Do Rural Roads Promote Inclusive Entrepreneurship?*

Ananyo Brahma† Vidhya Soundararajan‡

August 31, 2024

Abstract

Achieving inclusivity in entrepreneurship has been challenging. Using data from firms in India, we examine how a national rural road construction program, which connected previously isolated villages, affects entrepreneurship across different social groups. Our findings reveal that new feeder roads boost the number of service enterprises for all caste categories, including lower-caste groups. However, manufacturing entrepreneurship increases only among upper-caste groups. The new roads expand industry diversity in services for lower-caste groups and in both services and manufacturing for upper-caste groups. For lower-caste groups, the positive impact on entrepreneurship likely stems from the negative effect of new roads on wage employment, which contrasts with the positive impact for upper-caste groups. Lower-caste groups capitalize on market opportunities by starting new businesses, supported by access to formal financing and primary education.

JEL Classification: L26, J15, O12, O18

Keywords: Infrastructure development, rural roads, social inclusion, entrepreneurship, discrimination.

*We thank Aparajit Mahajan, Ashwini Deshpande, Dweepobotee Brahma, Nirupama Kulkarni, Sourav Bhattacharya, Tanika Chakrborty, and the audience members in PacDev 2024 (Stanford King Center on Global Development), and seminar participants at CAFRAL, IIT Jodhpur, Ashoka University, and IIM Calcutta, for valuable comments and feedback. Aishwarya Bhamidipati and Vijayshree Jayaraman provided excellent research assistance. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

†University of California, Santa Cruz. abrahma@ucsc.edu

‡Centre for Advanced Financial Research and Learning (CAFRAL), IZA, and GLO. vidhya.srajan@cafral.org.in

1 Introduction

An extensive literature has documented the role of entrepreneurship in economic growth and development (Baumol, 2002; Glaeser et al., 2015; Kerr and Nanda, 2009; Schumpeter, 1934). Yet, entrepreneurship is often contained within certain socio-demographic groups, which may slow the attainment of inclusive growth. In the United States, the share of firm ownership among Blacks is only 3% and among Hispanics is only 7%. By comparison, Black and Hispanic adults comprised 12.4% and 19% of the overall U.S population in 2020, respectively, according to the U.S Census Bureau's Annual Business Survey (Leppert, 2023). In India, Scheduled Caste (SC) and Scheduled Tribe (ST), the historically disadvantaged social groups, owned 7.6% and 3.9% of the enterprises respectively in 1998. The SC and ST population in contrast was higher at 18.2% and 9.7%, respectively (Deshpande and Sharma, 2013; Iyer et al., 2013). Removing constraints to excluded groups has positive implications for poverty alleviation, employment generation, and inclusive growth and development.

This study explores the effects of rural infrastructure development, particularly the construction of feeder roads that enhance connectivity, on entrepreneurship across different social groups. We use the census of establishments in India comprising a sample of over 6.5 million observations across four waves (in the years 1990, 1998, 2005, and 2013), uniquely containing details on the social group of the owner, and various other aspects of the firms including the industry of operation, size, type of labor employed, registration status, and power availability. Our identifying variation utilizes a shock to the rural road network that started in 2000 and continued over this period arising from a national rural road-building program (the Pradhan Mantri Gram Sadak Yojana or Prime Minister's Village Road Program, or PMGSY) that connected previously unconnected habitations.

Can rural roads promote inclusive entrepreneurship? Rural roads have been shown to promote greater connectivity between settlements and promote market integration (Aggarwal, 2018; Asher and Novosad, 2020). While market integration is expected to reduce social exclusion (Becker, 1957), it may alone not be effective in stimulating entrepreneurship among excluded social groups due to various reasons. In India, for example, there are discriminatory social norms against lower caste groups (Deshpande, 2011). These groups, traditionally relegated to working on stigmatized industries, such as, leather or footwear manufacturing, may not have the scope to expand to other profitable industries even if markets present possibilities. They have restricted access to finance (Banerjee and Munshi, 2004), or sparse access to social-networks or poorer education levels due to which they maybe unable to start and manage enterprises. The empirical literature on infrastructure provision (both small and trunk infrastructure), albeit large, does not focus on its impact across socio-demographic groups in an attempt to study inclusion related outcomes (Aggarwal, 2018; Asher and Novosad, 2020; Banerjee et al., 2020; Donaldson, 2018; Faber, 2014; Maitra and Yu, 2022)[1]. This paper aims to fill this gap. The paper also examines the channels through which individuals belonging to excluded groups could potentially become entrepreneurs when markets open up possibilities through new roads, overcoming discriminatory cultural and social norms.

Using a difference-in-difference framework, we exploit variation in timing of road construction, comparing the evolution of entrepreneurship across social categories in villages that receive a program road to villages that are yet to receive a program road. In India, the Scheduled Castes (SCs) and Scheduled Tribes (STs) are historically-disadvantaged groups recognized in the constitution. The 'Other Backward caste' (OBC) is a collective term used by the Government of India to classify castes which are educationally and socially disadvantaged, although not as acutely as SCs and STs. All other castes are grouped as the 'General Category' (GC). Our baseline results are presented for SC and ST entrepreneurs separately, and for the 'Others' comprising the OBCs and GC together.

Our results indicate that new roads increase the number of service sector enterprises across all caste categories, including lower caste groups. In the manufacturing sector, roads increase entrepreneurship among the non-SC/ST caste groups, and decrease entrepreneurship among SCs without any significant effect on STs. New roads do not change the division of the ownership pie in services sector, but tilts it towards higher non-SC/STs ownership

[1]Trunk transportation infrastructure is large scale infrastructure that connects regions across longer distances, namely, rail networks or highways. In contrast, small transport infrastructure connects closer settlements and habitations using feeder roads.

share in the manufacturing sector although these latter effects are not statistically significant. These results are based on a fixed effects model, including village fixed effects that controls for unobserved village level characteristics, and state-year fixed effects that controls for unobserved factors that may vary across states over time. Due to staggered treatment of villages, we allow for dynamic heterogeneous treatment effects and employ the estimator proposed by De Chaisemartin and d'Haultfoeuille (2020). Our results prevail using this estimator but are more conservative and with the effect on service sector ST enterprises being null. Based on the baseline average number of enterprises, rural roads led to 14% increase in enterprises among the SCs and 10% increase among OBC/Others.

Next, we explore the type of firms that are being created among the lower caste groups, specifically in the services sector where we find a large impact on lower-caste entrepreneurship. Results show that the increase in entrepreneurship among the lower caste groups, namely, the SCs and STs, are concentrated among small firms, single-employee firms, unregistered firms, and firms without any power supply. This suggests that even though we see a rise in firm ownership among the lower caste groups, these firms themselves have limited capabilities to extend beyond sustenance.

Why do SC/ST individuals start enterprises in response to new roads? Can wage employment instead be an alternative engine of development and inclusion for lower caste groups? Rural roads has been shown to increase employment opportunities, even if not by directly generating jobs in the village but at least by providing access to jobs in the nearest towns (Aggarwal, 2018; Asher and Novosad, 2020). Are all caste groups able to obtain these jobs? We expect this not to be the case because studies suggest evidence of homophily in employment (Deshpande and Sharma, 2013; Iyer et al., 2013), whereby the employment share of non-SC/ST-owned firms is titled towards non-SC/ST employees. Further, there is evidence of discrimination in the job market against lower caste groups and racial minorities (Banerjee et al., 2009; Bertrand and Mullainathan, 2004). We test for potential differential access to jobs based on caste groups in the labor market using data from the Employment-Unemployment survey (EUS) of the National Sample Survey Organization in the years 2004, 2004-2005, 2009-2010, and 2011-2012. We define a district-level treatment variable because the EUS provides only district identifiers, and not village identifiers. We show that districts more exposed to the rural roads program saw an increase in employment probabilities (consistent with Aggarwal, 2018). However, these effects were significantly differentially negative for the SC/STs. Given the lack of success in the job market for SC/STs, it is not surprising that the margin of response for SC/STs entrepreneurship is positive.[2]

To better understand the mechanisms through which road construction affects entrepreneurship possibilities, we explore heterogeneity in treatment effects. The first is the financial channel whereby with the markets opening from new roads, lower caste entrepreneurs could obtain credit especially in areas with better access to formal financing. Formal financing can lead to the starting of enterprises among lower caste groups who are typically considered "outsiders" in the business of business and who lack the network capital to obtain starting capital (Banerjee and Munshi, 2004). The second is the human capital channel whereby rural roads can lead to the accumulation of human capital (Adukia et al., 2020; Aggarwal, 2018). Thus, a plausible mechanism is that after rural road construction, new roads could have lead to lower caste individuals with higher exposure to formal schooling to start enterprises to take advantage of the new market opportunities.

We show evidence consistent with both channels. Using bank branch presence data from the Central Information System for Banking Infrastructure (CISBI), Reserve Bank of India, we find that the impact of rural roads on SC/ST entrepreneurship is significantly higher in areas with more bank branches. We find larger effects for bank branches of state-owned banks compared to private banks, consistent with the greater outreach of public banks in India (Berger et al., 2008) and particularly in their role in reaching out to the excluded groups.[3] Second, with respect to the human capital channel, we show evidence that the

[2]There is also a rising consciousness in *dalit* (often synonymous with the Scheduled Castes) entrepreneurship. The *dalit* community increasingly questions the adequacy of employment opportunities in primarily upper-caste owned enterprises where their growth and human capital accumulation capabilities are minimal. The Dalit Indian Chamber of Commerce and Industry (DICCI) was established in 2005 with an aim to bring together *dalit* entrepreneurs, allowing aspiring *dalit* entrepreneurs to take advantage of the 'network effects'.

[3]We also utilize credit data at the bank branch level from the Basic Statistical Returns of the Reserve Bank of India. We aggregate bank-branch level credit to the village-level. We show that rural roads improve entrepreneurship outcomes for lower caste groups in areas with higher credit activity, indicating that credit flow is an important channel for lower caste groups. Admittedly, higher credit can also flow

impact of rural roads on entrepreneurship is differentially higher for all caste groups in areas with a larger presence of primary schools.

We contribute to three strands of literature. First, we speak to the literature on the impact of infrastructure creation on economic outcomes. While a large literature exists on the economic impacts of infrastructure creation, few studies examine the distributional impacts.[4] Ghani et al. (2016) studied investments in the Golden Quadrilateral (GQ) construction in India and found that female entrepreneurship and employment was higher in districts far away from the highway construction compared to districts closer, contradicting the expectations of Becker's theory on the impact of competitive forces on discrimination (Becker, 1957). A recent set of studies have found that highway construction in the United States led to more spatial segregation for the Black population, but they have not directly examined economic impacts across racial lines (Bagagli, 2023; Mahajan, 2023; Weiwi, 2023).[5] We contribute to this literature by examining the impact of rural road construction in India on the potential of inclusivity in entrepreneurship for excluded social categories.

In contrast to studies focusing on trunk infrastructure, we study feeder road construction in the rural hinterlands in India. Previous studies show that rural roads reduced poverty in Bangladesh (Khandker et al., 2009), enabled households to switch from agriculture to non-agricultural service based activities in Vietnam (Mu and Van de Walle, 2011), and increased the use of agricultural inputs and market orientation among farmers in Ethiopia (Nakamura et al., 2019). In the Indian context, evidence shows that the PMGSY road access

due to road presence. Therefore, we consider this as suggestive evidence.

[4]There is an entire literature studying the broad impacts of infrastructure on growth. Banerjee et al. (2020) find that proximity to transportation networks in China have a moderate positive causal effect on per capita GDP levels across sectors, but no effect on per capita GDP growth. Faber (2014) finds that Chinese highways decreases local GDP in rural areas newly connected to more productive urban centers. Maitra and Yu (2022) found that railways had an immediate impact on trade and development in the predominantly agricultural India, and that these positive effects have persisted over more than a century. Districts in the Indian subcontinent that gained railway access earlier still exhibit greater economic prosperity and lower rural poverty rates even a century later.

[5]Mahajan (2023) found an increase in the share of the black population close to highways, driven by the white population moving out and black population moving into these neighborhoods. Weiwi (2023) found that the highway system mostly benefited suburban regions. The Black population's initial concentration in central areas and their low mobility away contributed to their welfare losses from the interstate highway system. Bagagli (2023) showed that expressway construction in Chicago in the 1950s increased racial segregation.

increases market integration (Aggarwal, 2018) and enables structural transformation by bringing workers out of agriculture (Asher and Novosad, 2020), improved crop variety and agricultural technology adoption especially in remote villages, and improves educational attainment (Adukia et al., 2020; Aggarwal, 2018). However, little is known about whether these improvements from road access is universal or if it only accrues to particular social groups. While trunk infrastructure has shown to increase spatial segregation across racial lines, we hypothesize that rural feeder roads that connect the hinterlands may improve inclusivity.

We also speak to the literature that examines caste based differences in economic and human capital outcomes, and how markets, institutions, or the state can potentially enhance social mobility in societies where cross-group inequality is entrenched. Lower caste groups have persistently lower consumption expenditure, education levels, social networks, and access to public goods (Banerjee and Somanathan, 2007; Desai and Dubey, 2012). Deshpande and Sharma (2016) use quantile decomposition to show that SC/ST-owned businesses at the lower and middle end of the conditional earnings distribution face greater discrimination in terms of earnings. Iyer et al. (2013) document that SCs and STs are underrepresented in the entrepreneurial sphere even decades after the economic reforms in 1991.

Previous literature showed that affirmative action policies increased transfers (Pande, 2003) and reduced poverty (Chin and Prakash, 2011) among lower caste groups; improvements in court quality positively impact entry and investment decisions among SC and ST firm owners (Chakraborty et al., 2023); reforming inheritance laws for women to inherit equal property as compared to men significantly increased firm creation by women without worsening the quality of new entrants (Naaraayanan, 2019); and also that the structural reforms in India post-1991 correlated with reduction in gaps between SC/STs and non-SC/STs in educational attainment, wages, occupational mobility rates (Hnatkovska et al., 2012, 2013). Adding to this literature, we examine how improvements in rural infrastructure, specifically rural feeder roads, contributes towards the entrepreneurial success of the marginalized castes.

Finally, we speak to the literature on determinants and correlates of entrepreneurship. Evidence shows that parental human capital and family background/wealth strongly influences entrepreneurship in advanced countries.[6] Institutional and spatial features play a role. Kerr and Nanda (2009) show that US banking reforms brought about exceptional growth in entrepreneurship. Glaeser et al. (2015) show that historic presence of coal/mineral deposits reduced entrepreneurship because mining and allied activities subsume resources towards themselves. In developing countries, culture and social norms play a role. In addition to the previous literature cited on lower entrepreneurship rates among lower caste groups, evidence shows that women-led small enterprises have lower growth in India (Coad and Tamvada, 2012), but also the presence of incumbent female-owned businesses is correlated with higher subsequent new female entrepreneurship highlighting the importance of role models (Ghani et al., 2013).

The rest of the paper is organized as follows: section 2 discusses the data and the context, section 3 presents the empirical methodology, section 4 presents the results, and section 5 concludes.

2 Context and Data

2.1 The PMGSY

The government of India launched the *Pradhan Mantri Gram Sadak Yojana* (PMGSY) in 2000 to provide connectivity to unconnected habitations. The main goal was to establish reliable, all-weather connectivity to selected villages by constructing paved roads with cross-drainage structures, ensuring they remain functional in all conditions. The roads constructed under PMGSY were designed to connect unconnected villages to the nearest village with an all-weather road, the nearest all-weather road, the market center, or the block headquarters. To be eligible for the program, a village had to be unconnected, meaning it was located at least 500 meters away from an all-weather road or another village

[6] See Dunn and Holtz-Eakin (2000) for the United States, and La Porta et al. (1999) for a study of 27 wealthy economies.

with such a road, and it should not already have a paved road.

PMGSY is fully funded by the central government, but is managed by the state governments. Each state generates a list of all unconnected villages within the state, with villages ranked in descending order by population size (as recorded in the 2001 Population Census). After discussions at the district level, a priority ranking of villages across all districts within a state was then generated. Highest priority is given to villages above the population size of 1000, second highest priority to villages with population size between 500 and 1000, and ultimately to villages above the size of 250 and below 500. The program also upgraded existing roads, but new construction has much higher priority and a state could upgrade roads only if it had completed all new construction. Therefore, we restrict our analysis to previously unconnected habitations. The program is still active as of writing this paper. Between 2000 and 2015, the program connected 124,935 habitations and built roads spanning 419,358 kilometers.

2.2 Data Sources and Description

The primary data source is the Economic Census of India, conducted by the Central Statistical Organisation (CSO), Government of India. We use the 3^{rd}, 4^{th}, 5^{th}, and 6^{th} waves of the census, conducted in 1990, 1998, 2005, and 2013 respectively. The Economic Census is a "complete count of all entrepreneurial units located within the geographical boundaries of the country, involved in any economic activities of either agricultural (excluding crop production and plantation) or non-agricultural sector of the Economy, engaged in production and/or distribution of goods and/or services not for the sole purpose of own consumption". The Economic Census provides data on the enterprise's employment, registration status, location, industry classification, power use, and caste category of the enterprise owner. We assemble the EC at the village level, creating aggregate number of enterprises owned by various caste categories (SCs, STs, OBCs, and other castes) and of different enterprise type at the village level.

We also use the Population Census of India for the years 1991 and 2001. These provide data on village characteristics and demographics that we use as control variables. These variables include the total population, the share of SC population, the share of ST population, number of primary schools, number of middle schools, number of secondary schools, number of higher secondary schools, number of colleges, availability of power supply, and presence of previously paved road. We merge the variables from the EC and PC using the Socioeconomic High-resolution Rural-Urban Geographic Platform for India (SHRUG, Version 2.0) created by Asher et al. (2021), which provides village-level identifiers compatible across these datasets.

The SHRUG database also contains information on roads built under PMGSY in each village. Asher and Novosad (2020) scraped the PMGSY website (Online Management, Monitoring and Accounting System) and mapped details of road construction, such as sanction cost, sanction year, completion date, road length, etc. to the SHRUG database.[7] We use the completion date to construct an indicator variable for whether a previously unconnected habitation received a road or not by that year, and employ this as our main treatment variable.

We also use the Employment-Unemployment Survey (EUS) conducted by the National Sample Survey Organization in the years 2004, 2004-2005, 2009-2010, and 2011-2012. The EUSs were are generally conducted between July of the first year to June of the second year. For example, the 2004-2005 survey is conducted from July 2004 to June 2005. The exception is the 2004 survey that was conducted between January to June 2004. The EUS obtains the status of individuals related to their job market (whether they are employed, unemployed, or out of the labor force). Since the EUS does not contain village identifiers, we merge the PMGSY data at the district level to the EUS. Specifically, in the spirit of Aggarwal (2018), we obtain the share of villages that obtained a PMGSY road by the previous year for each district and employ this as our treatment variable.

We also use bank-branch opening date and their addresses from the Central Information System for Banking Infrastructure (CISBI), Reserve Bank of India, to examine the financial channel. Using bank branch opening date along with their pincodes, we obtain the number of existing bank branches at the 6-digit pincode level. We then match pincodes to villages

[7] https://omms.nic.in/

since the Economic Census is at the village level. The village names corresponding to the pincodes were available from the Indian Postal Service website.[8] A pincode can be matched to multiple villages. By assigning bank branch presence from pincode to village level, we assume that all villages that are a part of the pincode have access to all the bank branches in that pincode. Finally, we merge the number of bank branches the village is exposed to with the Economic Census based on village names using a fuzzy text matching approach. In performing this fuzzy match, we lose about 25,000 villages from our original sample of villages. Finally, we are able to obtain village-level data on bank presence for 146,882 villages.

Additionally, we also use confidential data on bank-credit flow at the bank-branch level from the Basic Statistical Returns (BSR) data from the Reserve Bank of India. When a village spans multiple pincodes, we assume that the credit to the village flows from the pincode where the largest share of the village lies. We overlay shapefiles at the village and pincode levels to calculate this.

Habitations that were already connected in 2000 (as per the information in population census) are dropped from our analysis sample. Our final analysis sample spans 172,365 habitations across 30 states and has 529,881 observations. The descriptive statistics for our estimation sample are presented in Table 1. The number of enterprises have grown across all caste groups, particularly significantly so in the services sector. In the manufacturing sector, in 1990, the share of ownership was the highest among the Others (OBC + General category), followed SCs and STs. In 2013, the shares increased slightly for OBC and STs, and decreased slightly for STs. In the services sector, the share of ownership in 1990 was highest among Others, followed by SCs and STs. In 2013, the shares for Others declined substantially, but increased for SCs and STs.

Table 2 presents the average number of enterprises in a village across caste groups. The average is highest for Others, followed by SC and then ST for both control and treated villages across all years. Treated villages here refers to those that ever got treated after 2000 during the program years. Treated villages had higher average number of enterprises than

[8] https://www.indiapost.gov.in/VAS/Pages/findpincode.aspx

control villages (although they are not statistically significantly different in every case) suggesting that larger villages were perhaps treated before smaller villages.[9] Table A1 provides descriptive statistics on the number of villages that obtained a road under the PMGSY each year based on road completion dates, and Table A2 presents the descriptive statistics on the number of new villages gaining a new road under PMGSY across each population group in each year.

3 Methodology

We estimate the impact of rural roads on entrepreneurship across caste group categories. Our primary specification is defined by the following equation:

$$Y_{vt} = \beta Road_{vt} + \gamma_{st} + \eta_v + \epsilon_{vt} \tag{1}$$

where $Road_{vt}$ is an indicator variable that equals 1 if the village v was connected by a PMGSY road by year t, and Y_{vt} is the outcome variable denoting the number of enterprises owned by a specific caste group in village v in year t.[10] Top 1% of number of enterprises are winsorized. We use the Economic Census data for the years 1990, 1998, 2005, and 2013 for estimating these models. The caste group could be Scheduled Caste (SC), Scheduled Tribe (ST), or Others (combining Other Backward Categories (OBC) or the general category). We combine OBC and the GC because the OBC category did not exist in 1990, and because our analysis sample starts in 1990. The OBC category was only created after the Mandal commission recommendations in 1991.[11] γ_{st} represents state-year fixed effect, and η_v represent village fixed effects. The ϵ_{vt} is the error term. We cluster standard errors at the village level.

Since we exploit the the timing of road completion varying across village and time, we

[9] We do not report the test statistics related to the statistical significance of the difference between average enterprises in the treatment and control groups, but these are available upon request.

[10] We do not use logarithmic transformation of outcome variables because a large proportion of them are zeros, indicating the absence of enterprises owned by specific caste groups in many villages (Chen and Roth, 2024). We also report Poisson regression results using a similar empirical model in section 4.2

[11] Using the 1999, 2005, and 2013 Economic Census data, in a separate specification in the appendix, we also estimate these regressions using OBC and Others as separate categories.

have, in essence, adopted a staggered treatment design. Recent studies indicate that estimates of Equation 1 may be biased if there are time-varying treatment effects and that the difference-in-differences estimates itself would be biased (Callaway and Sant'Anna, 2021; De Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2021). Therefore, we also estimate dynamic treatment effects and difference-in-difference estimates using the DIDM method proposed by De Chaisemartin and d'Haultfoeuille (2020). We estimate immediate (period 0) treatment effects and effects for the subsequent two periods as well as placebo estimates for one period before a policy was implemented. These treatment effects are based on the counterfactual comparison of one-period differences in villages that switched status compared to those units that did not switch status.

In other variants of Equation 1, we also use the share of enterprises owned by different caste groups as outcome variables. We also explore the heterogeneity of the above results across firms. For this, we replace the outcome variable in Equation 1 with the same measure constructed for different sub-groups of firms: small firms/large firms using employee size 10 as the cutoff (based on the lowest threshold size for formality of firms in India under the Factories Act, 1948); single-employee/multiple employee firms; firms without power/firms with power availability; and finally, registered/unregistered firms.

For robustness checks, in separate specifications, we include district-time fixed effects instead of state-time fixed effects. We also include district-time trend along with state-time fixed effects in a separate specification. In another specification, instead of village fixed effects, we control for a set of village-level demographic controls in 1991 interacted with year fixed effects, and separately interacted with time-trends. These controls include the number of primary schools, presence of educational facility, adult literacy centre, communication facility, market facility, and power supply. As another robustness check, we also control for the 2001 SC population interacted with time trend, and 2001 ST population interacted with time trend.

We address the endogeneity concerns related to timing of rural road construction using various ways. Under PMGSY road construction guidelines, states were instructed to first target villages with populations greater than 1000 in the population census, and then villages with population greater than 500. Since we use village-year variation in the completion of new roads, there is potentially an endogeneity issue in the timing of road construction in a particular village and whether we are using similarly sized villages as the control group. To address this concern, we define a dummy variable called "eligible" village taking the value 1 if their population is above 500, and 0 otherwise. We conduct multiple analyses. First, we estimate equation 1 only for eligible villages so that we compare eligible treated villages with eligible control villages. Next, we create a eligible-village-state dummy that takes the value 1 for eligible village in a specific state and 0 otherwise. Since eligible villages in a specific state could have a different time-trend (or year-specific outcomes) compared to eligible villages in other states, we control for these dummy variables interacted with year-fixed effects, and dummy variables interacted with time trends. Next, we estimate the dynamic estimator based on De Chaisemartin and d'Haultfoeuille (2020) that allows us to test whether treated and control districts in the pre-treatment period were on similar trajectory with respect to the outcome variables. If the placebo estimate in the period before the treatment was null, it mitigates the endogeneity concerns posed by the timing of the road competition.

Next, we test if SC/ST individuals residing in rural areas take up wage employment in response to new road construction and if the response differs by caste group. For this, we estimate the following equation:

$$Y_{idt} = \lambda Villageroadshare_{dt} \times SC/ST_{idt} + \beta Villageroadshare_{dt} + \gamma_t + \eta_d + \epsilon_{idt} \qquad (2)$$

where $Villageroadshare_{dt}$ is the share of villages in district d that had new PMGSY roads by time t. SC/ST_{idt} is a dummy variable that takes the value 1 for a SC/ST household and 0 otherwise. The outcome variable Y_{idt} is a dummy variable that takes the value 1 if individual i is engaged in wage employment in district d in time t, 0 otherwise. Using Equation 2, we estimate the impact of higher exposure to new roads at the district level on wage employment for the base category (non-SC/ST) (coefficient β) households and for SC/ST households (λ).

Next, we explore the channels through which road construction might affect entrepreneurship among lower caste groups. First, we test the financial channel. For this, we estimate the following specification, a variant of Equation 1:

$$Y_{vt} = \delta Road_{vt} \times Branches_{vt} + \beta Road_{vt} + \gamma_{st} + \eta_v + \epsilon_{vt} \tag{3}$$

$Branches_{vt}$ refers to the number of branches in village v at time t. We test if villages that have a large number of bank branches see a differentially higher effect on the number of enterprises owned by lower caste groups. The other parts of the equation are same as Equation 1. The parameter of interest is δ. A positive δ indicates that the impact of rural roads on entrepreneurship of a specific caste group is higher with a higher number of bank branches. We also estimate the same specification separately for public bank branches and private bank branches to examine if the results differ. Second, we test the human capital channel. For this, we estimate the following specification, a slight variant of Equation 1.

$$Y_{vt} = \kappa Road_{vt} * Primaryschool_{vt} + \beta Road_{vt} + \gamma_{st} + \eta_v + \epsilon_{vt} \tag{4}$$

$Primaryschool_{vt}$ takes the value 1 if village v in time t has a higher than median share of primary schools, and 0 otherwise. We test if villages that have a large primary school presence see a differentially higher effect on the number of enterprises owned by lower caste groups. The parameter of interest is κ. A positive γ indicates that the impact of rural roads on entrepreneurship of a specific caste group is higher with a large primary school presence, indicating that the human capital channel is salient.

4 Results

4.1 Main Results

Table 3 shows estimates of the effect of a new road on the number of enterprises by owner's caste. Columns (1)–(3) report the estimates for the manufacturing sector and columns

(4)–(6) report results for the services sector. Panel A presents the results using two-way fixed effects regression (Equation 1). The results indicate that the impact of rural roads on manufacturing sector is positive for Others-owned enterprises, negative for SC-owned enterprises, and there is no effect on ST enterprises. There is an increase in enterprises among all caste categories in the services sector. In panel B, we allow for heterogeneous treatment effects following De Chaisemartin and d'Haultfoeuille (2020). In the manufacturing sector, the positive effects of rural roads presence is observed for Others-owned enterprises along with a decline in SC-owned enterprises that is not statistically significant. In the services sector, positive and significant effects are observed and is highest for OBCs followed by SCs. Based on the baseline average number of service enterprise across caste groups (Table 2), rural roads lead to 14% increase in enterprises among the SCs and 10% increase among OBC/Others. The rise in service enterprises compared to manufacturing enterprises among lower caste groups aligns with the expectation that manufacturing businesses require more initial capital and are therefore more challenging to establish.

Table A3 presents the results separately for number of enterprises in OBC and the general category enterprises as outcome variables utilizing three years of the Economic Census (1998, 2005 and 2013). Panel A shows that rural roads have a positive impact on enterprises for the General category and OBC, but the effect is not statistically significant for OBC. De Chaisemartin and d'Haultfoeuille (2020) estimator from panel B indicates that the effects on OBC enterprises are significant while that on the General Category is not. The positive impact in the services sector is seen for both OBC and General category.

We next test if rural roads presence changed the share of ownership among caste groups using shares of enterprises for each caste group as the outcome variable. The results are presented in table Table 4.[12] Columns (1)–(3) report the estimates for the manufacturing sector and columns (4)–(6) report results for services sector. OLS and De Chaisemartin and d'Haultfoeuille (2020) estimators show that there is no effect on the share of enterprises in the manufacturing and services sectors, implying that the composition of caste groups remain unchanged as a result of rural roads. Broadly, these results indicate that rural roads

[12]The sample size in Table 3 is lower than Table 4 because there are many villages where the total number of manufacturing firms or service firms is zero. Division by zero yields a missing value.

helped increase the number of enterprises across the social groups in the services sector. While the pie has considerably increased, its composition remains broadly unchanged particularly from the lower-caste ownership viewpoint. Table A4 shows that all caste group owners are producing across a wider variety of two-digit industries in the services sector, taking cues from the market opportunities. Still, in the manufacturing, only Others-owned enterprises are diversifying after the construction of new roads.

4.2 Robustness Checks

Table A5 presents the results at alternate levels of clustering, namely, state-level and the state-year-level. Our results remain robust. Due to the large number of zeroes in the dependent variable, we also estimate our model using Poisson regression. Table A6 confirms that the results remain robust. Next, we test the impact on the number of enterprises per capita at the village level across caste groups. The results are presented in Table A7. The results indicate that the positive impact on services are robust for SCs and Others. Our results are also robust when we consider a balanced panel of villages across the four years of the EC (Table A8). To ensure that our results are not driven by the largest villages in our sample, we drop the top decile of villages by population and re-estimate Equation 1. Our results are qualitatively similar (Table A9).

We show additional robustness checks for services sector, as that is where the increase in entrepreneurship is observed for the lower caste groups. Table 5 show that our estimates in the service sector are robust to the inclusion of district-year trends along with state-year fixed effects, or district-year fixed effects in place of state-year fixed effects. Table 6 presents the results for the service sector after including village-specific characteristics and demographics at the baseline year (1991) multiplied by year dummies (columns (1)–(3)) or multiplied by a time trend (columns (4)–(6)). The results remain robust. Table 7 presents the results after including SC population in 2001 $\times$ time trend, and ST population in 2001 $\times$ time trend both in the manufacturing (columns (1)–(3)) and in the services sectors (columns (4)–(6)). The results remain robust.

4.3 Heterogeneity by Firm Characteristics

What types of firms are created by lower caste group owners? Next, we examine the heterogeneity across various firm characteristics. Results in Table 8 and Table 9 show that the effects on SC and ST enterprises are larger for small firms, single-employee firms, firms using non-hired labor, firms without power, and firms that are unregistered (see panel A in both tables). The impact on ST enterprises, particularly, is absent in large and registered firms, and those using power (panel B, Table 9). These results collectively point out that even though there are gains to the lower caste groups in terms of their entrepreneurship opportunities, these remain constricted to firms with limited growth potential.

4.4 Endogeneity Concerns

The dynamic estimators from De Chaisemartin and d'Haultfoeuille (2020) are presented in Figure 1 for Others-owned, Figure 2 for SC-owned, and Figure 3 for ST-owned enterprises. The results indicate a strong positive effect on the number of enterprises for Others in both services and manufacturing, and an absence of pre-trends. For SC and ST manufacturing enterprises, there are no significant effects both before the program and after the program. For SC service enterprises, there is a strong positive effect on the number of enterprises, especially in the second period post treatment, and an absence of pre-trend. For ST service enterprises, there are no effects post treatment, but a small negative pre-trend one wave before treatment. Overall, results indicate that there are no strong pre-trends driving our main results in Table 3.

Table 10 addresses further endogeneity concerns. Columns (1)–(3) presents the main results only for eligible villages so that we compare eligible treated villages with eligible control villages. Columns (4)–(6) presents results after controlling for dummy variables for eligible village specific state interacted with year. Columns (7)–(9) presents results after controlling for dummy variables for eligible village specific state specific time trends. In all these instances, the results remain robust.

To further alleviate endogeneity concerns, we estimate placebo regressions by randomizing the road completion year for each village in our sample, and then estimating Equation 1 using the placebo year as the treatment year. We do this randomization and estimation 500 times. The average of these 500 regressions coefficients and standard errors are reported in Table A10. The estimated placebo coefficients are statistically indistinguishable from zero.

4.5 Mechanisms

We investigate the impact of rural road construction on wage employment. Table 13 estimates Equation 2 to show that the districts more exposed to the rural roads program saw an increase in employment probabilities, consistent with Aggarwal (2018). However, these effects were significantly differentially negative for the SC/STs. Given this evidence, it is not unexpected that the margin of responses by the SC/STs is positive for entrepreneurship.

We test the financial channel by estimation Equation 3. This checks if bank presence enables credit flow to particularly lower caste groups in helping them start a business once a market potential through roads opens up. The results are shown in Table 11. Results for interaction with all branches are reported in columns (1)-(3), for interaction with public branches in columns (4)-(6), and for interaction with private branches in columns (7)-(9). The effect of new roads on the number of enterprises is differentially higher with more bank branches for SC-owned enterprises (column 1). This is consistent with lower caste group having low levels of network capital unlike the OBCs, and with formal finance channel intermediation plausibly easing access to capital for these groups (Banerjee and Somanathan, 2007). Not surprisingly, the results are driven by public banks, consistent with their greater emphasis and initiatives on inclusion. Further, we test for heterogeneity based on credit activity. Table A11 presents the results using number of credit accounts (columns 1-3), and amount outstanding (columns 4-6). The results indicate that positive effects for SC/ST entrepreneurship is observed in areas with higher credit accounts and higher loan amount outstanding.

Table 12 presents the results from testing the human capital channel based on estimating Equation 4. Columns (1) and (2) indicate that the impact of new roads on SC and ST owned enterprises is differentially higher in villages with above median number of primary schools indicating that education is playing a crucial role in SC/ST individuals setting up enterprises.

5 Conclusion

Using data from the universe of enterprises in India, we show that entrepreneurship across social groups increased in the services sector following a national road program that connected previously unconnected habitations. We find significant increases in SC and ST owned enterprises in the services sector but not in the manufacturing sector. We also do not find evidence of any change in the entrepreneurial pie in both the services and the manufacturing sectors, although there is increase in the diversity of industries owned by all caste groups in the service sector. The positive effects on SC and ST entrepreneurship is higher in small firms (employing less than 10 workers), firms run by a single employee, firms without power, and unregistered firms, as compared to large firms, multiple employee firms, firms using power supply, and registered firms.

The positive impacts of rural roads on SC/ST entrepreneurship makes sense because the impact of new roads on wage employment is differentially negative for SC/STs, compared to OBCs/General Category for whom there is a positive impact. Thus, SC/STs are exploiting the opening up of markets and the ensuing new opportunities by starting new businesses. This is fueled by the financial and the human capital channel. The effects on SC and ST entrepreneurship are primarily in villages with higher number of banks, particularly public banks, suggesting that the formal financing channel is important for lower caste groups to start enterprises. The human capital channel also emerges to be important. The impact of new roads on SC/ST firm ownership is particularly higher in villages that have a above-median number of primary schools. Overall, we conclude that rural infrastructure program enables lower caste group individuals to take advantage of the market opportunities by taking up entrepreneurial activities fueled by formal credit and aided by human capital.

References

Adukia, A., Asher, S., and Novosad, P. (2020). Educational investment responses to economic opportunity: evidence from Indian road construction. *American Economic Journal: Applied Economics*, 12(1):348–376.

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

Asher, S., Lunt, T., Matsuura, R., and Novosad, P. (2021). Development research at high geographic resolution: An analysis of night-lights, firms, and poverty in India using the SHRUG open data platform. *The World Bank Economic Review*, 35(4).

Asher, S. and Novosad, P. (2020). Rural Roads and Local Economic Development. *American Economic Review*, 110(3):797–823.

Bagagli, S. (2023). The (Express)Way to Segregation: Evidence from Chicago. Working paper.

Banerjee, A., Bertrand, M., Datta, S., and Mullainathan, S. (2009). Labor market discrimination in delhi: Evidence from a field experiment. *Journal of comparative Economics*, 37(1):14–27.

Banerjee, A., Duflo, E., and Qian, N. (2020). On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics*, 145:102442.

Banerjee, A. and Munshi, K. (2004). How efficiently is capital allocated? evidence from the knitted garment industry in tirupur. *The Review of Economic Studies*, 71(1):19–42.

Banerjee, A. and Somanathan, R. (2007). The political economy of public goods: Some evidence from India. *Journal of development Economics*, 82(2):287–314.

Baumol, W. J. (2002). Entrepreneurship, innovation and growth: The David-Goliath symbiosis. *Journal of Entrepreneurial Finance, JEF*, 7(2):1–10.

Becker, G. (1957). The economics of discrimination.

Berger, A. N., Klapper, L. F., Peria, M. S. M., and Zaidi, R. (2008). Bank ownership type and banking relationships. *Journal of Financial Intermediation*, 17(1):37–62.

Bertrand, M. and Mullainathan, S. (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review*, 94(4):991–1013.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2):200–230.

Chakraborty, T., Mukherjee, A., Saha, S., and Shukla, D. (2023). Caste, courts and business. *Journal of Economic Behavior & Organization*, 212:333–365.

Chen, J. and Roth, J. (2024). Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics*, 139(2):891–936.

Chin, A. and Prakash, N. (2011). The redistributive effects of political reservation for minorities: Evidence from India. *Journal of development Economics*, 96(2):265–277.

Coad, A. and Tamvada, J. P. (2012). Firm growth and barriers to growth among small firms in India. *Small Business Economics*, 39:383–400.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

Desai, S. and Dubey, A. (2012). Caste in 21st century India: Competing narratives. *Economic and political weekly*, 46(11):40.

Deshpande, A. (2011). *The grammar of caste: Economic discrimination in contemporary India*. Oxford University Press.

Deshpande, A. and Sharma, S. (2013). Entrepreneurship or survival? Caste and gender of small business in India. *Economic and Political Weekly*, pages 38–49.

Deshpande, A. and Sharma, S. (2016). Disadvantage and discrimination in self-employment: Caste gaps in earnings in Indian small businesses. *Small Business Economics*, 46:325–346.

Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5):899–934.

Dunn, T. and Holtz-Eakin, D. (2000). Financial capital, human capital, and the transition to self-employment: Evidence from intergenerational links. *Journal of labor economics*, 18(2):282–305.

Faber, B. (2014). Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System. *Review of Economic Studies*, 81(3):1046–1070.

Ghani, E., Kerr, W. R., and O'Connell, S. D. (2013). Local industrial structures and female entrepreneurship in India. *Journal of Economic Geography*, 13(6):929–964.

Ghani, S. E., Grover, A., Kerr, S. P., and Kerr, W. (2016). Will market competition trump gender discrimination in India? *World Bank Policy Research Working Paper*, (7814).

Glaeser, E. L., Kerr, S. P., and Kerr, W. R. (2015). Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics and Statistics*, 97(2):498–520.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of econometrics*, 225(2):254–277.

Hnatkovska, V., Lahiri, A., and Paul, S. (2012). Castes and labor mobility. *American Economic Journal: Applied Economics*, 4(2):274–307.

Hnatkovska, V., Lahiri, A., and Paul, S. B. (2013). Breaking the caste barrier: Intergenerational mobility in india. *Journal of Human Resources*, 48(2):435–473.

Iyer, L., Khanna, T., and Varshney, A. (2013). Caste and Entrepreneurship in India. *Economic and Political Weekly*, pages 52–60.

Kerr, W. R. and Nanda, R. (2009). Democratizing entry: Banking deregulations, financing constraints, and entrepreneurship. *Journal of Financial Economics*, 94(1):124–149.

Khandker, S. R., Bakht, Z., and Koolwal, G. B. (2009). The Poverty Impact of Rural Roads: Evidence from Bangladesh. *Economic Development and Cultural Change*, 57(4):685–722.

La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (1999). Corporate ownership around the world. *The journal of finance*, 54(2):471–517.

Leppert, R. (2023). A look at Black-owned businesses in the U.S. *Pew Research Center*.

Mahajan, A. (2023). Highways and segregation. *Journal of Urban Economics*, page 103574.

Maitra, P. and Yu, W. (2022). The long shadow of infrastructure development: Long run effects of railway construction in colonial india. *Available at SSRN 4010598*.

Mu, R. and Van de Walle, D. (2011). Rural Roads and Local Market Development in Vietnam. *The Journal of Development Studies*, 47(5):709–734.

Naaraayanan, S. L. (2019). Women's inheritance rights and entrepreneurship gender gap. *Department of Finance, HKUST*.

Nakamura, S., Bundervoet, T., and Nuru, M. (2019). Rural Roads, Poverty, and Resilience: Evidence from Ethiopia. *World Bank Policy Research Working Paper*, (8800).

Pande, R. (2003). Can mandated political representation increase policy influence for disadvantaged minorities? Theory and evidence from India. *American economic review*, 93(4):1132–1151.

Schumpeter, J. A. (1934). *The theory of economic development*. Cambridge, MA: Harvard University Press.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of econometrics*, 225(2):175–199.

Weiwi, L. (2023). Unequal Access: Racial Segregation and the Distributional Impacts of Interstate Highways in Cities. Working paper.

Figure 1: Impact of Rural Roads on the Number of Other Enterprises

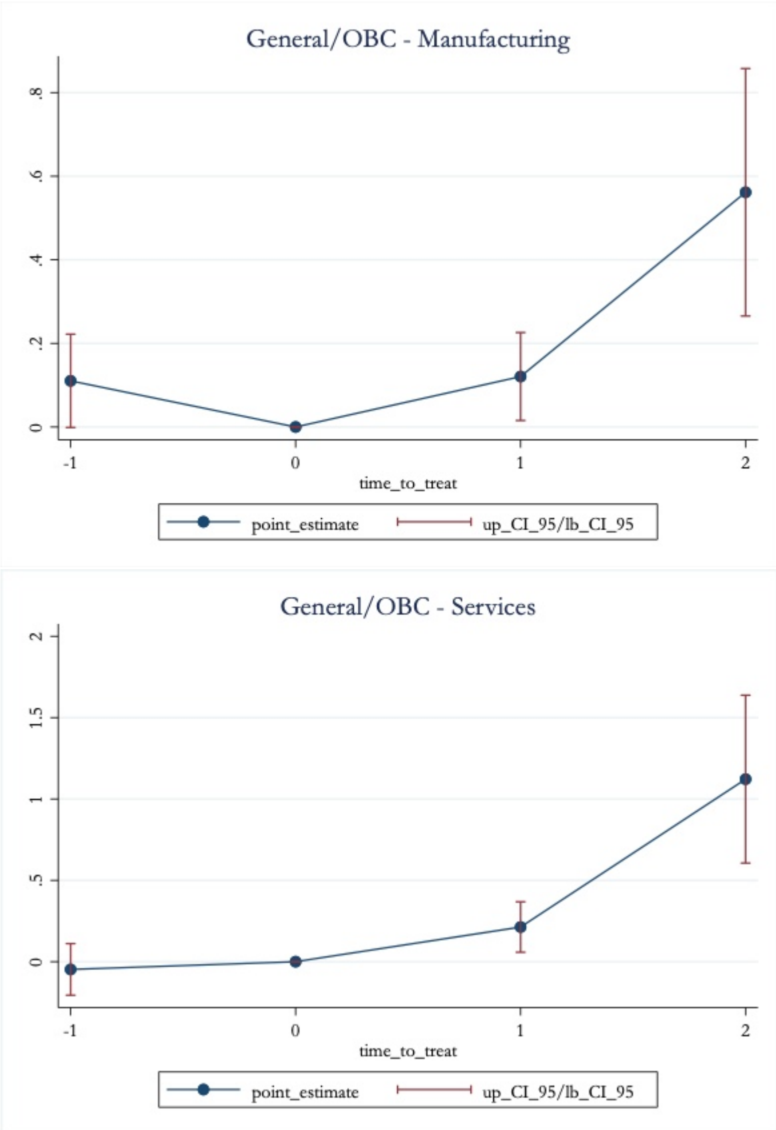
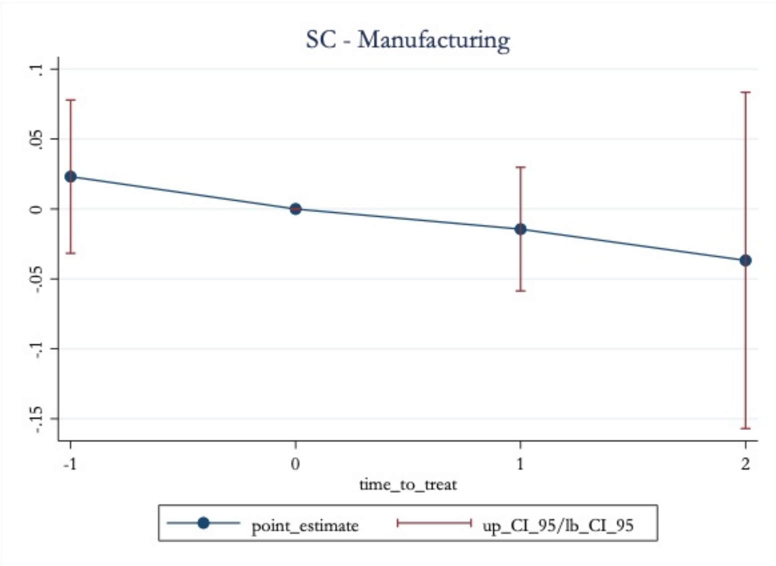


Figure 2: Impact of Rural Roads on the Number of SC Enterprises



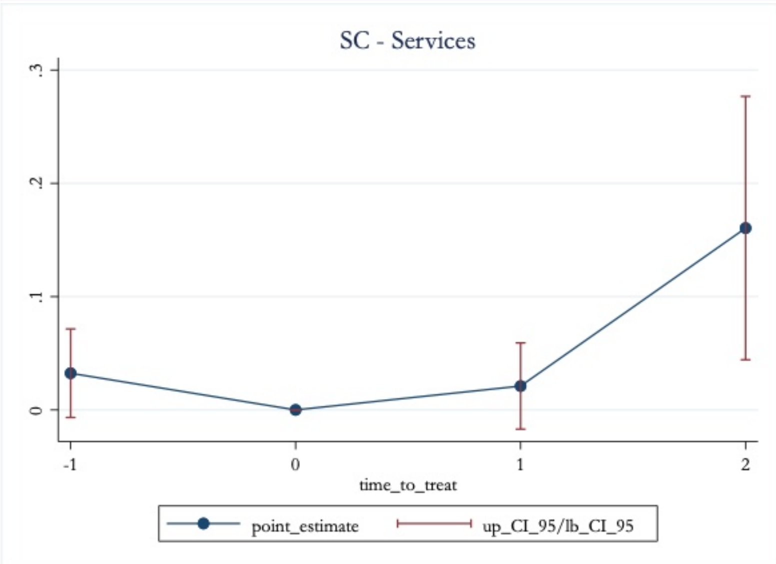


Figure 3: Impact of Rural Roads on the Number of ST Enterprises

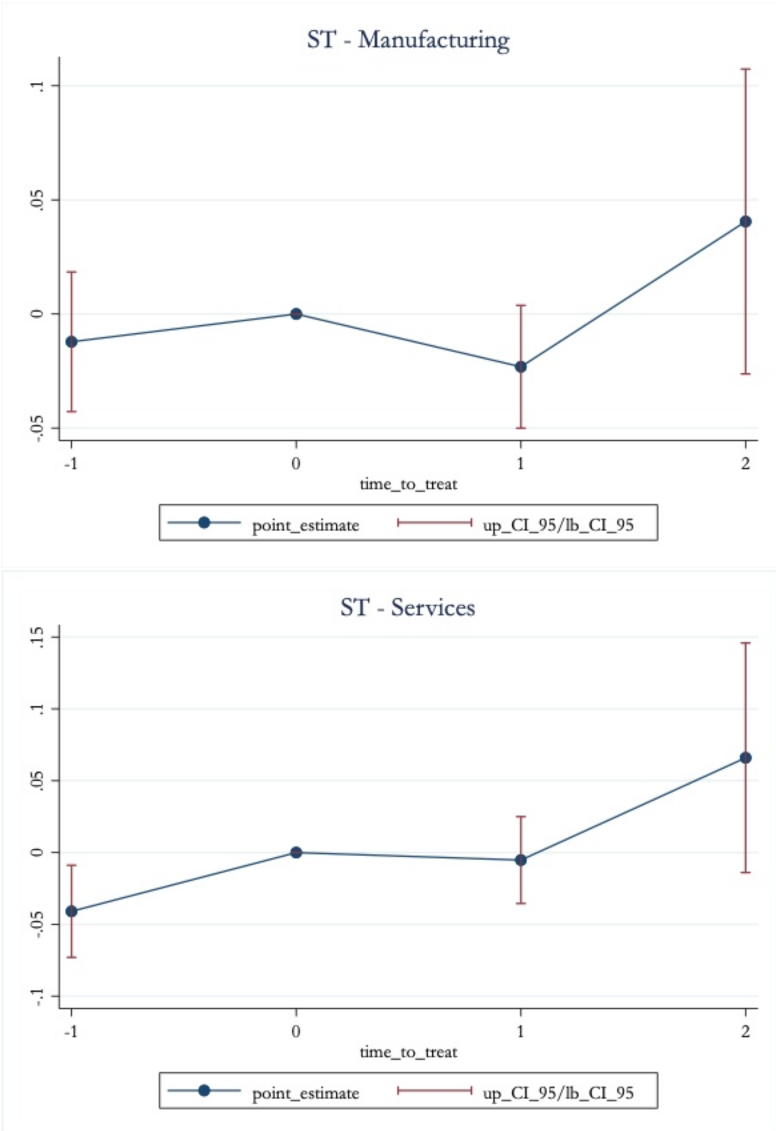


Table 1: Descriptive Statistics, Estimation Sample

	(1)	(2)	(3)	(4)	(5)	(6)
	SC		ST		Others	
	Count	Share	Count	Share	Count	Share
EC 1990						
Manufacturing	106,117	21.65%	42,252	8.62%	341,768	69.73%
Services	51,513	9.91%	23,425	4.51%	444,661	85.58%
EC 1998						
Manufacturing	91,095	17.98%	59,695	11.79%	355,737	70.23%
Services	65,027	9.13%	54,139	7.60%	593,127	83.27%
EC 2005						
Manufacturing	118,789	16.42%	77,561	10.72%	527,159	72.86%
Services	126,680	10.70%	92,376	7.80%	965,398	81.51%
EC 2013						
Manufacturing	115,546	15.27%	81,740	10.80%	559,549	73.93%
Services	196,991	12.05%	154,527	9.45%	1,283,568	78.50%

Note: SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC).

Table 2: Average Number of Enterprises per Village across Caste Groups

		1990	1998	2005	2013
Panel A: Manufacturing					
Control	SC	0.789	0.575	0.601	0.598
	ST	0.263	0.323	0.348	0.369
	Others	2.813	2.581	3.048	2.984
Treated	SC	1.042	0.776	0.842	0.667
	ST	0.334	0.415	0.472	0.427
	Others	3.068	2.877	3.562	3.456
Panel B: Services					
Control	SC	0.411	0.443	0.715	1.005
	ST	0.186	0.377	0.512	0.719
	Others	3.883	4.562	5.956	7.214
Treated	SC	0.452	0.542	0.808	1.061
	ST	0.248	0.478	0.717	0.857
	Others	3.956	4.863	6.499	8.193

Note: Treated villages here refers to those that ever got treated between 2000 and 2013. Control villages are those that never received a road during until 2013. Top 1% of outcomes winsorized. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC).

Table 3: Impact of Rural Roads on the Number of Enterprises across Caste Groups

	(1)	(2)	(3)	(4)	(5)	(6)
	Manufacturing			Services		
	Number of Enterprises owned by			Number of Enterprises owned by		
	SC	ST	Others	SC	ST	Others
Panel A: OLS						
New Road	-0.038**	0.001	0.220***	0.092***	0.047***	0.862***
	(0.018)	(0.011)	(0.044)	(0.016)	(0.012)	(0.070)
Observations	529,879	529,879	529,879	529,879	529,879	529,879
R-squared	0.545	0.546	0.616	0.551	0.603	0.681
Panel B: De Chaisemartin and d'Haultfoeuille (2020)						
New Road	-0.034	-0.014	0.139**	0.063***	0.014	0.472***
	(0.022)	(0.012)	(0.061)	(0.018)	(0.014)	(0.077)
Observations	304,932	304,932	304,932	304,932	304,932	304,932
Village FE	YES	YES	YES	YES	YES	YES
State x Year FE	YES	YES	YES	YES	YES	YES

Note: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are clustered at the village level. Panel A estimates are using OLS, and panel B estimates are using De Chaisemartin and d'Haultfoeuille (2020). New Road takes the value 1 in villages and in years where a PMGSY road is built by that year, and 0 otherwise. The outcome variables are the number of enterprises in the respective caste groups. All regressions include village fixed effects and state-year fixed effects. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC). Top 1% of outcomes winsorized.

Table 4: Impact of Rural Roads on Share of Enterprises across Caste Groups

	(1)	(2)	(3)	(4)	(5)	(6)
	Manufacturing			Services		
	Share of Enterprises owned by			Share of Enterprises owned by		
	SC	ST	Others	SC	ST	Others
Panel A: OLS						
New Road	-0.003	0.001	0.002	0.000	-0.001	0.000
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Observations	344,575	344,575	344,575	457,521	457,521	457,521
R-squared	0.561	0.715	0.646	0.543	0.742	0.671
Panel B: De Chaisemartin and d'Haultfoeuille (2020)						
New Road	-0.004	0.000	0.004	0.001	-0.001	0.000
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Observations	173,970	173,970	173,970	263,231	263,231	263,231
Village FE	YES	YES	YES	YES	YES	YES
State x Year FE	YES	YES	YES	YES	YES	YES

Note: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are clustered at the village level. Panel A estimates are using OLS, and panel B estimates are using De Chaisemartin and d'Haultfoeuille (2020). New Road takes the value 1 in villages and in years where a PMGSY road is built by that year, and 0 otherwise. The outcome variables are the number of enterprises in the respective caste groups. All regressions include village fixed effects and state-year fixed effects. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC). Top 1% of outcomes winsorized. The sample size in Table 3 is lower than Table 4 because there are many villages where the total number of manufacturing firms or service firms is zero. Division by zero yields a missing value.

Table 5: Impact of Rural Roads in the Service Sector: Robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)
	Number of Enterprises owned by			Number of Enterprises owned by		
	SC	ST	Others	SC	ST	Others
New Road	0.091***	0.049***	0.605***	0.088***	0.055***	0.614***
	(0.015)	(0.012)	(0.067)	(0.016)	(0.012)	(0.067)
Observations	529,879	529,879	529,879	529,835	529,835	529,835
R-squared	0.561	0.618	0.697	0.567	0.627	0.704
Village FE	YES	YES	YES	YES	YES	YES
State x Year FE	YES	YES	YES			
District-Time Trend	YES	YES	YES			
District x Year FE				YES	YES	YES

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of service enterprises in the respective caste groups. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC). Top 1% of outcomes winsorized.

Table 6: Impact of Rural Roads in the Service Sector: Controlling for Baseline (1991) Variables X Trend, Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
	SC	ST	Others	SC	ST	Others
New Road	0.052**	0.040***	0.432***	0.047**	0.039***	0.425***
	(0.021)	(0.014)	(0.089)	(0.021)	(0.014)	(0.090)
Observations	318,978	318,978	318,978	318,972	318,972	318,972
R-squared	0.563	0.587	0.676	0.564	0.588	0.677
Village FE	YES	YES	YES	YES	YES	YES
State × Year FE	YES	YES	YES	YES	YES	YES
Controls-Time Trend	YES	YES	YES			
Controls × Year FE				YES	YES	YES

Note: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are clustered at village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of service enterprises in the respective caste groups. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC). The controls include presence of educational facility, adult literacy centre, communication facility, market facility, power supply, and number of primary schools. Top 1% of outcomes winsorized. Data from EC 1990, 1998, 2005, and 2013.

Table 7: Impact of Rural Roads in the Services Sector: Controlling for 2001 SC/ST Population × Trend

	(1)	(2)	(3)
	SC	ST	Others
New Road	0.080***	0.031**	0.670***
	(0.016)	(0.013)	(0.072)
Observations	408,578	408,578	408,578
R-squared	0.606	0.651	0.720
Village FE	YES	YES	YES
State × Year FE	YES	YES	YES

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of enterprises in the respective caste groups. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC). All specifications include village fixed effects, state-year fixed effects, SC population share in 2001 × trend, and ST population share in 2001 × trend. Standard errors are clustered at village level. Data from Economic Censuses 1998, 2005, and 2013.

Table 8: Heterogeneity in Service Enterprises for Scheduled Caste

	(1)	(2)	(3)	(4)	(5)
Panel A: Number of Enterprises					
	Small	Single Employee	Non-hired labor	No power	Unregistered
New Road	0.092***	0.076***	0.103***	0.134***	0.132***
	(0.016)	(0.012)	(0.018)	(0.032)	(0.041)
R-squared	0.550	0.533	0.536	0.592	0.674
Panel B: Number of Enterprises					
	Large	Multiple Employees	Hired labor	Power	Registered
New Road	0.001	0.018***	0.011**	0.025***	0.013**
	(0.001)	(0.005)	(0.004)	(0.008)	(0.005)
R-squared	0.550	0.533	0.536	0.592	0.674
Observations	529,879	529,879	529,879	341,034	204,366

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of enterprises. All specifications include village fixed effects and state-year fixed effects. Power data not reported in Economic Census 2013, and registration data not reported in Economic Censuses 1990 and 2013. Top 1% of outcomes winsorized.

Table 9: Heterogeneity in Service Enterprises for Scheduled Tribe

	(1)	(2)	(3)	(4)	(5)
Panel A: Number of Enterprises					
	Small	Single Employee	Non-hired labor	No power	Unregistered
New Road	0.048***	0.036***	0.073***	0.055**	0.030
	(0.012)	(0.009)	(0.014)	(0.022)	(0.030)
R-squared	0.603	0.565	0.585	0.628	0.691
Panel B: Number of Enterprises					
	Large	Multiple Employees	Hired labor	Power	Registered
New Road	-0.000	0.011**	0.006*	0.008	0.004
	(0.000)	(0.005)	(0.003)	(0.007)	(0.003)
R-squared	0.333	0.499	0.440	0.508	0.553
Observations	529,879	529,879	529,879	341,034	204,366

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of enterprises. All specifications include village fixed effects and state-year fixed effects. Power data not reported in Economic Census 2013, and registration data not reported in Economic Censuses 1990 and 2013. Top 1% of outcomes winsorized.

Table 10: Impact of Rural Roads: Controlling for Eligible Village-State Trends, Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Number of service enterprises owned by								
	SC	ST	Others	SC	ST	Others	SC	ST	Others
New Road	0.043**	0.022	0.323***	0.028*	0.026**	0.439***	0.052***	0.007	0.403***
	(0.019)	(0.014)	(0.080)	(0.016)	(0.012)	(0.071)	(0.015)	(0.012)	(0.066)
Observations	347,509	347,509	347,509	529,877	529,877	529,877	529,881	529,881	529,881
R-squared	0.567	0.612	0.696	0.556	0.605	0.690	0.552	0.602	0.688
State × Year FE	YES	YES	YES						
District-Time Trend	YES	YES	YES						
Eligible vil.-State x Year FE				YES	YES	YES			
Eligible vil.-State-Trend							YES	YES	YES
Control group	Eligible Villages			All villages			All villages		

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of service enterprises in the respective caste groups. All specifications include village fixed effects. Top 1% of outcomes winsorized.

Table 11: Impact of rural roads on number of service enterprises: Heterogeneity based on the number of banks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	SC	ST	Others	SC	ST	Others	SC	ST	Others
New Road	0.074***	0.050***	0.567***	0.071***	0.052***	0.533***	0.092***	0.050***	0.627***
	(0.020)	(0.015)	(0.089)	(0.020)	(0.014)	(0.087)	(0.017)	(0.012)	(0.074)
New Road × Bank Branches	0.005*	0.001	0.009						
	(0.003)	(0.002)	(0.011)						
New Road × Public branches				0.006**	0.001	0.017			
				(0.003)	(0.002)	(0.012)			
New Road × Private branches							0.009	0.009	-0.027
							(0.008)	(0.008)	(0.036)
Observations	463,531	463,531	463,531	463,531	463,531	463,531	463,531	463,531	463,531
R-squared	0.556	0.609	0.692	0.556	0.609	0.692	0.556	0.609	0.692
Village FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
State × Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
District-Time trend	YES	YES	YES	YES	YES	YES	YES	YES	YES

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of enterprises in the respective caste groups. The data on bank presence is sourced from the Central Information System for Banking Infrastructure (CISBI) data from RBI, reported at the branch level. We extract 6-digit pincodes from the bank branches addresses, and then in turn map the pincodes to village names from Indian post database. To match the village names from the India Post data to the Economic Census data, we use a text matching algorithm.

Table 12: Impact of rural roads on number of service enterprises: Heterogeneity based on large primary school presence

	(1) SC	(2) ST	(3) Others
New Road	0.026	0.023	0.170**
	(0.019)	(0.015)	(0.078)
New Road × High Number of Primary Schools Dummy	0.117***	0.057**	0.890***
	(0.034)	(0.026)	(0.151)
Observations	408,291	408,291	408,291
R-squared	0.612	0.659	0.733
Village FE	YES	YES	YES
State × Year FE	YES	YES	YES
District-time Trend	YES	YES	YES

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcomes variables are the number of enterprises in the respective caste groups. High number of Primary schools dummy takes the vale 1 if the village has higher than the median number of primary schools, and 0 otherwise. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC). Top 1% of outcomes winsorized.

Table 13: Impact of rural roads on Wage Employment

	(1)	(2)	(3)	(4)
Panel A: All industries				
Share of villages exposed to roads as on year	0.143***	0.154***	0.144	0.159***
	(0.049)	(0.046)	(0.092)	(0.048)
Share of villages exposed to roads * SC/ST HH	-0.168***	-0.178***	-0.180***	-0.189***
	(0.040)	(0.038)	(0.051)	(0.040)
Observations	328,229	320,670	168,452	320,651
R-squared	0.173	0.223	0.278	0.227
Panel B: Manufacturing and Services Only				
Share of villages exposed to roads	0.177*	0.175**	0.144	0.144
	(0.092)	(0.089)	(0.092)	(0.088)
Share of villages exposed to roads * SC/ST HH	-0.178***	-0.185***	-0.180***	-0.185***
	(0.047)	(0.047)	(0.051)	(0.051)
Observations	168,473	165,520	168,452	165,498
R-squared	0.267	0.283	0.278	0.294
District FE	YES	YES	YES	YES
Year FE	YES	YES		
State time trend	YES	YES		
Industry-time trend	YES	YES		
Industry-year FE			YES	YES
State-year FE			YES	YES
Controls		YES		YES

Note: These estimations use four rounds of Employment-Unemployment data from the National Sample Survey, from the years 2004, 2004-05, 2009-10, and 2011-12. The dependent variable takes the value 1 if an individual is engaged in wage employment in the manufacturing and services sector, and 0 otherwise. The main independent variable is the share of villages exposed to roads in a district. The sample comprises working age individuals in rural areas (ages 16 to 65 years). Controls include an individual's gender, education, marital status, religion, age and age-squared. Regressions are weighted using NSS sample weights.

Appendix

Table A1: Villages Gaining a New PMGSY Road by year

New Road	1990	1998	2005	2013	Total
No	108,664	120,342	144,985	126,954	500,945
Yes	0	0	3,572	25,364	28,936
Total	108,664	120,342	148,557	152,318	529,881

Note: Each cell represents the number of villages in our estimation sample that received a new road under the PMGSY program. We use road completion dates available in the SHRUG database from Asher et al. (2021).

Table A2: Villages Gaining a New PMGSY Road by Population Size Group

	Population Group				
Year	< 250	$250 - 499$	$500 - 999$	≥ 1000	Total
2005	88	278	803	2,403	3,572
2013	805	3,128	10,095	11,336	25,364
Total	893	3,406	10,898	13,739	28,936

Note: Each cell represents the number of villages in our estimation sample that received a new road under the PMGSY program. We use population figures from Population Census of 2001 and road completion dates available in the SHRUG database from Asher et al. (2021).

Table A3: Impact of Rural Roads on Number of Enterprises across Caste Groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Manufacturing				Services			
	Number of Enterprises owned by				Number of Enterprises owned by			
	SC	ST	OBC	Others	SC	ST	OBC	Others
Panel A: OLS								
New Road	-0.031*	-0.004	0.043	0.066***	0.065***	0.037***	0.269***	0.286***
	(0.019)	(0.012)	(0.034)	(0.019)	(0.016)	(0.013)	(0.049)	(0.044)
Observations	408,652	408,652	408,652	408,652	408,652	408,652	408,652	408,652
R-squared	0.580	0.580	0.620	0.586	0.602	0.646	0.649	0.695
Panel B: De Chaisemartin and d'Haultfoeuille (2020)								
New Road	-0.033**	-0.014	0.080***	0.010	0.065***	0.013	0.307***	0.086*
	(0.015)	(0.014)	(0.030)	(0.021)	(0.023)	(0.014)	(0.053)	(0.046)
Observations	304,932	304,932	304,932	304,932	304,932	304,932	304,932	304,932

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of enterprises in the respective caste groups. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; OBC refers to Other Backward Caste; Others refer to all other castes. All regressions include village fixed effects and state-year fixed effects. Top 1% of outcomes winsorized. Data are from Economic Census 1998, 2005, and 2013.

Table A4: Impact of Rural Roads on Diversity

	(1)	(2)	(3)	(4)	(5)	(6)
	Manufacturing			Services		
	No. of 2-digit Industries			No. of 2-digit Industries		
	SC	ST	Others	SC	ST	Others
New Road	0.003	0.001	0.039***	0.046***	0.025***	0.083***
	(0.005)	(0.004)	(0.010)	(0.006)	(0.005)	(0.010)
Observations	529,879	529,879	529,879	529,879	529,879	529,879
R-squared	0.514	0.539	0.610	0.565	0.619	0.671

Note: All specifications include village and state-year FE. SE clustered at village level. Diversity based on the number of 2-digit industries.

Table A5: Impact of Rural Roads: Alternate Levels of Clustering of Standard Errors

	(1)	(2)	(3)	(4)	(5)	(6)
	Manufacturing			Services		
	Number of Enterprises owned by			Number of Enterprises owned by		
	SC	ST	Others	SC	ST	Others
Panel A: Standard error clustered at state						
New Road	-0.038	0.001	0.220	0.092***	0.047	0.862**
	(0.040)	(0.015)	(0.184)	(0.025)	(0.029)	(0.386)
Observations	529,879	529,879	529,879	529,879	529,879	529,879
R-squared	0.545	0.546	0.616	0.551	0.603	0.681
Panel B: Standard error clustered at state × year						
New Road	-0.038	0.001	0.220	0.092***	0.047*	0.862**
	(0.044)	(0.017)	(0.165)	(0.028)	(0.025)	(0.334)
Observations	529,879	529,879	529,879	529,879	529,879	529,879
R-squared	0.545	0.546	0.616	0.551	0.603	0.681

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of enterprises in the respective caste groups. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC). All regressions include village fixed effects and state-year fixed effects. Top 1% of outcomes winsorized.

Table A6: Impact of Rural Roads: Poisson Regression

	(1)	(2)	(3)	(4)	(5)	(6)
	Manufacturing			Services		
VARIABLES	SC	ST	Others	SC	ST	Others
New Road	-0.035	0.017	0.027**	0.031*	0.022	0.026***
	(0.023)	(0.026)	(0.012)	(0.017)	(0.018)	(0.009)
Observations	216,955	138,854	451,451	260,663	183,603	495,111
Village FE	YES	YES	YES	YES	YES	YES
State × Year FE	YES	YES	YES	YES	YES	YES

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of enterprises in the respective caste groups. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC). All regressions include village fixed effects and state-year fixed effects. Top 1% of outcomes winsorized.

Table A7: Impact of Rural Roads on Enterprises Per Capita

	(1)	(2)	(3)	(4)	(5)	(6)
	Manufacturing			Services		
VARIABLES	SC	ST	Others	SC	ST	Others
New Road	0.00022	0.00020	0.00017	0.00039**	0.00017	0.00056**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	405,639	225,141	471,526	405,639	225,141	471,526
R-squared	0.499	0.491	0.556	0.444	0.456	0.579
Village FE	YES	YES	YES	YES	YES	YES
State × Year FE	YES	YES	YES	YES	YES	YES

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of enterprises in the respective caste groups divided by the population of that caste group. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC). All regressions include village fixed effects and state-year fixed effects. Top 1% of outcomes winsorized.

Table A8: Impact of Rural Roads: Using a Balanced Panel of Villages

	(1)	(2)	(3)	(4)	(5)	(6)
	Manufacturing			Services		
VARIABLES	SC	ST	Others	SC	ST	Others
New Road	-0.038	0.021	0.300***	0.088***	0.066***	0.795***
	(0.032)	(0.014)	(0.076)	(0.026)	(0.015)	(0.110)
Observations	249,956	249,956	249,956	249,956	249,956	249,956
R-squared	0.540	0.482	0.610	0.541	0.531	0.672
Village FE	YES	YES	YES	YES	YES	YES
State × Year FE	YES	YES	YES	YES	YES	YES

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of enterprises in the respective caste groups. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC). Top 1% of each outcome is winsorized. Data from Economic Census of 1990, 1998, 2005, and 2013.

Table A9: Impact of Rural Roads: Drop Top Decile

	(1)	(2)	(3)	(4)	(5)	(6)
	Manufacturing			Services		
VARIABLES	SC	ST	Others	SC	ST	Others
New Road	-0.047***	-0.003	0.049	0.069***	0.039***	0.507***
	(0.016)	(0.011)	(0.036)	(0.013)	(0.012)	(0.054)
Observations	476,929	476,929	476,929	476,929	476,929	476,929
R-squared	0.541	0.547	0.569	0.525	0.609	0.615
Village FE	YES	YES	YES	YES	YES	YES
State $\times$ Year FE	YES	YES	YES	YES	YES	YES

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of enterprises in the respective caste groups. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC). Top 1% of each outcome is winsorized. The top decile of villages by population in 2001 is dropped.

Table A10: Placebo Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	Manufacturing			Services		
VARIABLES	SC	ST	Others	SC	ST	Others
New Road	0.0005623	0.0000948	-0.0007504	0.0007079	0.0002647	-0.0013458
	(0.0002359)	(0.0000903)	(0.0014902)	(0.0002152)	(0.0001145)	(0.0037607)
Observations	529,835	529,835	529,835	529,835	529,835	529,835
Village FE	YES	YES	YES	YES	YES	YES
State × Year FE	YES	YES	YES	YES	YES	YES

Note: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcomes variables are the number of enterprises in the respective caste groups. SC refers to Scheduled Caste; ST refers to Scheduled Tribe; Others refer to Other Backward Class (OBC) and General Category (GC). The coefficients and standard errors are averages across 500 placebo regressions from randomly sampling road completion years across villages in our sample.

Table A11: Impact of rural roads on number of enterprises: Heterogeneity with credit activity

	(1)	(2)	(3)	(4)	(5)	(6)
	SC	ST	Others	SC	ST	Others
New Road	0.023	0.020	0.400***	0.039*	0.038**	0.479***
	(0.022)	(0.017)	(0.094)	(0.020)	(0.016)	(0.087)
New Road × Number of credit accounts	0.125***	0.072***	0.258			
	(0.039)	(0.028)	(0.157)			
New Road × Amount Outstanding				0.058***	0.015	0.033
				(0.016)	(0.012)	(0.065)
Observations	371,418	371,418	371,418	371,418	371,418	371,418
R-squared	0.613	0.657	0.733	0.613	0.657	0.733
Village FE	YES	YES	YES	YES	YES	YES
State x Year FE	YES	YES	YES	YES	YES	YES
District-Time Trend	YES	YES	YES	YES	YES	YES

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the village level. New Road takes the value 1 in villages and in years where a PMGSY road is built, and 0 otherwise. The outcome variables are the number of service enterprises in the respective caste groups. The data on credit accounts and amount outstanding is sourced from the Basic Statistical Returns (BSR) data from the Reserve Bank of India, reported at the branch level. We aggregate credit flow to 6-digit pincodes. Whenever a village spans multiple different pincodes, we assume that the credit in the pincode with a larger area of the village is available in the village. We overlay shapefiles at the village and pincode levels to calculate this. BSR data is only available from the late 1990s, so this analysis uses Economic Censuses 1998, 2005, and 2013.