The Roads to Hospital and Healthcare Access

Tanika Chakraborty*
Yuko Mori†

Abstract

Adoption and utilization of social insurance programs remain low in India despite high out-of-pocket health costs. While these programs subsidize private tertiary healthcare, a fundamental challenge is that majority of the households, with poor healthcare access, reside in rural regions, and private hospitals are mostly located in urban regions. In this paper, we study whether rural road connectivity can mitigate the barriers to accessing healthcare services. Using the rural road construction program under the Pradhan Mantri Gram Sadak Yojana (PMGSY) program, we estimate the effect of road connectivity on utilization of Arogyasri - India's pioneering public health insurance program, introduced in erstwhile Andhra Pradesh in 2007. Using habitation level road construction data from the PMGSY portal, a population-threshold based road allocation rule, and administrative insurance claim records from Arogyasri, we find that access to a new road increases the likelihood of making an insurance claim by a large margin - 6 percentage points, from a baseline mean of 16%. At the intensive margin, our results suggest that a village registers approximately 1 more health-insurance claim when connected by a new road, once again a remarkable increase considering a baseline mean of 1 claim per village per year. These findings are also supported by our Instrumental Variable estimates, using the discontinuity created by the population-thresholds as an instrument.

Keywords: Crime-against-women, Female Labor Supply, Instrumental Variable, Alcohol Regulation

JELcodes: E24, J08, J16, J18

Declarations of interest: none

*IIM Calcutta, tanika@iimcal.ac.in
†Tsuda University, $y_m ori@tsuda.ac.jp$

1 Introduction

There are two prominent models in healthcare delivery world over - (a) direct public provision and (b) privately provided with targeted public subsidy programs. Under direct public provision, the government provides healthcare services directly by building hospitals and healthcare facilities. The second model is where healthcare is provided by the private market for a fee and the government may subsidize the fee for the poor. Social health insurance, which is becoming increasingly popular among governments in developing countries, is an example of the second model of public-private-partnership (PPP) where the private market provides healthcare and the government subsidizes health insurance for various sections of the population.

In many countries, like in India, the two frameworks co-exist. However, the fiscal stress imposed by direct provision has led more and more governments to expand social health insurance programs with a view to achieve universal health coverage. For instance, estimates from the National Health Accounts in India show two big changes - a significant decline in spending on tertiary health care by the government and a large increase in spending on social security (including social health insurance) between 2014-2020 (NHSRC, 2023). This suggests a shift of focus from the direct provision model to the PPP model, for tertiary healthcare. However, if there is not enough private healthcare supply, reducing direct public provision could reduce healthcare access. If private healthcare supply increases in response to the shift towards a PPP model, it could offset the reduced access. However, private healthcare is more likely to be urban biased. Chaudhuri and Datta (2020) estimate that more than 70% of private hospitals were located in urban regions in 2015-2016. For rural households, this uneven spatial distribution may create a barrier in utilizing the social health insurance programs even if it makes private care free.

Indeed, adoption and utilization of public health insurance programs remain low in India despite high out of pocket costs.(Debnath and Jain, 2020a) Distance to healthcare services have been shown to affect health services and outcomes even in developed countries with well developed health infrastructure.Hare and Barcus (2007) In poor settings, with weaker formal healthcare infrastructure, this problem is likely to be more critical. In particular, if most hospitals are located in urban regions and majority of the households are located in rural regions, with poor road connectivity, utilization of insurance for tertiary healthcare would remain low, even when free. Indeed, previous research has shown that the utilization of Arogyasri is largely determined by the proximity of a household to a hospital.Bhattacharya and De (2024).

In this paper we examine whether road connectivity can improve access and mitigate the demand gap for healthcare services. However, the implications of our study are relevant beyond healthcare and are more general. There are many social programs implemented every year at the national as well as the state level. At least some of these programs are likely to have complementarity effects. In this paper, we explore two different public programs. One is provided nationally and the other is provided by the state. We want to understand whether there are complementarities between the two programs. Specifically, we estimate the impact of a national rural road construction program on utilization of Andhra Pradesh's state health insurance program - Arogyasri.

The fundamental problem in asking such a question is that the governance quality in a state is likely to affect the effectiveness of both programs. What we might be picking up as complementarity of two different programs could simply reflect better governance quality which leads to better implemenatation of all public programs. To address this, we leverage a nationally determined population-based rule governing the implementation of the rural-roads program. The previous literature raises the possibility of deviation from the national mandate in Andhra Pradesh.(Asher and Novosad, 2020) In section 4, we provide evidence of the rule being followed to distribute the rural roads construction program in Andhra Pradesh.

Further following Asher and Novosad (2020) the growing literature assumes that the population threshold rule was applied at the village level. However, extensive policy documentation of the PMGSY program makes it very clear that the program was implemented at the habitation level so that the population threshold rule was applied at the habitation level. While Asher and Novosad (2020) note this condition, they assume that one village is equivalent to one habitation in India. We find that this is tenuous. On average there are 3 habitations per village. This implies a significant potential for mis-measurement of the policy indicator. Indeed, the null-effects found in Asher and Novosad (2020) could be due to attenuation bias arising from measurement error the treatment indicator for which villages received a road. For instance, consider a village with 3 habitations each with a population of 200. While each habitation does not qualify for a rural road under PMGSY, the population-threshold rule when applied to the village implies that the village is considered treated.

In this paper, we address this issue by constructing a treatment indicator which takes in to account that the program implementation was done at the habitation level. Specifically, we exploit the population-cutoff based implementation strategy which determined the eligibility of a habitation to get roads. We elaborate on this in Section 4 but in essence starting in 2001 budget was sanctioned for rural roads construction under PMGSY in unconnected habitations with more than 1000 people. Following this, habitations with more than 500 people and then habitations more than 250 people were sanctioned to receive roads provided that a large part of the habitations with larger population sizes were already connected. We use the last leg of the program to identify the effects of the rural road connectivity on Arogyasri utilization. We consider a village to be treated if there is at least one eligible habitation in the village. In a second approach, we consider a village to be treated if at least 50% of the population in the village reside in eligible habitations.

The data for this analysis comes from two sources. The administrative records of PMGSY and the administrative records of Arogyasri. The details about road construction, sanction and/or completion dates, are obtained from the PMGSY records. In addition, it also provides information on habitaion characteristics, including population size which is essential for constructing the treatment indicator. The Arogyasri claim records between 2007, the year of program inception, and 2018, last year for which we could obtain the data, are used to construct measures of utilization. We combine these two data sources with information on hospital empanelment under Arogyasri program. Road connectivity is likely to matter more depending on the distance to a hospital.

We follow two different approaches for identification of the effects. We start with a difference-in-differences framework, akin to Adukia et al. (2020), except the difference in definition of treatment assignment to a village. We estimate whether a yet-unconnected village experiences an increase in insurance utilization after receiving a road under PMGSY. We restrict our primary analysis to villages that are treated between 2008 and 2018. However, our results are robust to extending the sample to villages that are not-yet-treated.

Our second approach involves estimating a regression discontinuity framework, where the population-threshold rule at the habitation level is used to predict whether a habitation received a road between 2008-2016. We then compare utilization of Arogyasri in 2017 and 2018, across habitations that did and did not receive a road between 2008-2016. Since we do not observe number of claims at the habitation level, we construct a binary indicator for whether a village had any claim and assign all habitations within that village to have registered a claim. Conversely, if a village did not register any claim, then we consider no habitation within that village to have registered a claim.

Our DID estimates suggest that access to a new road increases the likelihood of making an insurance claim by 6 percentage points. At the intensive margin, our results suggest that a village registers approximately 1 more health-insurance claims when connected by a new road. These findings are also supported by our Instrumental Variable estimates, using the discontinuity created by the population-thresholds as an instrument in the first stage.

As found in Dupas and Jain (2021), we observe a gender gap in utilization of the insurance which remains unmitigated by road connectivity. We also find significant caste heterogenity - the impact of roads on insurance utilization is much larger for OBCs than for SCs.

Further, the utilization of private hospitals increases more than the utilization of government hospitals due to new roads. This is perhaps driven by the fact that a majority of the private hospitals are located in urban regions, and road connectivity reduces the relative cost of reaching a private hospital compared to a government hospital. We explore this possibility further using the information on distance to the nearest empanelled hospital.

Our paper contributes to two strands of recent and growing literature. First, following Asher and Novosad (2020) there is now a significant body of work that evaluates the effect of the rural roads program on various economic outcomes. **?** examine employment and income effects, Aggarwal (2021) and Dasgupta et al. (2024) examine reproductive healthcare and child health outcomes, Garg et al. (2024) and Asher et al. (2020) examine the environmental impacts. Finally, the paper that is closest to ours in this literature is Asher et al. (2018a) where they examine the effect of rural roads on access to public goods. We contribute to this body of work by estimating the effect of PMGSY on a particular public good - a state level health insurance program. Apart from Aggarwal (2021) and Dasgupta et al. (2024)

who exploit district level policy exposure, all the other papers rest on village level variation assuming that the village and habitation are identical and that the discontinuity in the policy design applied at the village level. We create village level exposure variables based on the information that the policy rule was applied at the habitation level.

The second line of work relates to understanding what influences the utilization of social health insurance programs. In India, there are only a few papers that explore why utilization remain low inspite of the programs being free. Debnath and Jain (2020a) show that caste networks matter in determining utilization by providing information. Dupas and Jain (2021) show that better women are less likely to utilize free insurance than men. Dupas and Jain (2023) examine the role of information in improving the service provided by the free insurance programs. By estimating the impact of PMGSY on utilization of Arogyasri, we contribute to this literature by establishing a potential cause of why utilization of these insurance programs remain low despite providing free healthcare services in a context with high out of pocket health expenses.

The rest of the paper is organized as follows. Section **??** provides the context of our study, including the details of the two public programs that form the backdrop of our paper. We next provide the data sources and status of the PMGSY and Arograsri programs in Section 3. 4 outlines the empirical approaches followed. 5 discusses the findings from the estimation exercise. 6 concludes.

2 Study Context

Our study is in the context of two different government programs. The first, introduced by the Government of India in 2001 involves building all-weather roads to connect all habitations in India. The second, introduced by the Government of ertwhile Andhra Pradesh in 2007, involves providing free health insurance to the people of Andhra Pradesh. We evaluate the effect of the first on utilization of the second program.

2.1 PMGSY program

The Government of India launched the PMGSY program in December 2000 with the aim to connect all rural habitations with all-weather roads. Habitations are clusters of households living in an area within a village and the the location of this area does not change over time. According to program guidelines "all-weather connectivity" means roads with adequate drainage to withstand the monsoon. This implies that the road-bed is drained effectively but it need not be paved with blacktopping or cement concrete. A gravel road can also be an all-weather road. The main aim of the program was to provide news roads such that any eligible habitation should not be more than 500 meters (1.5 km of path distance in case of Hills) away from an all-weather road or a connected habitation. The program also permitted upgrading pre-existing roads in eligible habitations where pre-existing ones were not all-weather roads (PMGSY, 2012). The stated goal was to provide all habitations with populations greater than 1000, 500 and 250, in that order of preference, with connectivity. The population, as recorded in the Census 2001, is used for for determining the eligibility of a habitation. Some exceptions to this rule was made for desert, tribal and hilly areas and districts with international borders. Further, even within rural areas, the program is restricted to roads that were formerly classified as 'Other District Roads' and 'Village Roads'.

The unit for connectivity under Pradhan Mantri Gram Sadak Yojana (PMGSY-I) is 'Habitation' and not a 'village'. However, Adukia et al. (2020) and Asher and Novosad (2020) use village as the unit of analysis implying that the population thresholds were applicable at the 'village' level. For instance, Adukia et al. (2020) writes, "We focus on villages as the unit of analysis because: many villages have only one habitation; many habitations were pooled to the village level for the purposes of the program; and little economic data is available at the habitation level."

In our data we find that about 10% of the villages have 1 habitation in a village, about 8% have 2 habitations, about 6% have 3 habitations, roughly 5% villages have 4 habitations On average there are about 6.6 habitations in a revenue-village or Gram Panchayat in our study sample, compared to an all India average of 3 per village Adukia et al. (2020). Hence, equating a village to a habitation is likely to create one-sided measurement error in the treatment assignment variable. For instance, consider a village with 3 habitations of size 200 each. While the entire village might be eligible for a road under PMGSY, none of the habitations would not be eligible while the treatment indicator will mistakenly consider this village to be treated. Thus, a unit in the control group will be considered treated. To the extent that this mis-measurement is purely a random error, based only on the distribution of habitations in a village, it is likely to lead to attenuated estimates. This could explain the modest to null effects of village roads obtained by Asher and Novosad (2020) on various developmental indicators.

2.2 Arogyasri program

In India, publicly provided healthcare has existed since its independence from the British rule. However, like in most developing countries, public healthcare has suffered from overcrowding, crumbling infrastructure, staff shortage, chronic funding shortfall, lack of equipments and medicines among others (Mavalankar and Rosenfield (2005)). While the private market co-exists, the high OOP expenditure for private healthcare makes it difficult to access for poor households. Even when available, high cost of insurance products make them unaffordable for poor households. According to the 2007 report of the planning commission, less than 10% of the Indian population was covered by any form of health insurance. Data from the most recent round of NFHS, in 2019-2021, shows that only about 40% of the households in India are covered by any health insurance (see Fig 1).

High OOP may have catastrophic financial consequences like reducing consumption or incurring high debt levels. At more than 50%, India has one of the world's highest OOP healthcare expenditure rates For more than 50 years after India gained independence in 1947, two mutually exclusive segments of the health care system co-existed, where low-income families could only access overburdened but free inpatient and outpatient care, and families with resources to pay full price access better-attended private facilities. This unequal delivery system led to substantial and rising levels of health inequity.(Joe et al., 2008)

During the last two decades, successive state and national governments have introduced free health insurance for poor households with an aim to increase healthcare access. One defining feature of such program is the public-private partnership. The aim of the free insurance program is to increase healthcare access by (a) providing access to private facilities and (b) freeing up space in public facilities. We study one such program which is India's pioneering social insurance program - The Rajiv Aarogyashri Program (RA), later renamed to Arogyasri.

The RA program's main objective was to provide health services for Below Poverty Level (BPL) families up to a value of Rs 2,0000 (roughly $300 at that time) per year for tertiary surgical and medical treatment of severe medical conditions. The program was conceived against the backdrop of at least two recent developments. First, there were many reports of distressed farmers, some committing suicide due to debt traps. This unfortunate phenomenon brought the lack of healthcare access in rural Andhra Pradesh to the fore (Ghosh (2015)). The second was a rapid proliferation of private healthcare facilities limited to urban areas (Shukla et al. (2011)). Rao et al. (2012) provide a detailed description of the program.

Here, we outline the salient features relevant to this study. First, private hospitals, government medical colleges, district hospitals, and area hospitals were eligible to enroll, provided that the private facilities were established chains and/or had at least 50 beds. Second, the scheme was implemented and supervised by a public-private partnership called the Aarogyasri Health Care Trust between state government bodies and insurance agencies Star Health and Allied Insurance. Finally, on the demand side, although the program was meant for BPL population, the eligibility cutoff was more lenient than the national definition making almost 90% of the population eligible (Debnath and Jain, 2020b). The program, launched in April 2007, entitles low-income households in erstwhile Andhra Pradesh to free tertiary care at public and empanelled private hospitals. Debnath and Jain (2020b) show that inspite of facilitating free access to private healthcare, the take-up and utilization of the program remain low and is driven to a large extent by association with a social network. In a recent study, Bhattacharya and De (2024) finds that the access to private healthcare due to the Arogyasri program increases primarily for households who reside close to a hospital. Coupled with the fact that more than 70% of the private hospitals are located in urban areas, these studues underscore the importance of distance to a hospital - either to reduce the cost of traveling and/or to reduce the cost of information acquisition.

It is in this context, that the rural road connectivity could become an important contributor to the success of other government programs. Indeed, Asher et al. (2018b) shows that public goods are less likely to be delivered to remote villages. Our question in this paper is related to Asher et al. (2018b) in that we explore the impact of rural road connectivity on the delivery of a specific public good - utilization of the public insurance program. We study the extent to which rural road connectivity, under PMGSY, enabled rural households to utilize the Arogyasri health insurance program.

3 Data

The analysis in this paper rests on two primary data sources. Administrative insurance claim records from the Arogyasri Health Insurance Program and administrative data on rural road construction under the PMGSY program for the state of Andhra Pradesh. In addition, we also combine these two data sources with the hospital network data under the Arogyasri Health Insurance Program.

3.1 Arogyasri Claims Record

We use administrative claims records from the Arogyasri program since its inception in 2007 to the 2018, the latest year for which we obtained the data. The data provides the universe of insurance claims since April 2007. It includes information on the location of the patient who underwent treatment. This allows us to identify the village of the claimant. It provides details of the care provided - specifically the medical code which identifies the specialization, either surgical or medicinal, as well as the name of the procedure. It provides claim registration, pre-authorization and claim dates as well as the date of treatment. We use the information on pre-authorization to identify whether the insurance-utilization was before or after road construction. We also use the information on age, sex and caste of the patient to explore heterogeneity in treatment effects. Finally, the data provides the details of the hospital where the treatment was done. We use information on distance from village as well as the type of hospital, public or private, to investigate implicit mechanisms.

Figure 1 and Table 11 provides summary statistics from the compiled dataset on hospital visits for the estimation sample in Table 12. We study all hospital visits between April 2007 and Decenmer 2018.

We find that number of claims in 2014 was much lower than in other years. This could be because the state of Telangana was officialy formed, by carving out parts of Andhra Pradesh, in June 2014. The division of the state may have created bureaucratic uncertainties leading to the drop in the number of claims in 2014.

3.2 Roads

We obtain the road construction records from the official website of the government of India, which provides detailed information related to the road construction projects under PMGSY.[1]

We begin by extracting habitation level information for all districts of Andhra Pradesh and Telangana. For each habitation the website provides information on population as per 2001 census as well as whether the habitation was connected by a road in 2001. In addition, the page provides information on whether a road project was sanctioned under the PMGSY program and the date of sanction. In our primary analysis we use this information to determine the treatment status of a habitation.

We also obtain data on all road projects that were completed until May 2024, at the time of our data extraction. We then map this data in to the habitation level road network data to obtain completion date for each road. Since the completion date is likely to be endogenous we use the sanction date in our main specification, providing an intent to treat effect. We also use the completion date for a robustness analysis. As can be seen in Figures 4 and 5 show the number of new roads constructed every year from the inception of the program till May 2024, and the number of habitations connected cumulatively during our study period. While some habitations were connected even before 2007, most of the construction in Andhra Pradesh took place after 2008.

Our sample comprises of habitations that were un-connected in 2001. Among these, we only retain habitations that were unconnected in 2007 since the outcome variable is observed from 2007. Next, we only consider villages that got a new connection between 2007 and 2019 and then we exclude villages which got a road in 2007 since they will not have a control comparison year.

About 40% of the habitations do not qualify for a road under the PMGSY program as they have a population size less than 250.(see Figure 3)

We conduct the analysis on all villages which received a road between 2007 and 2018. This is because for villages that have already received a road before 2007, cannot be used since they are always treated with no switch in treatment status. For the same reason we also eliminate all villages that received a road in 2007 since they are also the always treated villages given that the outcome variable, utilization of the Arogyasri insurance program, also began in 2007.

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

4 Empirical Strategy

As discussed in earlier sections, the PMGSY program was implemented using population thresholds at the habitation level. However, all other variables that we observe are at the village level. Hence, measurement of treatment assignment raises several challenges. To address them we follow two different strategies to estimate the effect of new road construction on the extent of utilization of the Arogyasri program.

We start by estimating the following model using a difference-in-differences framework:

$$\text{Number of Claims}_{vt} = \beta_0 + \beta_1 \text{Road}_{vt} + \alpha_v + \tau_t + X_{vt} + u_{vt} \tag{1}$$

where, $NumberofClaims_{vt}$ is the total number of insurance claims submitted in village v at time t. Road_{vt} denotes whether village v received a road in year t. α_v are village fixed effects and τ_t are fixed effects for each year between 2008 and 2018.

Since roads were not assigned at the village level, we use the habitation level information to aggregate to the village level Road-indicator. Specifically, we follow 2 methods. In the first, we define a village to have a road, if at least one habitation in the village crosses the threshold under PMGSY rules. Effectively, this is an underestimate of the actual extent of village connectivity. Hence, in a second approach, we construct a measure of connectivity based on whether at least 50% of the village population is connected by a road. Specifically, we define connectivity status as:

$$\text{Connected} = \begin{cases} 1, & \text{if } \frac{\sum_h p^c_{hv}}{p_v} \geq 50\% \\ 0, & \text{otherwise} \end{cases}$$

where, p^c_{hv} is the population of a connected habitation, h, in village v. p_v is the population of village v.

We include only those villages that received a road between 2008 and 2018. Thus the identification rests on switchers from not-connected to connected between 2008 and 2018. We cannot include those villages that received a road in 2007 or before since Arogyasri started in 2007 and hence we do not observe claim records before 2007. In other words, these villages are always treated. We also do not include the not-yet-treated villages, those who did not receive a road by 2018, in the main sample.

Since many villages have zero claims in a year, we also estimate a specification with a binary outcome:

$$\text{Any Claim}_{vdt} = \beta_0 + \beta_1 \text{Road}_{vdt} + \alpha_v + \tau_t + X_{vdt} + u_{vdt} \tag{2}$$

where $AnyClaim_{vdt}$ is an indicator whether there was at least one insurance claim from village v in year t.

We also employ a regression discontinuity design (RDD) to estimate the causal effects of road construction by leveraging the population threshold. As discussed in Section 2, the PMGSY programme bases the implementation of road construction on the population size of habitations, rather than villages. In our analysis, the unit of observation is the habitation. However, the claim data are reported at the village level. To address this, we apply the village-level claim status uniformly to all habitations within the village, and construct a binary dependent variable that indicates whether the claim is zero or not. This dummy variable is then used as the outcome in the regression analysis. Although the PMGSY uses population as the criterion for new road construction, this guideline is not applied perfectly in practice. Therefore, we employ a fuzzy regression discontinuity design to account for this imprecision in the assignment rule following Asher 2022:

$$\begin{aligned} \text{Any Claims}_{vt} = \beta_0 + \beta_1 \text{Connected2016}_{vh} + \beta_2 (Pop_{vh} - 250) + \\ \beta_3 \ I[\text{Pop}_{vh} \geq 250 \ ] \cdot (Pop_{vh} - 250) + \varepsilon_{hvt} \end{aligned}$$

where the first stage is given by:

$$\begin{aligned} \text{Connected2016}_h = \phi_0 + \phi_1 \ I[\text{Pop}_h \geq 250 \ ] + \phi_2 (Pop_h - 250) + \\ \phi_3 \ I[\text{Pop}_h \geq 250 \ ] \cdot (Pop_h - 250) + \varepsilon_{ht} \end{aligned}$$

where Any Claims_{ht} denotes whether at least a person residing in habitation h made a claim against a treatment in year t. Pop_h is the population of habitation h based on Census of India 2001.

We restrict the sample to habitations with populations below 500 to avoid the inclusion the next population threshold. Additionally, habitations with zero population are excluded from the analysis. Standard errors are clustered at the habitation level.Connected2016_h is a dummy variable indicating whether a habitation received a new road by 2016. Habitations that received new roads in 2017 or 2018 are excluded from the analysis.

One concern is that Andhra Pradesh may have deviated from the strict population threshold rule. The summary stats about the fraction of connected villages in each population category should help us understand this. In addition to the fact that we find a strong first stage correlation, road connection jumps at population cutoff of 250 but not at 1000 or 500.(See Figure 9)

Second, we use sanction year and not year of road completion. Sanction year is exogenous but the completion year is more likely to be endogenous. We find a gap of 3 years on average between sanction and completion and our event study also points to the an effect delay of 3-years post sanction date.

5 Results

The estimation results from equation 1 and 2 are presented in Table 12.

Column 1 presents the baseline results for β_1 in equation 2. The coefficient on road access suggests that if villages gain access to a new road, the number of claims increases by 1. This is large given that mean number of claims per village in any year is 1 (see Table 11).

In Column 2, the dependent variable is a binary indicator for whether a village had any claims in a specific year. This corresponds to estimation equation 2. The coefficient implies that access to a new road increases the likelihood that a claim is submitted by at least one person from a village by 6 percentage points. What does this coefficient mean? On the average 16 out of 100 villages make at least one insurance claim in a year. When connected by a road there is a 37% increase in the number of insurance claims so that among connected villages, 22 register at least one claim in a year.

Columns 1 and 2 include village and year fixed effects. In column 3, we also include linear time trends at the village level to account for secular increases in insurance use over time that varies by village. The results indicate a smaller increase compared to those reported in column 1 but still indicate large effect sizes.

We next explore heterogeneity in treatment effects. Gender differences in access to healthcare in India, especially tertiary services, is well known. Infact previous research has shows that the gender gap in access to healthcare goes up with distance to the hospital, perhaps because the costs of travel are higher for women due to social norms.(Kapoor et al., 2019) In columns 4 and 5 we examine gender-specific heterogeneity in treatment effects. The magnitude of the coefficient for males is larger, indicating that men may benefit more from road construction compared to women. This is even when we include delivery care which is covered by Arogyasri.

Columns 6 and 7 show that road construction also leads to an increase in the number of claims made on behalf of children, although the claims are much higher for men. We define children as those under the age of 14 and adults as those above age 15.

The main feature of the Arogyasri program was to subsidize private care. Since government care is anyway very low-cost we expect most of the action to take place in private care. Results in columns 8 and 9 show that while Arogyasri was used to get more of both public and private healthcare when a village is connected by road, the effect on private healthcare is stronger.

Columns 10 and 11 suggest that road connectivity has increased access to both emergency and non-emergency care. We define 'Emergency' when the pre-authorization date and surgery date are same for a particular claim.

We next turn to discussing the results from our regression-discontinuity(RDD) estimation. The results from the estimation of Equations **??** and **??** are reported in Table 13. In addition, Column 1 presents the results of a reduced form regression. The reduced form regression estimates the effect of the village population threshold (250 or more) on any claim, with the population dummy and the running variable as regressors. The dummy for villages with a population of 250 or more is positive and statistically significant implying that the findings in Table 12 are supported by the RDD estimates.

Column 2 shows the results of the instrumental variable regression - Equation **??**, while Column 3 reports the first stage regression results - Equation **??**. In Column 3, the dummy for villages with a population of 250 or more is positive and statistically significant at the 1 percent level, indicating that there is no concern about weak instruments. In Column 2, the coefficient is 0.954, compared to 0.8 in Table 12-column 1. In Columns 4 and 5, we account for the possibility that the effect of the running variable differs on either side of the threshold. The coefficient decreases to 0.548.

6 Conclusion

Adoption and utilization of social insurance programs remain low in India despite high out-of-pocket health costs. While these programs subsidize private tertiary healthcare, a fundamental challenge is that majority of the households, with poor healthcare access, reside in rural regions, and private hospitals are mostly located in urban regions. In this paper, we study whether rural road connectivity can mitigate the barriers to accessing healthcare services. Using the rural road construction program under the PMGSY program, we estimate the effect of road connectivity on utilization of Arogyasri - India's pioneering public health insurance program, introduced in erstwhile Andhra Pradesh in 2007.

The data for this analysis comes from two sources. The administrative records of PMGSY and the administrative records of Arogyasri. The details about road construction, sanction and/or completion dates, are obtained from the PMGSY records. In addition, it also provides information on habitaion characteristics, including population size which is essential for constructing the treatment indicator. The Arogyasri claim records between 2007, the year of program inception, and 2018, last year for which we could obtain the data, are used to construct measures of utilization.

We follow two different approaches for identification of the effects. We start with a difference-in-differences framework, akin to Adukia et al. (2020), except the difference in definition of treatment assignment to a village. Our second approach involves estimating a regression discontinuity framework, where the population-threshold rule at the habitation level is used to predict whether a habitation received a road between 2008-2016. We then compare utilization of Arogyasri in 2017 and 2018, across habitations that did and did not receive a road between 2008-2016.

The previous literature exploiting the PMGSY program for identification, rest on village level variation assuming that the village and habitation are identical and that the discontinuity in the policy design applied at the village level. However, the unit for road allocation rule followed under the Pradhan Mantri Gram Sadak Yojana (PMGSY-I) is 'Habitation' and not a 'village'. In our data we find that on average there are about 6.6 habitations in a revenue-village or Gram Panchayat, compared to an all India average of 3 per village Asher and Novosad (2020). Hence, equating a village to a habitation is likely to create one-sided measurement error in the treatment assignment variable. For instance, consider a village with 3 habitations of size 200 each. While the entire village might be eligible for a road under PMGSY, none of the habitations would not be eligible while the treatment indicator will mistakenly consider this village to be treated. Thus, a unit in the control group will be considered treated. To the extent that this mis-measurement is purely a random error, based only on the distribution of habitations in a village, it is likely to lead to attenuated estimates. This could explain the modest to null effects of village roads obtained by Asher and Novosad (2020) on various developmental indicators.

We address these issues in the construction of our treatment indicator in our DID framework and find that access to a new road increases the likelihood of making an insurance claim by a large margin - 6 percentage points, from a baseline mean of 16%. At the intensive margin, our results suggest that a village registers approximately 1 more health-insurance claim when connected by a new road, once again a remarkable increase considering a baseline mean of 1 claim per village per year. These findings are also supported by our Instrumental Variable estimates, using the discontinuity created by the population-thresholds as an instrument.

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

Asher, S., Garg, T., and Novosad, P. (2020). The ecological impact of transportation infrastructure. *The Economic Journal*, 130(629):1173–1199.

Asher, S., Nagpal, K., and Novosad, P. (2018a). The cost of distance: Geography and governance in rural india. *World Bank Working Paper*.

Asher, S., Nagpal, K., and Novosad, P. (2018b). The cost of distance: Geography and governance in rural india. *World Bank Working Paper*.

Asher, S. and Novosad, P. (2020). Rural roads and local economic development. *American economic review*, 110(3):797–823.

Bhattacharya, Titir, C. T. and De, P. (2024). Public private partnership models in social health insurance schemes- evidence from a pioneering program in india. *https://congress-files.s3.amazonaws.com/2024-07/*$DE_E SEM_2 024.pdf?VersionId = UdcD20hKB8qct5NJrHwtgBj60u7Nb.rE$.

Chaudhuri, C. and Datta, P. (2020). Analysis of private healthcare providers. *Economic & Political Weekly*, 55(44):59–64.

Dasgupta, A., Karandikar, A., and Raghav, D. (2024). Road access, fertility, and child health in rural india. *Population and Development Review*, 50(1):117–147.

Debnath, S. and Jain, T. (2020a). Social connections and tertiary health-care utilization. *Health Economics*, 29(4):464–474.

Debnath, S. and Jain, T. (2020b). Social connections and tertiary health-care utilization. *Health Economics*, 29(4):464–474.

Dupas, P. and Jain, R. (2021). Women left behind: gender disparities in utilization of government health insurance in india. Technical report, National Bureau of Economic Research.

Dupas, P. and Jain, R. (2023). Can beneficiary information improve hospital accountability? experimental evidence from a public health insurance scheme in india. *Journal of Public Economics*, 220:104841.

Garg, T., Jagnani, M., and Pullabhotla, H. K. (2024). Rural roads, farm labor exits, and crop fires. *American Economic Journal: Economic Policy*, 16(3):420–450.

Ghosh, J. (2015). Report of the commission on farmers' welfare, government of andhra pradesh.

Hare, T. S. and Barcus, H. R. (2007). Geographical accessibility and kentucky's heart-related hospital services. *Applied Geography*, 27(3-4):181–205.

Joe, W., Mishra, U. S., and Navaneetham, K. (2008). Health inequality in india: evidence from nfhs 3. *Economic and Political Weekly*, pages 41–47.

Kapoor, M., Agrawal, D., Ravi, S., Roy, A., Subramanian, S., and Guleria, R. (2019). Missing female patients: an observational analysis of sex ratio among outpatients in a referral tertiary care public hospital in india. *BMJ open*, 9(8):e026850.

Mavalankar, D. V. and Rosenfield, A. (2005). Maternal mortality in resource-poor settings: policy barriers to care. *American Journal of Public Health*, 95(2):200–203.

NHSRC (2023). National health accounts estimates for india (2019–20). *https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1919582*, April.

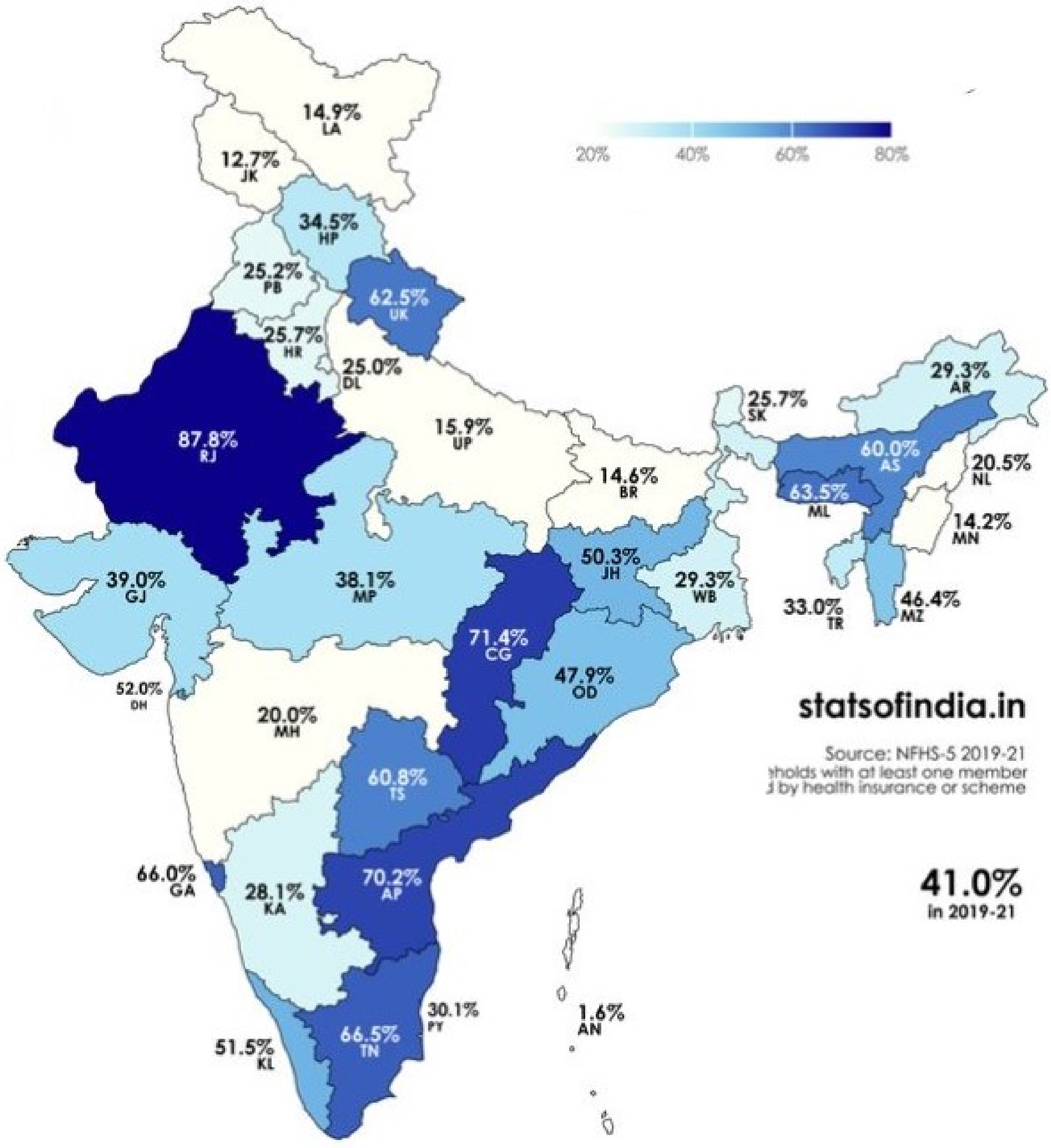
PMGSY (2012). Pradhan mantri gram sadak yojana programme guidelines. *https://omms.nic.in/ReferenceDocs/PMGSY$_G$uidelines.pdf, April*.

Rao, M., Kadam, S., Sathyanarayana, T., Shidhaye, R., Shukla, R., Ramachandra, S. S., Bandyopadhyay, S., Chandran, A., Anitha, C., Sitamma, M., et al. (2012). A rapid evaluation of the rajiv aarogyasri community health insurance scheme in andhra pradesh, india. In *BMC proceedings*, volume 6, pages 1–2. Springer.

Shukla, R., Shatrugna, V., and Srivatsan, R. (2011). Aarogyasri healthcare model: Advantage private sector. *Economic and Political Weekly*, pages 38–42.

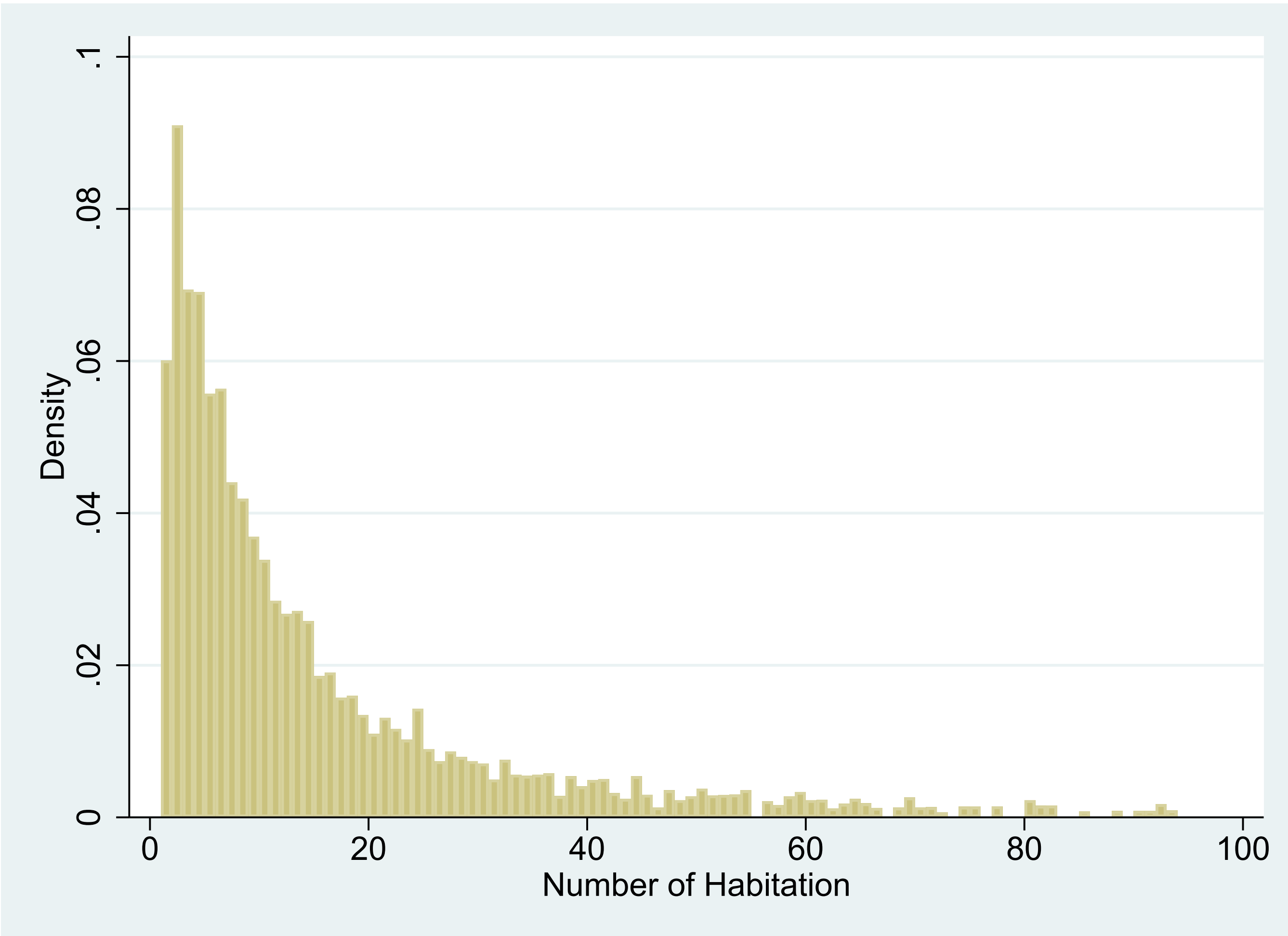
7 Tables & Graphs

Figure 1: Health insurance coverage



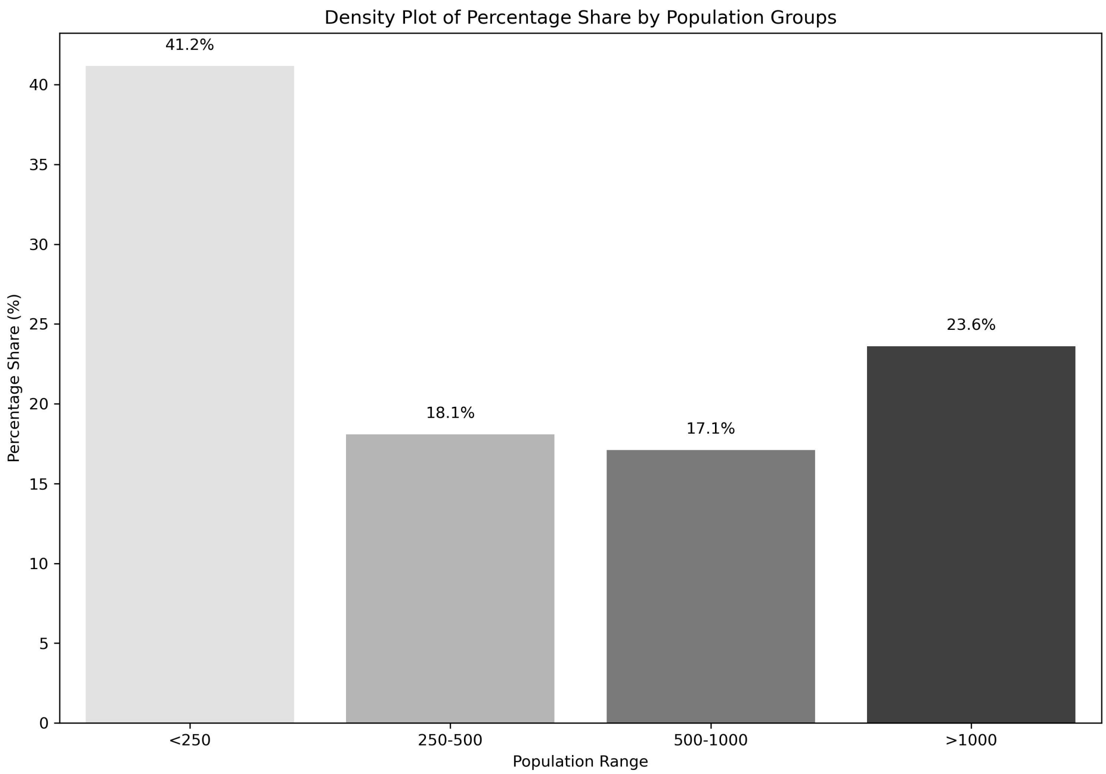
Notes: This figure presents the all-India coverage of public and private health insurance during 2019-2021.
Source: statsofindia.in

Figure 2: Distribution of habitations



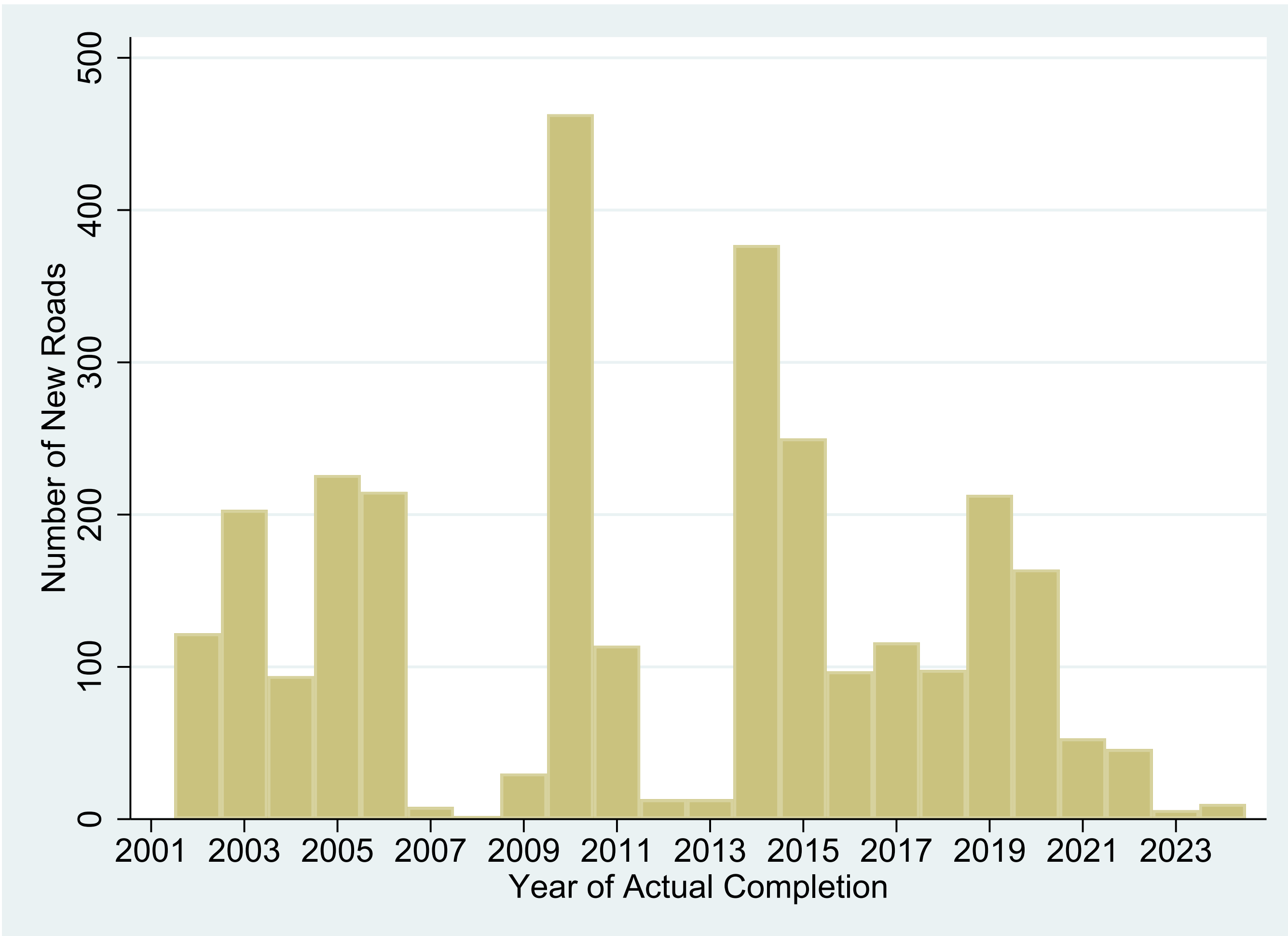
Notes: This figure shows the distribution of the number of habitations in a village in our study sample. We restrict to villages with number of habitations <= 100
Source: own calculation based on PMGSY data

Figure 3: Distribution of habitation-size



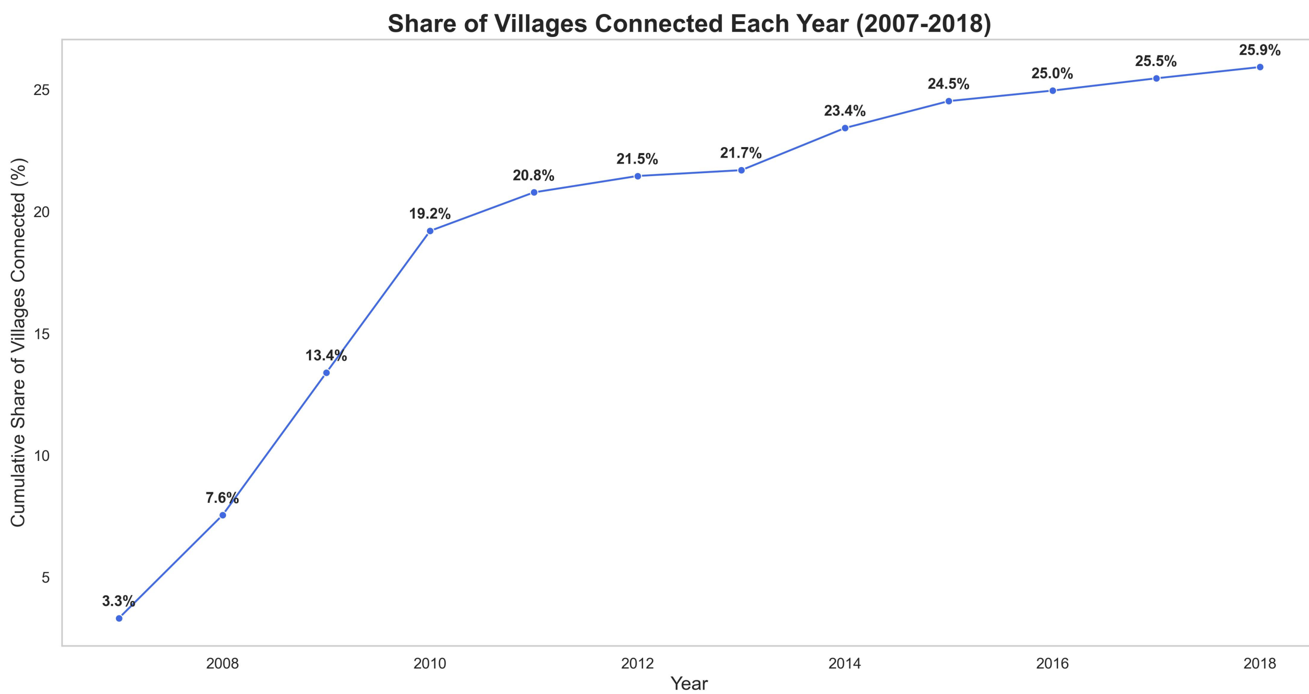
Notes: This figure shows size distribution of habitations.
Source: own calculation based on PMGSY data

Figure 4: Roads completed between 2001-2024



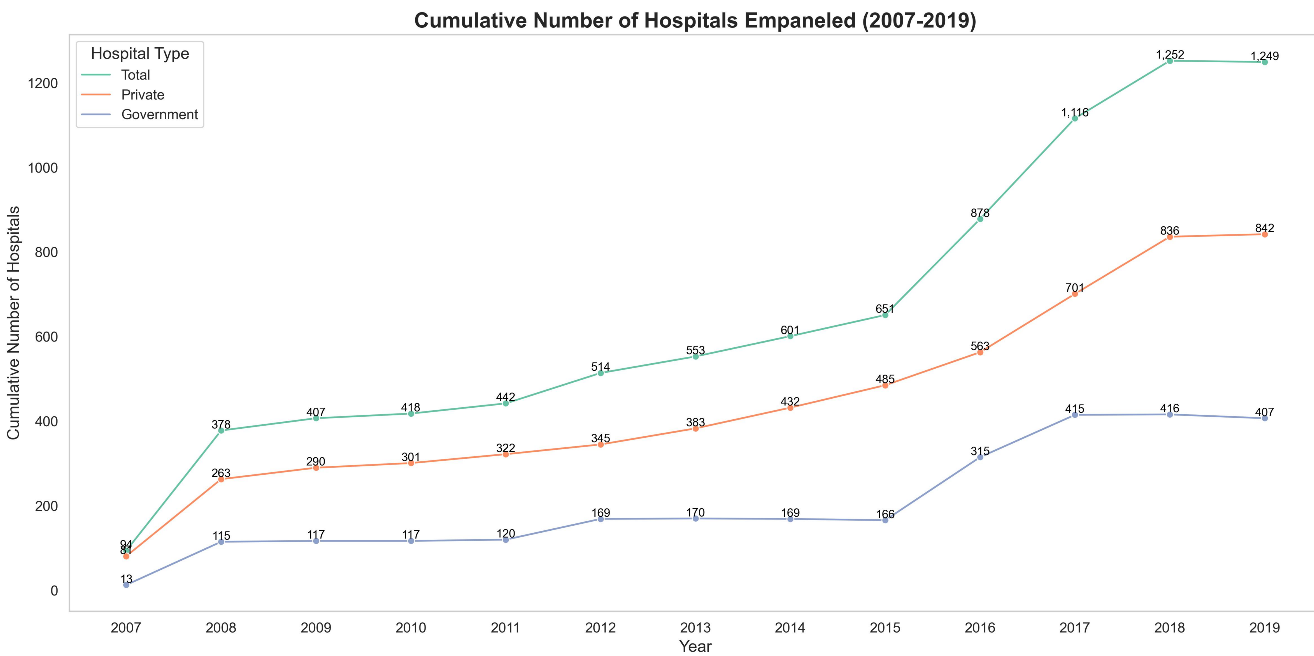
Notes: This figure shows the number of roads completed under the PMGSY program between 2001-2024. The period of our analysis is 2007-2018.
Source: own calculation based on PMGSY data

Figure 5: Fraction of villages connected between 2001-2024



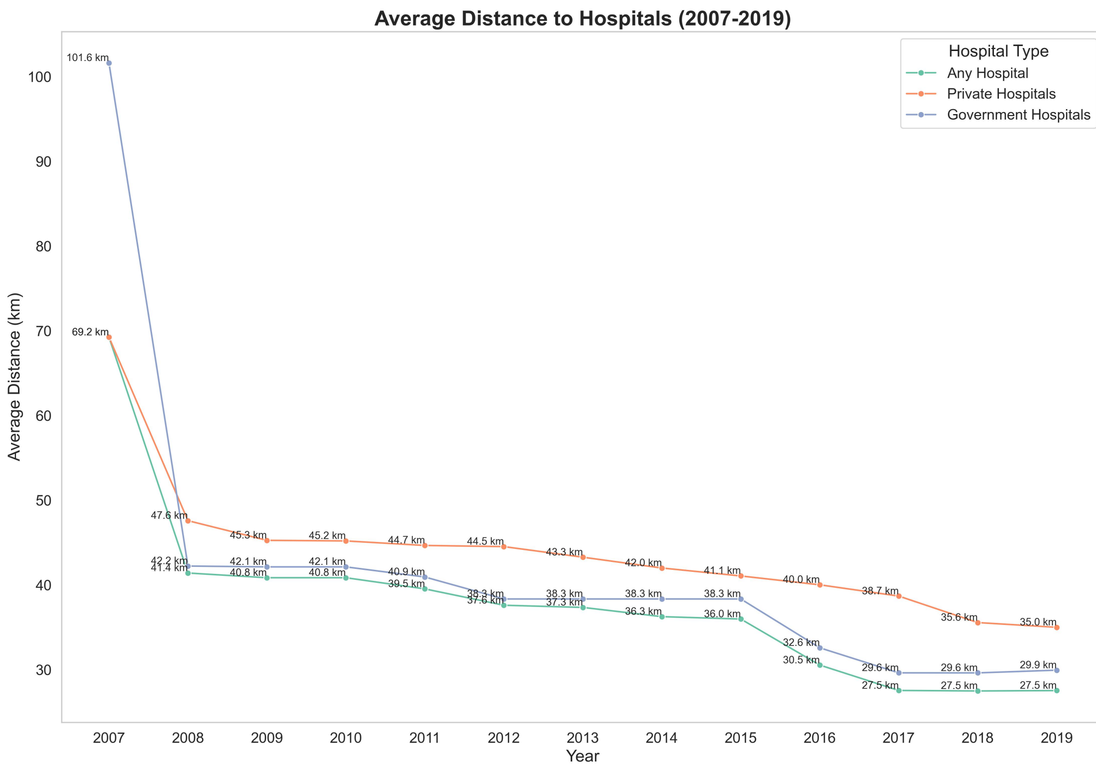
Notes: This figure shows the share of un-connected villages connected under the PMGSY program between 2001-2024. The period of our analysis is 2007-2018.
Source: own calculation based on PMGSY data

Figure 6: Number of hospitals empanelled under Arogyasri



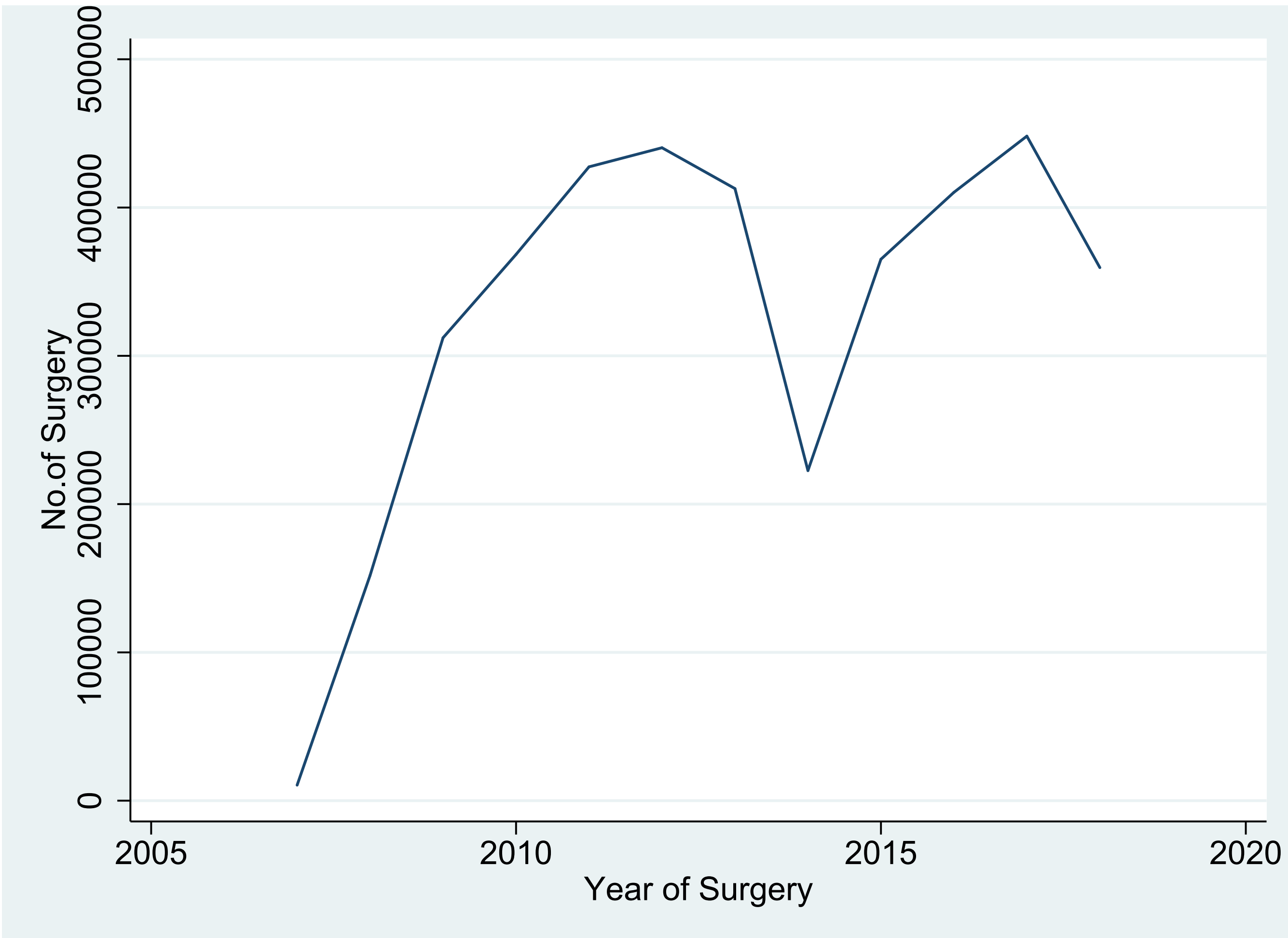
Notes: This figure shows the Number of hospitals empanelled under the Arogyasri Health Insurance program since inception.
Source: own calculation based on Arogyasri hospital records

Figure 7: Distance to the nearest Arogyasri-hospital from a village



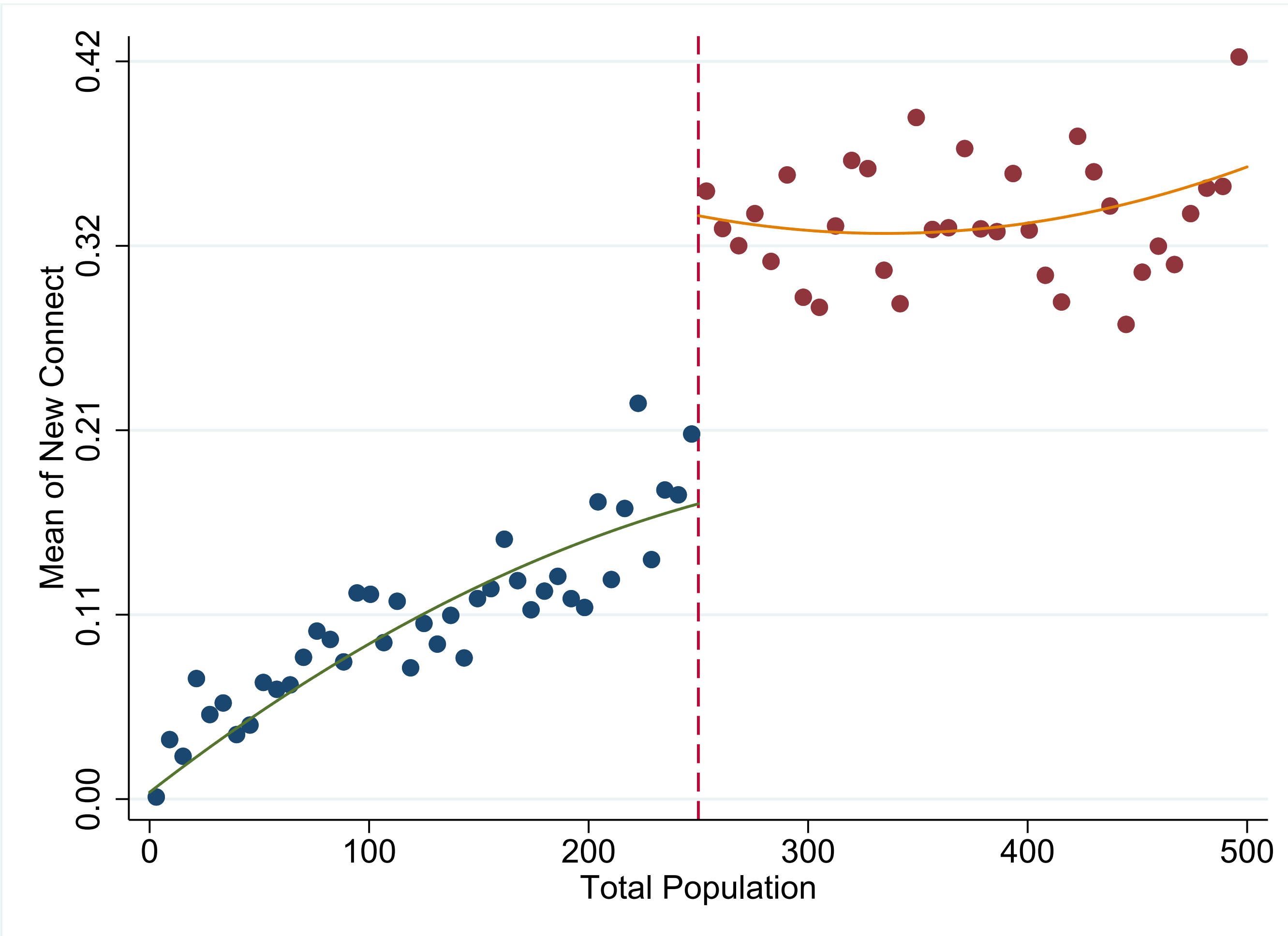
Notes: This figure shows the average distance to the nearest Arogyasri-empanelled hospital from a village over time.
Source: own calculation based on Arogyasri records and census village registry

Figure 8: Number of claims submitted under Arogyasri



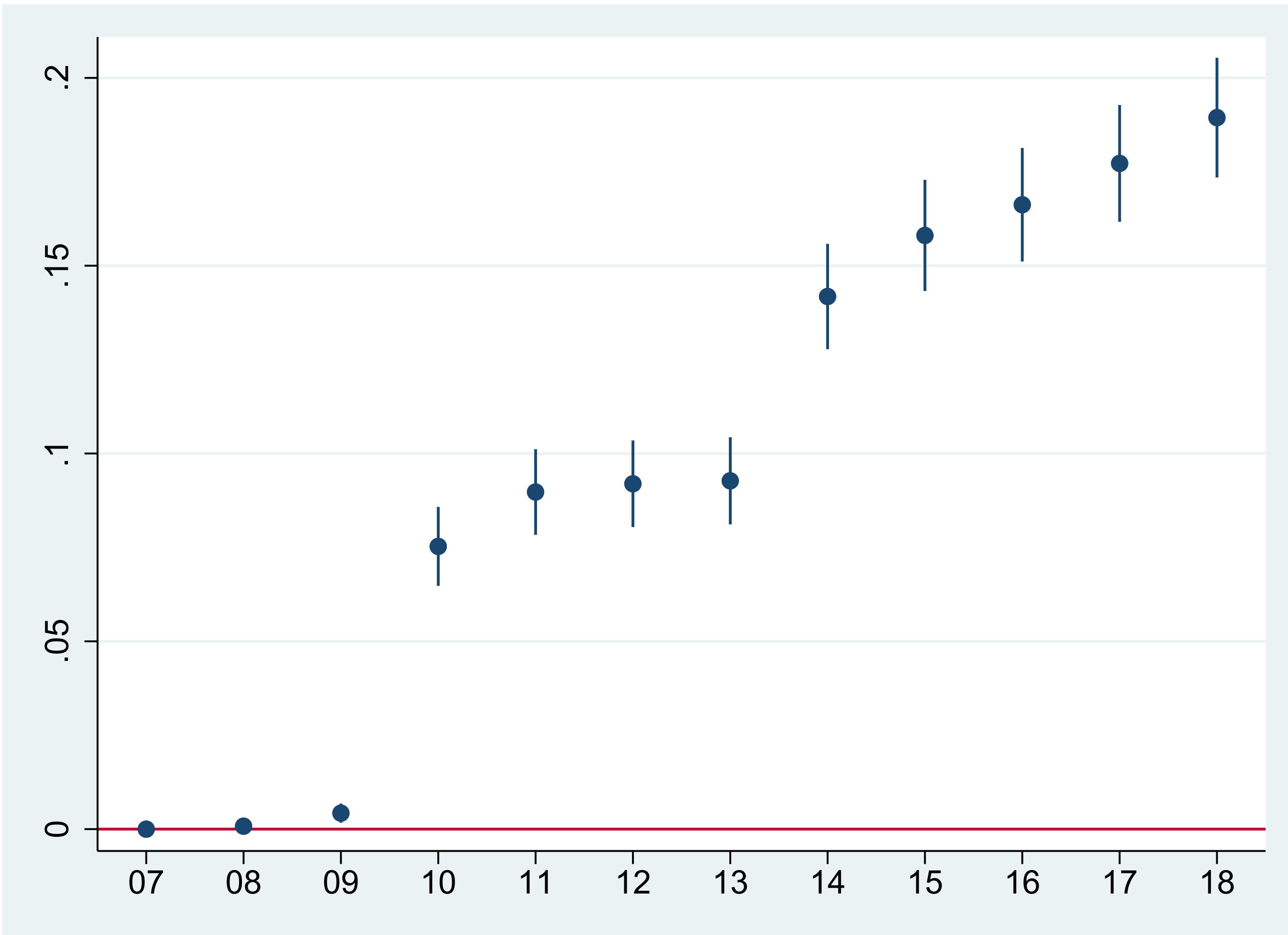
Notes: This figure shows the number of claims submitted under Arogyasri since inception until 2018.
Source: own calculation based on Arogyasri claim records

Figure 9: Probability of road construction by population size



Notes: This figure shows the share of habitations of different population sizes in the population range 1-500 that received a road between 2001 and 2018. The vertical line corresponds to population size 250.
Source: own calculation based on PMGSY data

Figure 10: Share of eligible villages connected under PMGSY between 2001-2018



Notes: This figure shows the share of eligible villages connected under PMGSY between 2001-2018. Eligibility is defined as whether a village has at least one habitation with population > 250 and < 500.
Source: own calculation based on PMGSY data

Figure 11: Village level summary

Variable	Mean	Std. dev.	Min	Max
Total Population	410.98	939.60	0	15662
Number of Claims	1.07	9.98	0	592
Dummy for claims	0.16	0.37	0	1

Figure 12: Effect of PMGSY on Arogyasri Utilization: DID

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	# of claims	D. for claims	# of claims	Male	Female	Child	Adult
Road	0.793***	0.063***	0.377**	0.485***	0.309***	0.092***	0.701***
	(0.232)	(0.015)	(0.168)	(0.144)	(0.100)	(0.025)	(0.219)
vill. F.E.	✓	✓	✓	✓	✓	✓	✓
vill. F.E*trend			✓				
Observations	32,412	32,412	32,412	32,412	32,412	32,412	32,412
R-squared	0.782	0.449	0.925	0.765	0.743	0.574	0.775

	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Government	Private	Emergency	Non-emerge.	SC	OBC	OC
Road	0.313***	0.480***	0.372***	0.421***	0.079*	0.449***	0.150**
	(0.073)	(0.175)	(0.108)	(0.133)	(0.045)	(0.124)	(0.075)
vill. F.E.	✓	✓	✓	✓	✓	✓	✓
vill. F.E*trend							
Observations	32,412	32,412	32,412	32,412	32,412	32,412	32,412
R-squared	0.608	0.794	0.757	0.775	0.715	0.717	0.731

Figure 13: Effect of PMGSY on Arogyasri Utilization: RDD

	(1) Reduced form	(2) IV	(3) 1st stage	(4) IV	(5) 1st stage
I(Population>=250)	0.100***		0.093***		0.095***
	(0.023)		(0.015)		(0.018)
New Road by 2016		1.077***		0.649**	
		(0.285)		(0.266)	
Population-250	0.000	-0.000	0.000***	-0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
(Population-250)*I(Population>=250)				0.001***	-0.000
				(0.000)	(0.000)
Observations	21,783	21,784	21,785	21,786	21,787
R-squared	0.009		0.060		0.060