threshold. These results complement prior work by Asher and Novosad (2017, 2020) and Agarwal,

Mukherjee, and Naaraayanan (2022) on the validity of PMGSY.

4.2 Summary statistics for the MSME sample

Table 3 reports the summary statistics for our sample of unregistered businesses in the matched dataset. Note that although the MSME data includes registered businesses, all villages in our sample are unregistered, and most are micro enterprises (below Rs. 25,00,000 investment in plant and machinery). This is to be expected given that we are examining rural entrepreneurship in sparsely-populated villages that get a paved road for the first time.

To save space, Table 3 focuses on villages within a bandwidth of 100 around the population thresholds. On average, four new businesses gets started per 1000 villagers; of which only 14.8% are started by women.[4] The average business is 7.4 years old with a size, based on investment in plant and machinery, of around Rs. 24,000 (around 350 USD), indicating these are on average microenterprises. On average, 0.7% of businesses suffer a decline in gross output between 2006-07 and 2007-08, while around 17% of businesses suffered a decline in net worth between 2006-07 and 2007-08. Lastly, the average firm in our sample has around 1.4 employees, but only about one in four new businesses employ a woman in their enterprise. This means that the ratio of female to male employment at new businesses is about 17.1% – a similar order of magnitude as the ratio of female to male entrepreneurs.

4.3 Change in rural entrepreneurship rates

Table 4 presents our main results from estimating Equation 2. Columns 1–5 (respectively, columns 6–10) present discontinuity estimates for a second (third) order polynomial fit to the underlying relationship between entrepreneurship rates and village population. Again, we present results across five bandwidths between 80 and 120.

We find that the average village above the population threshold saw a significantly larger increase in entrepreneurship rates than the average village below. The economic magnitude of these effects suggests that relative to villages below, villages above saw four more new businesses started per 1000 villagers, using the most conservative estimate across specifications. Relative to a baseline of four new businesses that were started in the average village in the pre-PMGSY period (from Table 3 above), micro-entrepreneurship rates doubled in above-cutoff villages in the post-PMGSY period. In contrast, the entrepreneurship rate in a below-cutoff village only increased

[4] A female enterprise is defined as an enterprise managed by one or more women entrepreneurs in a proprietary concern where she/they individually or jointly have a share of capital of not less than 51 percent as partners.

from 4 to 4.36 in the same period. These results are robust to modifications in specifications and bandwidths, as the Table shows.

Overall, our evidence suggests significantly higher growth in micro-entrepreneurship in villages more likely to get a new paved road under PMGSY.

4.4 Why do rural roads lead to new business creation?

In this section we examine two important mechanisms through which policymakers ((e.g., ILO, or the World Bank) expect roads to affect small business creation.

First, as in our example of Cecilia's bakery in Timor-Leste in the introduction, paved roads are supposed to provide all-weather access to input and/or output markets, enabling villagers to set-up businesses. All-season access is a particularly important issue in countries like India, where many villages get cut-off from the rest of the world due to heavy rain washing out or flooding mud roads during the infamous southwest monsoon season. In fact, the PMGSY program guidelines clearly state:

> *"The primary focus of the PMGSY is to provide All-weather road connectivity to the eligible unconnected Habitations. An All-weather road is one which is negotiable in all seasons of the year. This implies that the road-bed is drained effectively (by adequate cross-drainage structures such as culverts, minor bridges and causeways).."*

To the benefit of the econometrician, however, not all districts in India are equally prone to such monsoonal disruptions. If this channel – i.e., all-season access to the outside world – is important, then we should see more pronounced effects of road-building in areas particularly susceptible to such seasonal disruptions.

In Table 5, we test this hypothesis, and find support for it in our sample. Using the maximum monthly rainfall in each district during monsoons (June-August) averaged over 2000-2005 calculated from the ISRO dataset, we find that microentrepreneurship effects are at least 33% stronger in areas with severe rainfall – where are roads are more likely to have reduced or removed seasonal disruptions to connectivity.

Second, we study the market access explanation mentioned popularly by many senior policy experts, for example, by Ms. Sylvie Bossoutrot, World Bank Country Manager for Armenia, Europe and Central Asia,

"It is vital that rural infrastructure is improved and maintained to promote agricultural trade, so that farmers and small businesses are able to bring their products to market more easily and at a lower cost, thereby stimulating economic growth and local employment."

Here, we design our test to reflect that timely market access, as provided by new rural roads and the resultant improved transport facilities they likely provide, is particularly important for some industries, such as food processing. In the food processing industry, for example, for dairies in Indian villages, it is often imperative to transport the produce quickly to outside markets. If this is an important channel, then we should see more new businesses in the food processing industry in response to roads.

In Table 6, we focus only on new businesses in the food and beverages industry.[5] As expected, we see a significant increase in such businesses across our cutoffs across many of our specifications, in spite of a substantial drop in sample size due to our focus on one industry here. In terms of economic magnitudes, out of the four extra new businesses in above-cutoff villages, one comes from the food and beverages industry.

4.5 New business performance

In Table 7, we examine the differences in performance of new businesses that are likely to have been started in response to road connectivity. Again, we lack high frequency data on financial statements over time. However, we do have information on whether the net worth or output from these businesses suffered declines in the survey year, so this is what we use. The main limitation is that these performance metrics refer to a specific point in time, rather than an average across different market conditions. While this is likely to create noise in our measures, it is unlikely that such noise will be different across the thresholds, especially in our narrow bandwidths. Finally, we present results only at the bandwidth of 100, but results are similar at other bandwidths considered in the previous tables. We still continue to present evidence from both our quadratic (first two columns) and cubic (next two columns) specifications.

In Panel A, we take the difference between the fraction of new businesses started after PMGSY suffering a decline in net worth, and the fraction of those started before. We contrast this difference across above- versus below-threshold villages. In Panel B, we examine a similar specification using

[5]Note that due to the limited number of observations in this subsample, we are unable to include district-cutoff fixed effects.

changes in business output in rupee terms.

Our results do not uncover any significant differences in the propensity of output or net worth declines between businesses in above- versus below-threshold villages. These results suggest that the new businesses are of similar quality as proxied by these measures of performance.

In Internet Appendix Table IA2, we use data on the rupee value of plant and machinery at the start of the business. This is a measure of the size of business started. Here, we take the difference between the average asset size, at origination, of all new businesses started after PMGSY and those started before; and contrast this difference across above- vs. below-cutoff villages. Our results suggest that the average business started after PMGSY in above-threshold villages was somewhat smaller than those started earlier, although, again, these differences are almost always statistically insignificant.

Overall, new businesses likely to have been started in response to rural roads do not seem statistically different from other new businesses.

4.6 Roads and the entrepreneurship gender gap

Here we examine changes in the number of businesses started by women and men separately. encouraging is an important exercise because rural Indian villages have a very large gender gap in entrepreneurship (Naaraayanan (2022)). Even in our sample, (e.g., Table 3), such a gender gap is stark – for every six new businesses started by men, there is only one started by a woman.

As with Table 4, here we study the change in entrepreneurship in the post-PMGSY vs. the pre-PMGSY period, and contrast this change in villages above road-building population thresholds with those below. Our evidence suggests that rural roads significantly increase the number of female entrepreneurs. The estimates suggest that relative to villages below, villages above saw two to three more new businesses (per 1000 villagers) started by women in the post-PMGSY period. This contrasts with 0.6 new businesses started by women in our sample in the pre-road period – about a four-fold increase in female entrepreneurship. On the other hand, while we also see a similar increase in the number of new businesses started by men in above-cutoff villages post roads (using our most conservative estimates), these effects are on a much larger base rate (3.4 new male microentrepreneurs in the average village pre-PMGSY), and never statistically significant.

In Internet Appendix Table IA3, we separately examine entrepreneurship rates among the lowest strata of rural Indian society – the Scheduled Castes (SC), Scheduled Tribes (ST), and the Other Backward Castes (OBCs) – vs. all other castes. Our evidence does not suggest any statistically robust differences here. This seems to suggest that there is something special about roads in

how they lead to more women entrepreneurs, but do not lead more generally to improvements in entrepreneurship rates for other potentially discriminated categories in rural India.

4.7 Why are there more female-owned small businesses?

So far, we have found that rural entrepreneurship activity – particularly by women – responds to new road connectivity in the previous sections. But can we say something more about the underlying mechanism? We consider three channels through which female entrepreneurship in rural India could be responding to roads, as mentioned in the introduction.

4.7.1 Financing

In many emerging markets, financing constraints act as a critical barrier to entry for aspiring entrepreneurs, especially women (Beck, 2008; Kerr and Nanda, 2009; Naaraayanan, 2022)). Besides, there is evidence that financing follows infrastructure improvements (Agarwal, Mukherjee, and Naaraayanan, 2022; Naaraayanan and Wolfenzon, 2022; Das et al., 2019). Hence, rural credit from banks might be one potential mechanism that facilitates productivity improvements for the rural poor and can lead to the documented increase in rural entrepreneurial rates.

Unfortunately, we do not have data that provides us with lending information at the same level of granularity which allows similar analysis. However, the Reserve Bank of India (RBI) does provide macro data, aggregated by districts, on overall lending activity by sector (rural, urban, etc.). We use this data to examine the effects of new rural roads and whether these effects are qualitatively consistent with our earlier results.

Table 9 presents the results. Here, we interact our indicator for whether the village is above the population threshold with the natural logarithm of total district-level rural lending. This interaction allows us to capture the effect of new rural roads on the change in female entrepreneurial activity, depending on the depth of the rural credit market in each district. Results from this table do not suggest any significant relation between the availability of finance and female entrepreneurship.

One reason behind this finding is that because the new businesses in our sample are small (requiring an average investment of about Rs. 24,000), capital requirements are met from savings, with help from family and friends, or with loans from local money-lenders. Unfortunately data limitations do not allow us to delve further into this issue.

4.7.2 Male outmigration

Asher and Novosad (2020) show evidence of migration out of newly connected villages. We hypothesize that when men outmigrate from the villages, it frees up time for women in their household to start small businesses. This could happen for at least two reasons. First, male outmigration in rural India could allow women to start a new business due to the relaxation of gender-biased social norms within the household. This relaxation, in part, could come due to their husbands and brothers seeing more women work in the cities to which they migrate, making less taboo the notion that a woman in their family can work for a living. Second, it could also be that such migration simply yields more time for the women in the household as they no longer spend it 'taking care of' the male family members – doing their laundry, cooking them meals, etc.

To test for the hypothesis mentioned above, we create a measure of net outmigration with data from the Economic Census, defined as male minus female outmigration rates from villages. In panel B of Table 9, the second row presents the interaction of such outmigration with our above-cutoff variable. Results show that female entrepreneurship in above-cutoff-villages – more likely to be connected with a road – is indeed higher in villages that more men emigrate from, consistent with our hypothesis.

4.7.3 School enrollment

Another possible link between roads and female entrepreneurship lies in improvements in schooling brought about by connectivity. With better schools – say, due to teacher vacancies being filled in connected villages – women would get more time on hand to start and manage a small business if they sent their children off to schools (see, e.g., Gelbach (2002); Adukia, Asher, and Novosad (2020)).

We test this hypothesis in panel C of Table 9 by examining whether change in female entrepreneurship rates are higher in above-cutoff villages where school enrolment is high. Here we use secondary school enrolment per capita in 2005, since most of the variation in official school enrolment across villages in rural India comes from this group. We obtain data on school enrolment in each village from the District Information System for Education (DISE). The evidence we uncover is consistent with the hypothesis that the ability to send their children to school frees time for women, allowing them to start their own enterprise – many more women start small businesses in above-cutoff villages where school enrolment is higher.

5 Employment at new businesses

Taken together, our results up to this point suggest that while there were large changes in the rate of new business creation, these effects are likely not dramatic enough to leave a trace in village-level income or consumption overall, given that these are micro enterprises started by villagers, and probably not transformational enterprises (Schoar, 2010; Lerner and Schoar , 2010). Our results therefore are likely to be consistent with Asher and Novosad (2020) on overall economic growth at village-level. However, we do offer a glimmer of hope that roads do have some positive effects even at the village-level – encouraging women to start new businesses is a policy goal that has long proved elusive, and roads seem to indeed affect the gender gap in entrepreneurship in a very meaningful way. Moreover, businesses started by women can have other effects in correcting gender biased norms in rural societies, for example, by employing more women than businesses started by men. Here, we explore such employment effects.

We examine aggregate employment effects in Table 10. Similar to our other tests, we first take the change in the (natural logarithm of) number of employees who work for new businesses started after PMGSY, relative to those started before PMGSY. Again, we contrast these differences between businesses in above and below-threshold villages. We do not find any systematic evidence to suggest that businesses likely to have been created in response to new roads employed more workers.

However, in Panel B of Table 10, we examine whether new businesses started in above-cutoff villages employed more women. This test is motivated by the evidence we found on the increase in women-led businesses in Table 8, and the fact that women tend to work with other women. Evidence in favor of the latter view is presented in Ghani et al. (2016), who find that about 90% of employees in female-owned business in unorganized manufacturing are females. Moreover, this is likely to be especially true in socially conservative rural settings.

We run a specification similar to that in Table 8 above, except we measure changes in the number of female employees at newly formed businesses in the current tests. We find robust evidence supporting the hypothesis that new businesses started in above-cutoff villages employed more women after PMGSY, relative to below-cutoff villages. While female employee participation in new businesses rose across the board, above-threshold villages saw a 40.1%-80.7% increase, compared to a much more modest 7.4% increase in below-threshold villages.

6 Conclusion

We use a population-based discontinuity setting around a large rural road construction program in India to examine changes in rural entrepreneurship. Our results show twice the number of new rural enterprises in villages above population thresholds used to prioritize road construction, relative to those below. Importantly, our evidence suggests that a large part of this increase comes from new women entrepreneurs in these villages, along with an increase in the number of women working for such new businesses. Gender-biased familial norms and time constraints for women due to child-minding duties seem to hinder rural female micro-entrepreneurship more than financial constraints.

References

Asian Development Bank. 2002. Impact of rural roads on poverty reduction: A case study-based analysis. *Report IE-68, Operations Evaluation Department.*

Adukia, A., S. Asher, and P. Novosad. 2020. Educational investment responses to economic opportunity: Evidence from Indian road construction. *American Economic Journal: Applied Economics* 12:348–76.

Agarwal, S., S. Alok, P. Ghosh, S. Ghosh, T. Piskorski, and A. Seru. 2017. Banking the unbanked: What do 255 million new bank accounts reveal about financial access? Working Paper, Columbia Business School.

Agarwal, S., M. Ayyagiri, Y. Cheng, P. Ghosh. 2021. Road to stock market participation. Working Paper, National University of Singapore.

Agarwal, S., A. Mukherjee, and S. L. Naaraayanan. 2022. Roads and Loans. *Review of Financial Studies*, Forthcoming.

Aggarwal, S. 2018. Do rural roads create pathways out of poverty? Evidence from India. *Journal of Development Economics* 133:375–95.

Aghion, P., and P. Bolton. 1997. A theory of trickle-down growth and development. *Review of Economic Studies* 64:151–72.

Allen, F., E. Carletti, R. Cull, J. Qian, L. Senbet, and P. Valenzuela. 2021. Improving access to banking: Evidence from Kenya. *Review of Finance* 25:403–47.

Artadi, E., and X. Sala-I-Martin. 2004. The economic tragedy of the twentieth century: Growth in Africa. The Africa Competitiveness Report 2004.

Asher, S., and P. Novosad. 2017. Market access and structural transformation: Evidence from rural roads in India. Working Paper, Imperial College.

———. 2020. Rural roads and local economic development. *American Economic Review* 110:797–823.

Aterido, R., T. Beck, and L. Iacovone. 2013. Gender and finance in Sub-Saharan Africa: Are women disadvantaged? *World Development* 47:102–120.

Banerjee, A., E. Duflo, and N. Qian. 2020. On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics* 145:10.1016/j.jdeveco.2020.102442.

Basu, P. 2006. Improving access to finance for India's rural poor. Washington, DC: World Bank Publications.

Baum-Snow, N., L. Brandt, J. V. Henderson, M. A. Turner, and Q. Zhang. 2017. Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics* 99:435–48.

Beck, T. 2008. The econometrics of finance and growth. In *Palgrave handbook of econometrics*, vol. 2. New York: Springer.

———. 2012. The role of finance in economic development — Benefits, risks, and politics. In *Oxford handbook of capitalism*, ed. Dennis Müller. Oxford, UK: Oxford University Press.

Beck, T., and A. Demirguc-Kunt. 2008. Access to finance - An unfinished agenda. *World Bank Economic Review* 22:383–96.

Beck, T., and A. Demirguc-Kunt, and R. Levine. 2007. Finance, inequality, and the poor. *Journal of Economic Growth* 12:27–49.

Beck, T., A. Demirguc-Kunt, and M. Martinez Peria. 2007. Reaching out: Access to and use of banking services across countries. *Journal of Financial Economics* 85:234–66.

Beck, T. L. Lu, and R. Yang. 2014. Finance and growth for microenterprises: Evidence from rural China. *World Development* 67:38–56.

Bernhardt, A, E. Field, R. Pande, and N. Rigol, 2017. Household matters: Revisiting the returns to capital among female micro-entrepreneurs. Working Paper.

Bertrand, M, E. Duflo, and S. Mullainathan. 2004. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 119:249–75.

Black, S. E., and P. E. Strahan. 2002. Entrepreneurship and bank credit availability. *Journal of Finance* 57:2807–33.

Blattman, C, N. Fiala, and S. Martinez, 2013. Generating skilled self employment in developing countries: Experimental evidence from Uganda. *Quarterly Journal of Economics* 129: 697–752.

Brown, J. R., J. A. Cookson, and R. Z.Heimer. 2019. Growing up without finance. *Journal of Financial Economics* 134:591–616.

Burgess, R., and D. Donaldson. 2012. Railroads and the demise of famine in colonial India. Working Paper, London School of Economics.

Burgess, R., and R. Pande. 2004. Do rural banks matter? Evidence from the Indian social banking experiment. *American Economic Review* 95:780–95.

Calderon, C., and L. Serven. 2004. The effects of infrastructure development on growth and income distribution. Washington, DC: World Bank Publications.

Chaurey, R. and D.T. Le. 2022. Infrastructure maintenance and rural economic activity: Evidence from India. *Journal of Public Economics* 214: 104725.

Citizen. 2018. JPM promises to support banks' rural expansion. April 23. https://www.thecitizen.co.tz/News/1840340-4495330-2muo81z/index.html

Coffey, E. 1998. Agricultural finance: Getting the policies right. *Agricultural Finance* Revisited: vol. 2, FAO and GTZ.

Conning, J., and C. Udry. 2005. Rural financial markets in developing countries. In *The handbook of agricultural economics, agricultural development: Farmers, farm production and farm markets*, eds. R. E. Evenson, P. Pingali, and T. P. Schultz, vol. 3. Amsterdam, the Netherlands: North Holland.

Cornell, B., and I. Welch. 1996. Culture, information, and screening discrimination. *Journal of Political Economy* 104:542–71.

Das, A., E. Ghani, A. Grover, W. R. Kerr, and R. Nanda. 2019. Infrastructure and finance: Evidence from India's GQ highway network. Working Paper, World Bank.

Demirguc-Kunt, A., and R. Levine. 2008a. Finance, financial sector policies, and long-run growth. Washington, DC: World Bank Publications.

———. 2008b. Finance and economic opportunity. *Proceedings of the Annual World Bank Conference on Development Economics*.

———. 2009. Finance and inequality: Theory and evidence. *Annual Review of Financial Economics* 1:287–318.

Demirguc-Kunt, A., E. Feijen, and R. Levine. 2013. The evolving importance of banks and markets in economic development. *World Bank Economic Review* 27:476–90. District Plan Report. 2010–2011. GANJAM District, ORISSA.

Donaldson, D. 2018. Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review* 108:899–934.

Donaldson, D., and R. Hornbeck. 2016. Railroads and american economic growth: A "market access" approach. *Quarterly Journal of Economics* 131:799–858.

Drabenstott, M., N. Novack and B. Abraham. 2003. Main streets of tomorrow: growing and financing rural entrepreneurs – a conference summary. *Economic Review – Federal Reserve Bank of Kansas City* 88(3): 73–84.

Faber, B. 2014. Trade integration, market size, and industrialization: Evidence from China's National Trunk Highway System. *Review of Economic Studies* 81:1046–70.

Field, E., S. Jayachandran, and R. Pande. 2010. Do traditional institutions constrain female entrepreneurship? A field experiment on business training in India. *American Economic Review* 100:125–29.

Fisman, R., D. Paravisini, and V. Vig. 2017. Cultural proximity and loan outcomes. *American Economic Review* 107:457–92.

Gelbach, J. 2002. Public schooling for young children and maternal labor supply. *American Economic Review* 92(1): 307-322.

Ghani, E., W. Kerr, and S. O'Connell. 2014. Political Reservations and Women's Entrepreneurship in India. *Journal of Development Economics* 108: 138–153.

Ghani, E., A. G. Goswami, and W. R. Kerr. 2016. Highway to success: The impact of the golden quadrilateral project for the location and performance of Indian manufacturing. *Economic Journal* 126:317–57.

———. 2017. Highways and spatial location within cities: Evidence from India. *World Bank Economic Review* 30:S97–S108.

Ghani, E., A.G. Goswami, S. Kerr, and W. Kerr. 2016. Will market competition trump gender discrimination in India? Working paper, World Bank.

Goldin, C. and Olivetti, C., 2013. Shocking labor supply: A reassessment of the role of World War II on women's labor supply. *American Economic Review* 103(3): 257–262.

Imbens, G., and T. Lemieux, 2008.Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142:615–35.

The World Bank. 2022. Improving Road Transport to Promote Sustainable Regional Development in Tocantins, Brazil. *The World Bank Results Brief*, May 2, 2022.

Jayachandran, S., 2020. Microentrepreneurship in Developing Countries. Working Paper.

Ji, Y., S. Teng, and R. Townsend. 2021. Branch expansion versus digital banking: The dynamics of growth and inequality in a spatial equilibrium model. Working Paper, National Bureau of Economic Research.

Kerr, W. and R. Nanda. 2009. Democratizing Entry: Banking Deregulations, Financing Constraints, and Entrepreneurship. *Journal of Financial Economics* 94: 124–149.

Khandker, S. R., Z. Bakht, and G. B. Koolwal. 2009. The poverty impact of rural roads: Evidence from Bangladesh. *Economic Development and Cultural Change* 57:685–722.

Kimmel, J., 1998. Child care costs as a barrier to employment for single and married mothers. *Review of Economics and Statistics* 80(2):287-299.

King, R. G., and R. Levine. 1993a. Finance, entrepreneurship and growth: Theory and evidence. *Journal of Monetary Economics* 32:513–42.

———. 1993b. Finance and growth: Schumpeter might be right. *Quarterly Journal of Economics* 108:717–37.

Lerner, J. and A. Schoar. 2010. Introduction to International Differences in Entrepreneurship. *University of Chicago Press* pp. 1–13.

Levine, R. 1997. Financial development and economic growth: Views and agenda. *Journal of Economic Literature* 35:688–726.

———. 2005. Finance and growth: Theory and evidence. In *Handbook of economic growth*, eds. P. Aghion and S. Durlauf. Amsterdam, the Netherlands: Elsevier.

———. 2008. Finance, growth, and the poor, the financial development report 2008. Cologny, Switzerland: World Economic Forum.

McElwee, G. and R. Smith, 2014. Rural entrepreneurship, in A. Fayolle ed., Handbook of research on entrepreneurship (Cheltenham: Edward Elgar), Pp. 423–470.

McCrary, J. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142:698–714.

Morten, M., and J. Oliveira. 2014. Migration, roads and labor market integration: Evidence from a planned capital city. Working Paper, Stanford University.

Mukherjee, M. 2012. Do better roads increase school enrollment? Evidence from a unique road policy in India. Working Paper, Syracuse University.

Mukherjee, A. and X. Zhang. 2007. Rural industrialization in China and India: role of policies and institutions. *World Development* 35 (10) pp. 1621–1634

Naaraayanan, S. L., and D. Wolfenzon. 2023, Business Group Spillovers. *Review of Financial Studies*, Forthcoming.

Naaraayanan, S. L., 2022. Women's Inheritance Rights and Entrepreneurship Gender Gap. Working Paper.

Petrin, T. and A. Gannon. 1997. Rural development through entrepreneurship. Rome: FAO.

PriceWaterhouse Coopers. 2010. Foreign banks in China. Report. https://www.pwc.com/th/en/publications/assets/pdf-foreign-banks-in-china2010.pdf

Rebello, J. 2013. New private banks upbeat on rural expansion. LiveMint, August 29. https://www.livemint.com/Industry/jY4qrzLL14GoR9qvMsFgXL/New-private-banks-upbeat-on-rural-expansion.html

Revenga, A. and M. Dooley. 2020. What works for women microentrepreneurs?. *Brookings Institution.*

Rijkers, B. and R. Costa. 2012. Gender and non-farm entrepreneurship. *World Development* 40 (12) pp. 2411–2426.

Satish, P. 2004. Rural finance: Role of state and state-owned institutions. *Economic and Political Weekly* 39:1321—2, 1324, 1326–8, 1330.

Schoar, A. 2010. The divide between subsistence and transformational entrepreneurship. *Innovation Policy and the Economy* 10(1), 57–81.

Schumpeter, J. 1911. The theory of economic development. Cambridge, MA: Harvard University Press.

Shamdasani, Y. 2021. Rural road infrastructure and agricultural production: Evidence from India. *Journal of Development Economics* 152:10.1016/j.jdeveco.2021.102686.

Shankar, N. 2010. McKinsey to tutor SBI in rural banking for Rs. 62.8 crore. Rupee Times, May 5.

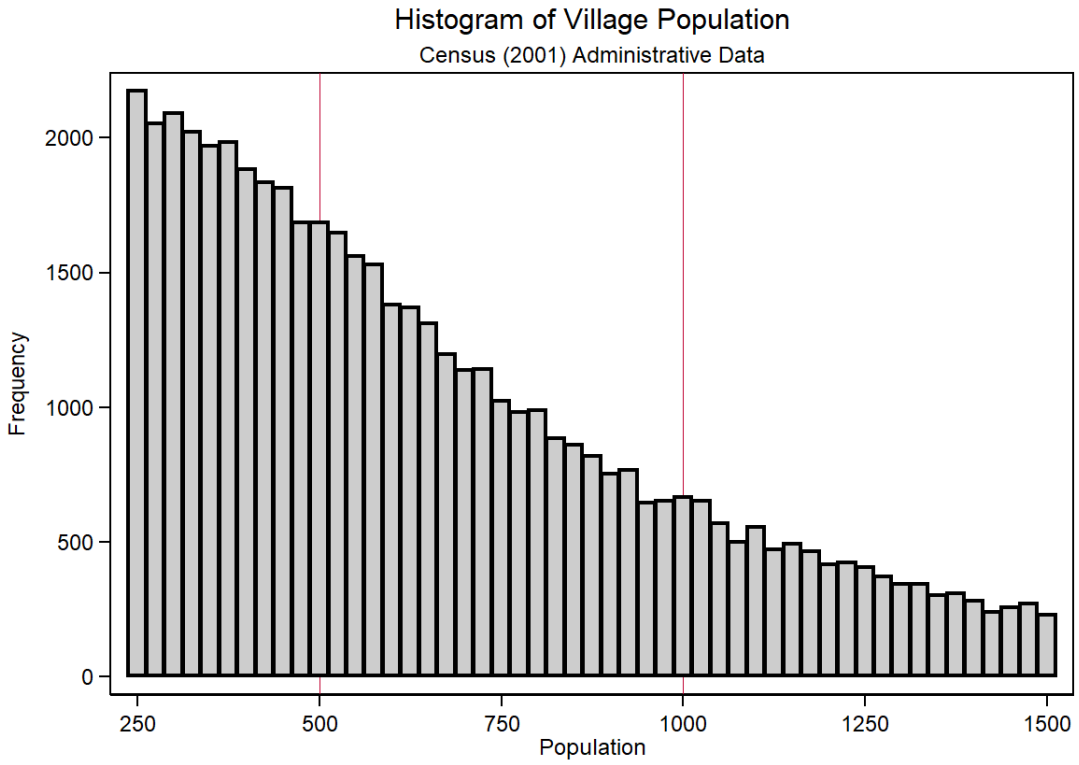
Storeygard, A. 2016. Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa. *Review of Economic Studies* 83:1263–95.

Vig, V. 2013. Access to collateral and corporate debt structure: Evidence from a natural experiment. *Journal of Finance* 68:881–928.

Visaria, S. 2009. Legal reform and loan repayment: The microeconomic impact of debt recovery tribunals in India. *American Economic Journal: Applied Economics* 1:59–81.

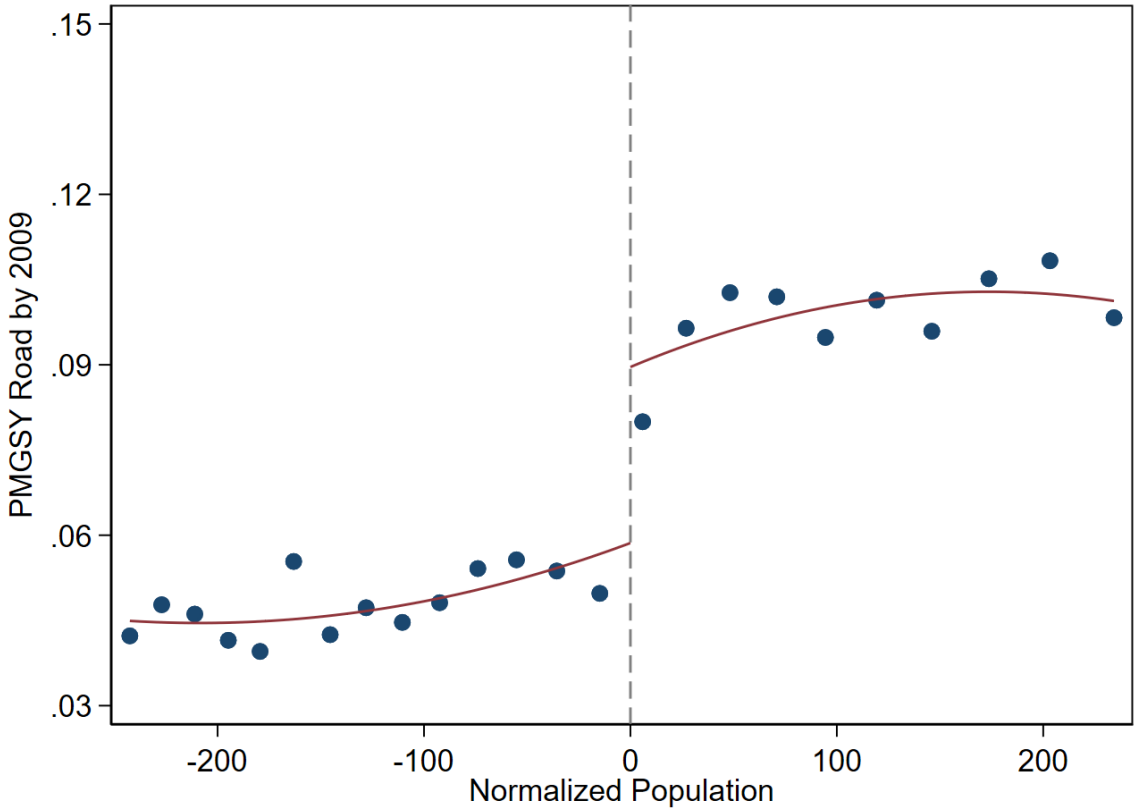
World Bank. 2007. A decade of action in transport: An evaluation of world bank assistance to the transport sector, 1995–2005. Washington, DC.

Figure 1: Distribution of running variable



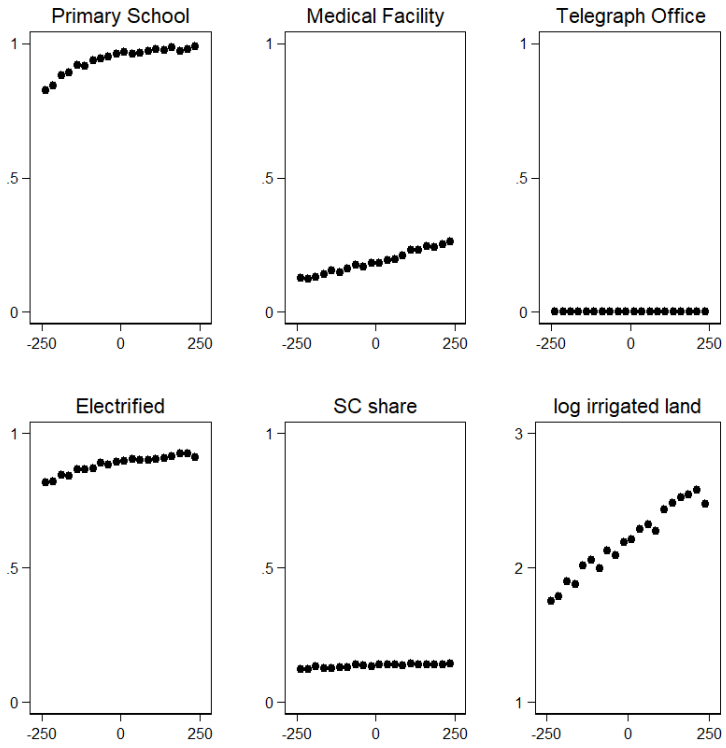
Notes: The figure shows the distribution of village population around the different population thresholds as outlined under PMGSY guidelines. We present the histogram of village population as recorded in the 2001 Population Census. The vertical lines depict the program eligibility cutoffs as defined in PMGSY at 500 and 1000. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of villages that did not have paved roads at the start of our sample. We also test for the discontinuity at 500 and 1000 thresholds by normalizing over village population as recorded in the 2001 Population Census, following McCrary (2008), with an estimate of 0.064 and the standard error of 0.055.

Figure 2: First stage: effect of road prioritization on probability of PMGSY road by 2009



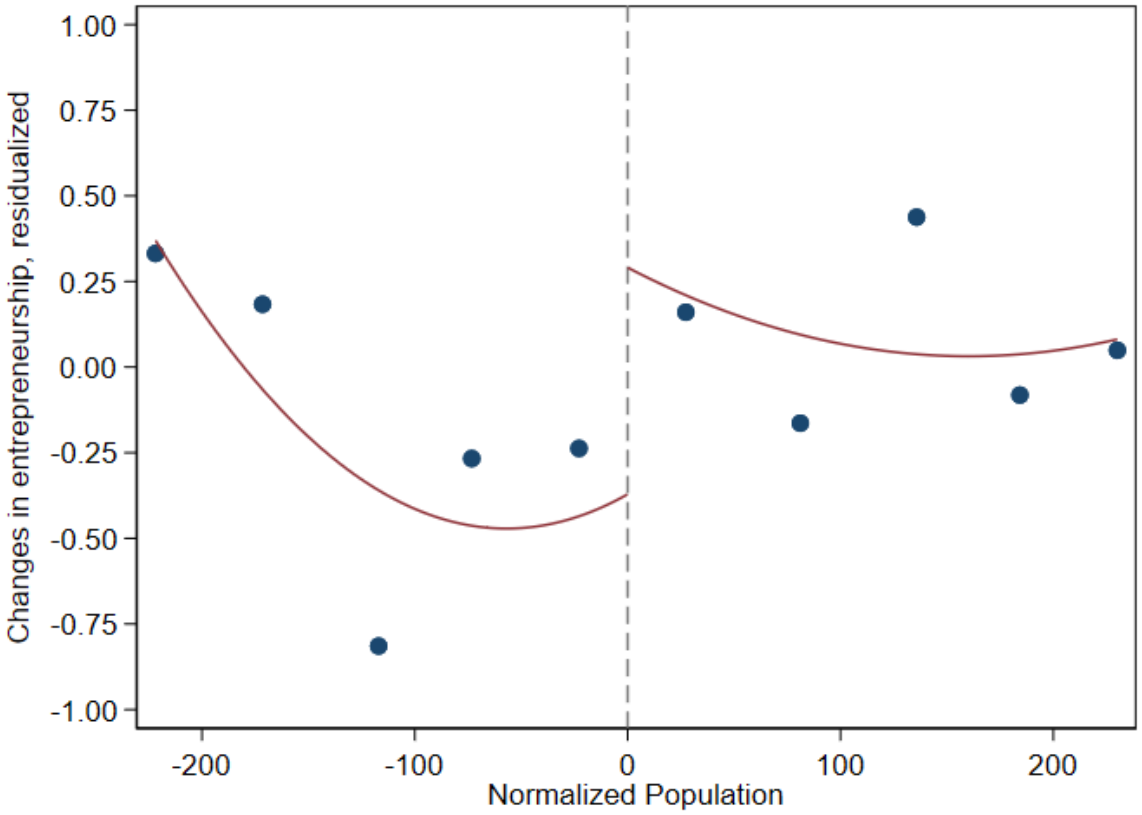
Notes: The figure plots the probability of receiving a road under a PMGSY by 2009 over village population as recorded in the 2001 Population Census. We also fit a quadratic polynomial on either sides of the cutoff. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of villages that did not have paved roads at the start of our sample.

Figure 3: Balance of baseline village characteristics



Notes: The figure plots means of baseline village characteristics over normalized popuation. Points to the right of zero are above treatment thresholds, while points to the left of zero are below treatment thresholds. The bin width is 25 on either side of the threshold and each point represents approximately fifteen observations. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of villages that did not have paved roads at the start of our sample.

Figure 4: Second stage: effect of new rural roads on changes in entrepreneurship



Notes: The figure plots the effect of new rural roads on changes in rural entrepreneurship. We residualize changes in entrepreneurship by adding district fixed effects to account for time-invariant district characteristics and fit a quadratic polynomial on either sides of the cutoff. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of villages that did not have paved roads at the start of our sample. Population as recorded in the 2001 Population Census.

Table 1: Covariate balance

The table presents mean values for baseline village characteristics, as recorded in 2001 Population Census. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. We use a bandwidth of ±125 around the population thresholds to define the sample of villages. Panel A reports balance for all unconnected villages while panel B reports balance for MSME sample. In both panels, column 1 presents the respective unconditional sample means. Columns 2 and 3 present the unconditional means for villages below the treatment threshold, and villages above the treatment threshold, respectively. Column 4 presents the difference in means between villages below the treatment threshold and villages above the treatment threshold. Column 5 reports the regression discontinuity estimate, following the main estimating equation, of the effect of being above the treatment threshold on the baseline variable and column 6 is the p-value for this estimate, using heteroskedasticity-robust standard errors.

	Panel A: Full sample					
	All	Below Threshold	Above Threshold	Difference	RD estimate	p-value on estimate
	(1)	(2)	(3)	(4)	(5)	(6)
Primary school	0.960	0.949	0.971	-0.022	-0.001	0.814
Medical facility	0.229	0.209	0.253	-0.044	-0.007	0.390
Telegraph office	0.003	0.002	0.003	-0.001	0.001	0.439
Electricity	0.902	0.891	0.913	-0.022	0.001	0.821
Log (irrigated land)	2.506	2.413	2.609	-0.196	-0.083	0.041**
Log (distance nearest town)	3.090	3.098	3.080	0.017	0.003	0.868
Scheduled caste (share)	0.134	0.133	0.135	-0.002	-0.002	0.613

	Panel B: MSME sample					
	All	Below Threshold	Above Threshold	Difference	RD estimate	p-value on estimate
	(1)	(2)	(3)	(4)	(5)	(6)
Primary school	0.964	0.956	0.975	-0.019	-0.041	0.640
Medical facility	0.292	0.281	0.305	-0.024	0.001	0.999
Telegraph office	0.004	0.007	0.000	0.007	0.023	0.543
Electricity	0.949	0.941	0.958	-0.017	0.079	0.389
Log (irrigated land)	2.872	2.790	2.967	-0.177	0.341	0.620
Log (distance nearest town)	2.950	2.971	2.925	0.046	-0.078	0.782
Scheduled caste (share)	0.136	0.143	0.127	0.016	-0.031	0.648

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: First stage effect of road priority on PMGSY road treatment

The table presents first stage estimates from Equation 1 of the effect of being above the population threshold on a village's probability of receiving a road under PMGSY by 2009. The dependent variable is an indicator variable that takes on the value one if a village has received a PMGSY road before 2009. The first (and sixth) column presents results for villages with populations within 80 of the population threshold (420-580 for the 500 threshold and 920-1080 for the 1000 threshold). Other columns expand the sample to include villages within 90, 100, 110, and 120 of the population thresholds, respectively. Columns 1 through 5 include polynomial of degree two while columns 6 through 10 include polynomial of degree 3. The specification in panel B includes baseline village-level controls for amenities and economic indicators, as well as district-cutoff fixed effects. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. Heteroskedasticity-robust standard errors are reported below point estimates.

	Panel A: Without controls									
Bandwidth	±80	±90	±100	±110	±120	±80	±90	±100	±110	±120
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above cutoff	0.030***	0.032***	0.032***	0.032***	0.031***	0.026***	0.026***	0.027***	0.028***	0.031***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
First stage (F-statistic)	41.6	53.5	60.8	67.6	66.3	18.6	21.1	24.7	29.9	38.6
Controls	No	No	No	No	No	No	No	No	No	No
District × Threshold f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial degree	2	2	2	2	2	3	3	3	3	3
R^2	0.040	0.038	0.036	0.036	0.036	0.040	0.038	0.036	0.036	0.036
Observations	40106	45070	50009	54991	60001	40106	45070	50009	54991	60001

	Panel B: With controls									
Bandwidth	±80	±90	±100	±110	±120	±80	±90	±100	±110	±120
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above cutoff	0.030***	0.032***	0.033***	0.033***	0.031***	0.027***	0.027***	0.028***	0.029***	0.031***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
First stage (F-statistic)	43.1	54.9	62.4	69.7	68.3	19.6	22.2	25.8	31.0	40.0
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial degree	2	2	2	2	2	3	3	3	3	3
R^2	0.044	0.043	0.041	0.041	0.040	0.045	0.043	0.041	0.041	0.040
Observations	40106	45070	50009	54991	60001	40106	45070	50009	54991	60001

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Summary statistics

The table reports the distribution of business characteristics for our matched dataset. We report the number of new business per capita (village population) and the number of new female businesses before PMGSY. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census.

	Mean	Std. dev.	p10	p25	Median	p75
	(1)	(2)	(3)	(4)	(5)	(6)
New business per capita	4	4	1	1.64	2.88	4.78
New female business (%)	14.8	35.5	0.0	0.0	0.0	0.0
Firm age (years)	7.4	4.8	2.7	4.0	6.1	9.5
Asset size ($\times 10^5$ Rs.)	0.24	0.52	0.034	0.06	0.14	0.28
Decline in gross output	0.007	0.058	0.0	0.0	0.0	0.0
Decline in net worth	0.17	.32	0.0	0.0	0.0	0.0
Employment	1.4	.69	1.0	1.0	1.0	1.8
Female employment	0.24	0.41	0.0	0.0	0.0	0.37

Table 4: New rural roads and change in entrepreneurial activity

The table presents reduced form estimates of the effect of new rural roads on the change in entrepreneurial activity within the villages. The dependent variable is the change in number of unregistered businesses within a village between 2000 and 2007, scaled by the village population as recorded in the 2001 Population Census. The first (and sixth) column presents results for villages with populations within 80 of the population threshold (420-580 for the 500 threshold and 920-1080 for the 1000 threshold). Other columns expand the sample to include villages within 90, 100, 110, and 120 of the population thresholds, respectively. Columns 1 through 5 include polynomial of degree two while columns 6 through 10 include polynomial of degree 3. Panel B includes baseline village-level controls for amenities and economic indicators. All specifications use district-cutoff fixed effects. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. The standard errors are robust to heteroscedasticity. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

	Panel A: Without controls									
Bandwidth	±80	±90	±100	±110	±120	±80	±90	±100	±110	±120
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above cutoff	0.006*	0.006**	0.005*	0.005*	0.004*	0.007*	0.006*	0.009**	0.006*	0.007**
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
Controls	No	No	No	No	No	No	No	No	No	No
District × Threshold f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial degree	2	2	2	2	2	3	3	3	3	3
R^2	0.43	0.38	0.35	0.33	0.32	0.44	0.38	0.39	0.34	0.33
Observations	89	109	130	146	165	89	109	130	146	165

	Panel B: With controls									
Bandwidth	±80	±90	±100	±110	±120	±80	±90	±100	±110	±120
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above cutoff	0.005	0.005*	0.004*	0.004*	0.004*	0.007	0.006	0.008**	0.005	0.006*
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial degree	2	2	2	2	2	3	3	3	3	3
R^2	0.58	0.47	0.47	0.39	0.37	0.59	0.48	0.51	0.40	0.38
Observations	89	109	130	146	165	89	109	130	146	165

Table 5: New rural roads and entrepreneurial activity, High vs. Low rainfall districts

The table presents reduced form estimates of the effect of new rural roads on the change in entrepreneurial activity within the villages. The dependent variable is the change in number of unregistered businesses within a village between 2000 and 2007, scaled by the village population as recorded in the 2001 Population Census. We interact *Above cutoff* with *High rainfall*, which takes the value 1 if the average monthly rainfall in June, July, and August within each district is in the top 10 percentile in the period before the survey was conducted. The first (and sixth) column presents results for villages with populations within 80 of the population threshold (420-580 for the 500 threshold and 920-1080 for the 1000 threshold). Other columns expand the sample to include villages within 90, 100, 110, and 120 of the population thresholds, respectively. In both panels, columns 1 through 3 includes polynomial of degree two while columns 4 through 6 includes polynomial of degree 3. The specification in panel B includes baseline village-level controls for amenities and economic indicators, as well as district-cutoff fixed effects. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. The standard errors are robust to heteroscedasticity. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

	Panel A: Without controls									
Bandwidth	±80	±90	±100	±110	±120	±80	±90	±100	±110	±120
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above cutoff	0.005	0.006*	0.005*	0.004*	0.004*	0.007	0.006*	0.009**	0.007*	0.007**
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
Above cutoff × High rainfall	0.003**	0.002**	0.004	0.008**	0.006**	0.004**	0.003	0.004**	0.008**	0.007**
	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial degree	2	2	2	2	2	3	3	3	3	3
R^2	0.44	0.37	0.34	0.33	0.32	0.44	0.38	0.39	0.34	0.34
Observations	89	109	130	146	165	89	109	130	146	165

	Panel B: With controls									
Bandwidth	±80	±90	±100	±110	±120	±80	±90	±100	±110	±120
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above cutoff	0.004	0.004*	0.004*	0.004*	0.003	0.007	0.006	0.008**	0.005	0.006*
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
Above cutoff × High rainfall	0.003**	0.002*	0.004*	0.008***	0.007**	0.003*	0.002	0.005**	0.008**	0.007**
	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial degree	2	2	2	2	2	3	3	3	3	3
R^2	0.59	0.47	0.47	0.39	0.37	0.60	0.47	0.51	0.40	0.38
Observations	89	109	130	146	165	89	109	130	146	165

Table 6: New rural roads and entrepreneurial activity, food and beverages sector

The table presents reduced form estimates of the effect of new rural roads on the change in entrepreneurial activity within the villages. The dependent variable is the change in number of unregistered businesses operating in the food and beverages sector within a village between 2000 and 2007, scaled by the village population as recorded in the 2001 Population Census. The first (and sixth) column presents results for villages with populations within 80 of the population threshold (420-580 for the 500 threshold and 920-1080 for the 1000 threshold). Other columns expand the sample to include villages within 90, 100, 110, and 120 of the population thresholds, respectively. In both panels, columns 1 through 3 includes polynomial of degree two while columns 4 through 6 includes polynomial of degree 3. The specification in panel B includes baseline village-level controls for amenities and economic indicators. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. The standard errors are robust to heteroscedasticity. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

	Panel A: Without controls									
Bandwidth	±80	±90	±100	±110	±120	±80	±90	±100	±110	±120
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above cutoff	0.001**	0.001	0.001*	0.001	0.001	0.001	0.001	0.001	0.001*	0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	No	No	No	No	No	No	No	No	No	No
District × Threshold f.e.	No	No	No	No	No	No	No	No	No	No
Polynomial degree	2	2	2	2	2	3	3	3	3	3
R^2	0.15	0.18	0.21	0.21	0.099	0.19	0.20	0.22	0.23	0.13
Observations	41	46	51	55	61	41	46	51	55	61

	Panel B: With controls									
Bandwidth	±80	±90	±100	±110	±120	±80	±90	±100	±110	±120
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above cutoff	0.001*	0.001	0.001*	0.001	0.001	0.000	0.001	0.001	0.001*	0.001*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District × Threshold f.e.	No	No	No	No	No	No	No	No	No	No
Polynomial degree	2	2	2	2	2	3	3	3	3	3
R^2	0.17	0.18	0.22	0.21	0.12	0.22	0.20	0.22	0.23	0.15
Observations	41	46	51	55	61	41	46	51	55	61

Table 7: Performance of new businesses

The table presents reduced form estimates of the effect of new rural roads on the performance of new businesses within the villages. In panel A, the dependent variable is the change in the fraction of businesses suffering a decline in net worth within a village between 2000 and 2007. In panel B, the dependent variable is the change in the fraction of businesses suffering a decline in gross output. All columns present results for villages with populations within 100 of the population threshold (400-600 for the 500 threshold and 900-1100 for the 1000 threshold). In all panels, columns 1 and 2 include polynomial of degree two while columns 3 and 4 include polynomial of degree 3. All specifications include district-cutoff fixed effects while specifications in columns 2 and 4 additionally include baseline village-level controls for amenities and economic indicators. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. The standard errors are robust to heteroscedasticity. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

Panel A: Change in the fraction of businesses suffering a net worth decline

	(1)	(2)	(3)	(4)
Above cutoff	-0.123	-0.117	-0.160	-0.118
	(0.190)	(0.195)	(0.250)	(0.259)
Controls	No	Yes	No	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes
Polynomial degree	2	2	3	3
R^2	0.42	0.48	0.45	0.50
Observations	130	130	130	130

Panel B: Change in the fraction of businesses suffering an output decline

	(1)	(2)	(3)	(4)
Above cutoff	0.067	0.081	0.046	0.058
	(0.072)	(0.083)	(0.058)	(0.069)
Controls	No	Yes	No	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes
Polynomial degree	2	2	3	3
R^2	0.30	0.34	0.31	0.35
Observations	130	130	130	130

Table 8: Who starts new businesses, by gender

The table presents reduced form estimates of the effect of new rural roads on the change in business characteristics within the villages. In panel A (panel B), the dependent variable is the change in female (male) entrepreneurship within a village between 2000 and 2007, scaled by the village population as recorded in the 2001 Population Census. All columns present results for villages with populations within 100 of the population threshold (400-600 for the 500 threshold and 900-1100 for the 1000 threshold). In all panels, columns 1 and 2 include polynomial of degree two while columns 3 and 4 include polynomial of degree 3. All specifications include district-cutoff fixed effects while specifications in columns 2 and 4 additionally include baseline village-level controls for amenities and economic indicators. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. The standard errors are robust to heteroscedasticity. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

	Panel A: Change in female entrepreneurship rates			
	(1)	(2)	(3)	(4)
Above cutoff	0.002*	0.002*	0.003**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
Controls	Yes	Yes	No	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes
Polynomial degree	2	2	3	3
R^2	0.49	0.50	0.50	0.51
Observations	130	130	130	130

	Panel B: Change in male entrepreneurship rates			
	(1)	(2)	(3)	(4)
Above cutoff	0.003	0.002	0.007	0.005
	(0.003)	(0.002)	(0.004)	(0.004)
Controls	Yes	Yes	No	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes
Polynomial degree	2	2	3	3
R^2	0.31	0.45	0.35	0.48
Observations	130	130	130	130

Table 9: What explains the increase in female entrepreneurial activity?

The table presents reduced form estimates of bank financing on the changes in female entrepreneurial activity within the villages. The dependent variable is the change in female entrepreneurship rates within a village between 2000 and 2007. We interact *Above cutoff* with, *Log(rural credit)*, defined as the logarithm of rural credit within a district (Panel A), $\Delta\%(Male - Female)$, defined as the net change in migration from agricultural employment between males and females (Panel B), *School enrolment (per capita)*, defined as the total school enrolment in the period before the survey, scaled by the village population as recorded in the 2001 Population Census (Panel C). All columns present results for villages with populations within 100 of the population threshold (400-600 for the 500 threshold and 900-1100 for the 1000 threshold). Columns 1 and 2 include polynomial of degree two while columns 3 and 4 include polynomial of degree 3. The specification in columns 2 and 4 includes baseline village-level controls for amenities and economic indicators, as well as district-cutoff fixed effects. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. The standard errors are robust to heteroscedasticity. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

	Panel A: Access to finance			
	(1)	(2)	(3)	(4)
Above cutoff	0.675	0.611	1.136**	1.090**
	(0.440)	(0.404)	(0.490)	(0.481)
Above cutoff × Log(rural credit)	-0.010	-0.005	-0.015	-0.012
	(0.019)	(0.018)	(0.018)	(0.018)
Controls	No	Yes	No	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes
Polynomial	2	2	3	3
R^2	0.51	0.53	0.53	0.55
Observations	130	130	130	130

	Panel B: Male outmigration			
	(1)	(2)	(3)	(4)
Above cutoff	0.283	0.276	0.543***	0.563***
	(0.171)	(0.177)	(0.195)	(0.202)
Above cutoff × Δ % (Male - Female)	0.023**	0.024**	0.020*	0.020*
	(0.011)	(0.011)	(0.011)	(0.011)
Controls	No	Yes	No	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes
Polynomial degree	2	2	3	3
R^2	0.48	0.49	0.52	0.52
Observations	130	130	130	130

	Panel C: Availability of schooling			
	(1)	(2)	(3)	(4)
Above cutoff	0.242	0.236	0.532***	0.561***
	(0.175)	(0.181)	(0.192)	(0.197)
Above cutoff × School enrolment (per capita)	1.700**	1.768**	2.305**	2.507**
	(0.762)	(0.818)	(0.966)	(0.962)
Controls	No	Yes	No	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes
Polynomial degree	2	2	3	3
R^2	0.48	0.49	0.53	0.54
Observations	130	130	130	130

Table 10: New rural roads and change in new business employment

The table presents reduced form estimates of the effect of new rural roads on the change in employment within the villages. In Panel A (Panel B), the dependent variable is the change total (female) employment in a village. All columns present results for villages with populations within 100 of the population threshold (400-600) for the 500 threshold and 900-1100 for the 1000 threshold). Columns 1 and 2 include polynomial of degree two while columns 3 and 4 include polynomial of degree 3. The specification in columns 2 and 4 includes baseline village-level controls for amenities and economic indicators, as well as district-cutoff fixed effects. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. The standard errors are robust to heteroscedasticity. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

	Panel A: Change in overall employment			
	(1)	(2)	(3)	(4)
Above cutoff	0.049	0.057	0.129	0.106
	(0.107)	(0.110)	(0.128)	(0.134)
Controls	No	Yes	No	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes
Polynomial degree	2	2	3	3
R^2	0.53	0.56	0.54	0.56
Observations	130	130	130	130

	Panel B: Change in female employment			
	(1)	(2)	(3)	(4)
Above cutoff	0.333**	0.327*	0.661***	0.733***
	(0.164)	(0.176)	(0.186)	(0.176)
Controls	No	Yes	No	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes
Polynomial degree	2	2	3	3
R^2	0.48	0.52	0.53	0.59
Observations	130	130	130	130

Rural roads, small business creation, and the gender-gap in entrepreneurship

INTERNET APPENDIX

FOR ONLINE PUBLICATION

Table IA1: New rural roads and entrepreneurial activity, Services sector

The table presents reduced form estimates of the effect of new rural roads on the change in entrepreneurial activity within the villages. The dependent variable is the change in number of unregistered businesses within a village between 2000 and 2007, scaled by the village population as recorded in the 2001 Population Census. The first (and sixth) column presents results for villages with populations within 80 of the population threshold (420-580 for the 500 threshold and 920-1080 for the 1000 threshold). Other columns expand the sample to include villages within 90, 100, 110, and 120 of the population thresholds, respectively. Columns 1 through 5 include polynomial of degree two while columns 6 through 10 include polynomial of degree 3. Panel B includes baseline village-level controls for amenities and economic indicators. All specifications use district-cutoff fixed effects. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. The standard errors are robust to heteroscedasticity. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

	Panel A: Without controls									
Bandwidth	±80	±90	±100	±110	±120	±80	±90	±100	±110	±120
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above cutoff	-0.000	0.000	-0.000	-0.000	-0.000	-0.002	-0.002	-0.001	-0.000	-0.001
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Controls	No	No	No	No	No	No	No	No	No	No
District × Threshold f.e.	No	No	No	No	No	No	No	No	No	No
Polynomial degree	2	2	2	2	2	3	3	3	3	3
R^2	0.0095	0.027	0.036	0.036	0.036	0.022	0.044	0.038	0.040	0.038
Observations	68	75	82	86	95	68	75	82	86	95

	Panel B: With controls									
Bandwidth	±80	±90	±100	±110	±120	±80	±90	±100	±110	±120
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above cutoff	-0.001	0.000	-0.000	-0.000	0.000	-0.001	-0.002	-0.001	-0.000	-0.001
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District × Threshold f.e.	No	No	No	No	No	No	No	No	No	No
Polynomial degree	2	2	2	2	2	3	3	3	3	3
R^2	0.051	0.065	0.062	0.060	0.048	0.055	0.078	0.067	0.064	0.051
Observations	68	75	82	86	95	68	75	82	86	95

Table IA2: Who starts new businesses, by asset size

The table presents reduced form estimates of the effect of new rural roads on the change in business characteristics within the villages. The dependent variable is the change in the average asset size of a new unregistered business within a village between 2000 and 2007, scaled by the village population as recorded in the 2001 Population Census. All columns present results for villages with populations within 100 of the population threshold (400-600 for the 500 threshold and 900-1100 for the 1000 threshold). Columns 1 and 2 include polynomial of degree two while columns 3 and 4 include polynomial of degree 3. All specifications include district-cutoff fixed effects while specifications in columns 2 and 4 additionally include baseline village-level controls for amenities and economic indicators. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. The standard errors are robust to heteroscedasticity. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

	(1)	(2)	(3)	(4)
Above cutoff	-0.910	-0.840	-1.040	-1.143*
	(0.549)	(0.568)	(0.640)	(0.677)
Controls	No	Yes	No	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes
Polynomial degree	3	3	3	3
R^2	0.37	0.42	0.37	0.42
Observations	130	130	130	130

Table IA3: Who starts new businesses, by minorities

The table presents reduced form estimates of the effect of new rural roads on the change in business characteristics within the villages. In Panel A (Panel B), the dependent variable is the change in the number of unregistered SC/ST/OBC (others) businesses within a village between 2000 and 2007, scaled by the village population as recorded in the 2001 Population Census. All columns present results for villages with populations within 100 of the population threshold (400-600 for the 500 threshold and 900-1100 for the 1000 threshold). In all panels, columns 1 and 2 include polynomial of degree two while columns 3 and 4 include polynomial of degree 3. All specifications include district-cutoff fixed effects while specifications in columns 2 and 4 additionally include baseline village-level controls for amenities and economic indicators. Our sample consists of villages from the following states which followed the population threshold prioritization rules as given by the national guidelines of the PMGSY: Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Odisha, and Rajasthan. The sample consists of all the villages that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. The standard errors are robust to heteroscedasticity. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

	Panel A: Change in the fraction of minority businesses			
	(1)	(2)	(3)	(4)
Above cutoff	0.003	0.002	0.008*	0.007*
	(0.002)	(0.002)	(0.004)	(0.004)
Controls	No	Yes	No	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes
Polynomial degree	2	2	3	3
R^2	0.31	0.41	0.37	0.46
Observations	130	130	130	130

	Panel B: Change in the fraction of other businesses			
	(1)	(2)	(3)	(4)
Above cutoff	0.002*	0.002**	0.002	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Controls	No	Yes	No	Yes
District × Threshold f.e.	Yes	Yes	Yes	Yes
Polynomial degree	2	2	3	3
R^2	0.55	0.62	0.55	0.62
Observations	130	130	130	130