For Online Publication: Internet Appendix for

Roads and Loans

Agarwal, Mukherjee, and Naaraayanan

Review of Financial Studies, 36(4), April 2023, 1508-1547

Table IA1: Covariate balance

The table presents mean values for baseline village characteristics, as recorded in 2001 Population Census. Panel A reports balance for all unconnected villages in Odisha and Uttarakhand while panel B reports balance for villages in the bank loan sample. In both panels, columns 1 and 2 present the unconditional means for villages below the treatment threshold, and villages above the treatment threshold, respectively. Column 3 presents the difference in means between villages below the treatment threshold and villages above the treatment threshold. Additionally, in panel A, column 4 shows the regression discontinuity estimate, following the main estimating equation, of the effect of being above the treatment threshold on the baseline variable and column 5 is the p-value for this estimate, using heteroskedasticity robust standard errors.

	Panel A: Unconnected villages in Odisha and Uttarakhand				
	Below	Above	Difference	RD estimate	p-value on estimate
	(1)	(2)	(3)	(4)	(5)
Primary school	0.85	0.90	-0.05	-0.01	0.34
Primary health centres	0.07	0.11	-0.04	-0.01	0.18
Telegraph office	0.00	0.00	-0.00	-0.00	0.30
Electricity	0.71	0.78	-0.07	-0.01	0.50
Scheduled caste share	49.43	45.39	4.04	-1.31	0.28
Irrigated land	0.28	0.34	-0.06	-0.03	0.41
Distance from nearest town (in kms)	32.30	30.75	1.55	-0.38	0.67
Observations	6,719	4,417			

	Panel B: Bank loan sample			
	Below	Above	Difference	p-value on difference
	(1)	(2)	(3)	(4)
Primary school	0.95	1.00	-0.05	0.16
Primary health centres	0.05	0.00	0.05	0.16
Telegraph office	0.03	0.00	0.03	0.32
Electricity	0.92	0.90	0.01	0.86
Scheduled caste share	23.53	17.09	6.43	0.16
Irrigated land	1.92	2.07	-0.15	0.77
Distance from nearest town (in kms)	27.57	29.10	-1.53	0.77
Observations	37	21		

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

Table IA2: First stage effect of road priority on PMGSY road treatment: (2009 - 2014)

The table presents first stage estimates from Equation 1 of the effect of being above the population threshold on a village's probability of receiving a road under PMGSY by 2014. The dependent variable is an indicator variable that takes on the value one if a village has received a PMGSY road before 2014. Column 1 presents results for villages within 100 of the population threshold (400-600 for the 500 threshold and 900-1100 for the 1000 threshold) while column 2 expands the sample to include villages within 150 of the population threshold. The regression specification includes state and threshold fixed effects. The sample consists of all the villages in Odisha and Uttarakhand that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. We report bootstrapped standard errors below point estimates.

	(1)	(2)
Bandwidth	±100	±150
Above cutoff	0.081***	0.065***
	(0.016)	(0.013)
Control group mean	0.13	0.12
F-statistic	25.13	24.32
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	5,537	8,246

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

Table IA3: Covariate balance for matching variables: Extensive margin

The table presents mean values for baseline village characteristics used in propensity score matching for villages in the bank loan sample and their counterfactuals constructed using propensity score matching. Specifically, we require the control group villages to be in the same block and match them on the following village-level covariates as recorded in the 2001 Population Census: the presence of a primary school, village population, the fraction of SC/ST population, and distance from the nearest town. Columns 1 and 2 present the unconditional means for villages where the bank never entered and villages where the bank enter during the sample period, respectively. Column 3 presents the difference in means, while column 4 reports the p-value on the difference.

	No Bank	Bank	Difference	p-value (difference)
	(1)	(2)	(3)	(4)
Primary school	0.97	0.91	0.06	0.25
Population	530.40	562.71	-32.31	0.49
Scheduled caste share	20.14	21.15	-1.01	0.78
Log (1+distance to nearest town)	3.08	3.10	0.02	0.91
Observations	58	58		

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

Table IA4: Impact of new roads on the Lending quantities: Unscaled dependent variable

The table presents reduced form estimates from Equation 2 of the effect of new rural roads on lending activity within the villages. Column 1 presents reduced form estimates for villages within 200 of the population threshold (300-700 for the 500 threshold and 800-1200 for the 1000 threshold) while column 2 presents reduced form estimates expanding the sample to include villages within 250 of the population threshold. The dependent variable, *NetDisburse*, is the natural logarithm of one plus total net loan amount disbursed. For each borrower, we compute the net loan amount disbursed as loan amount disbursed minus any repayment made by the end of the calendar year 2014. Our bank loan sample consists of individuals who had a loan with the bank by the end of the calendar year 2014. We include villages in Odisha and Uttarakhand that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. The specification also includes baseline borrower-level controls for age, land ownership, level of household assets, education, gender, and household income. All specifications include state and threshold fixed effects. For each regression, the outcome mean for the control group (villages with population below the threshold) is also reported. We report bootstrapped standard errors below point estimates.

Bandwidth	±200		±250	
	(1)	(2)	(3)	(4)
Above cutoff	1.530**	1.150*	1.159*	0.939*
	(0.666)	(0.601)	(0.621)	(0.557)
Age (years)		-0.032**		-0.032**
		(0.013)		(0.012)
Land		0.902***		0.916**
		(0.347)		(0.357)
Log (1+assets)		0.117*		0.108*
		(0.060)		(0.057)
School education		0.683**		0.671**
		(0.342)		(0.332)
Female		-2.772***		-2.647***
		(0.321)		(0.303)
Log (1+income)		1.358***		1.376***
		(0.238)		(0.232)
Control group mean	6.33	6.33	6.44	6.44
State fixed effects	Yes	Yes	Yes	Yes
Threshold fixed effects	Yes	Yes	Yes	Yes
Observations	1,032	1,032	1,084	1,084

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

Table IA5: First stage effect of road priority on PMGSY road treatment, robust standard errors

The table presents first stage estimates from Equation 1 of the effect of being above the population threshold on a village's probability of receiving a road under PMGSY by 2014. The dependent variable is an indicator variable that takes on the value one if a village has received a PMGSY road before 2014. Column 1 presents results for villages within 200 of the population threshold (300-700 for the 500 threshold and 800-1200 for the 1000 threshold) while column 2 expands the sample to include villages within 250 of the population threshold. The regression specification includes state and threshold fixed effects. The sample consists of all the villages in Odisha and Uttarakhand that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. We report heteroscedasticity standard errors below point estimates.

	(1)	(2)
Bandwidth	±200	±250
Above cutoff	0.068***	0.066***
	(0.011)	(0.010)
Control group mean	0.12	0.11
F-statistic	39.08	45.82
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	11,136	14,205

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

Table IA6: Impact of new roads on lending activity, robust standard errors

The table presents estimates from reduced form estimates of the effect of new rural roads on lending activity within villages. Panels A presents odds ratio from logit estimation as in panel A of Table 3 while panel B present estimates for Table 4. Panel C present estimates of Table 5 while panels D presents estimates of Table 6. Panels E presents estimates of Table 8. The dependent variable in panel A, *ExtMargin*, is an indicator variable that takes on the value one if an individual in the village received a loan from the bank. We construct the control group villages using propensity score matching. Specifically, we require the control group villages to be in the same block and match them on the following village-level covariates as recorded in the 2001 Population Census: fraction of SC/ST population, village population, presence of primary school, and distance from the nearest town. Internet Appendix Table IA 2 presents the covariate balance. The dependent variable for panels B and E, *NetDisburse/Inc*, is the net loan amount disbursed divided by household income of each borrower. For each borrower, we compute the net loan amount disbursed as loan amount disbursed minus any repayment made by the end of the calendar year 2014. We measure loan performance using two measures: (1) % Overdue amount captures the fraction of loan amount disbursed that was overdue (2) Total loan amount that was overdue. The dependent variable in columns 1 and 4 of panel C is natural logarithm of loan maturity. In columns 2 and 5, the dependent variable is Total Overdue amount while in columns 3 and 6 it is % Overdue amount. The dependent variable in panel D is the average interest rate across loans for each borrower. Our sample consists of individuals who had a loan with the bank by the end of the calendar year 2014. We include villages in Odisha and Uttarakhand that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. All specifications include state and threshold fixed effects and baseline borrower-level controls for age, land and asset ownership, education and gender. Panel A reports the odds ratio from a logit framework for estimation while all the remaining panels use Ordinary Least Squares (OLS) estimation and reports the coefficient estimates. For each regression, the outcome mean for the control group (villages with population below the threshold) is also reported. We report heteroscedasticity standard errors below point estimates.

	Panel A: Extensive margin	
Bandwidth	± 200	± 250
	(1)	(2)
Above cutoff	2.003**	1.733**
	(0.875)	(0.781)
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	93	116

	Panel B: Lending activity	
Bandwidth	±200	±250
	(1)	(2)
Above cutoff	0.025**	0.030***
	(0.012)	(0.011)
Control group mean	0.083	0.085
Controls	Yes	Yes
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	1,032	1,084

Continued...

	Panel C: Loan maturity and quality					
Bandwidth	±200			±250		
	(1)	(2)	(3)	(4)	(5)	(6)
	Ln(Maturity)	ODAmount	%OD Amount	Ln(Maturity)	ODAmount	%OD Amount
Above cutoff	-0.009	-164.872	0.057	-0.028	-184.306	-0.066
	(0.020)	(205.843)	(0.350)	(0.019)	(189.697)	(0.360)
Control group mean	1.11	104.7	0.12	1.11	100.6	0.12
Loanpurpose fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Threshold fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	630	630	630	665	665	665

	Panel D: Interest rates	
Bandwidth	± 200	± 250
	(1)	(2)
Above cutoff	-0.002	-0.005
	(0.006)	(0.005)
Control group mean	0.15	0.15
Loanpurpose fixed effects	Yes	Yes
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	630	665

	Panel E: Lending quantity by loan type			
	Productive Loans		Non-Productive Loans	
	(1)	(2)	(3)	(4)
	±200	±250	±200	±250
Above cutoff	0.043***	0.045***	-0.044***	-0.038***
	(0.011)	(0.010)	(0.011)	(0.009)
Control group mean	0.047	0.047	0.066	0.067
State fixed effects	Yes	Yes	Yes	Yes
Threshold fixed effects	Yes	Yes	Yes	Yes
Observations	1,032	1,084	1,032	1,084

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

Table IA7: Impact of new roads on lending activity, stratified bootstrapping

The table presents estimates from reduced form estimates of the effect of new rural roads on lending activity within villages. Panels A presents odds ratio from logit estimation as in panel A of Table 3 while panel B present estimates for Table 4. Panel C present estimates of Table 5 while panels D presents estimates of Table 6. Panels E presents estimates of Table 7. The dependent variable in panel A, *ExtMargin*, is an indicator variable that takes on the value one if an individual in the village received a loan from the bank. We construct the control group villages using propensity score matching. Specifically, we require the control group villages to be in the same block and match them on the following village-level covariates as recorded in the 2001 Population Census: fraction of SC/ST population, village population, presence of primary school, and distance from the nearest town. Internet Appendix Table IA 2 presents the covariate balance. The dependent variable for panels B and E, *NetDisburse/Inc*, is the net loan amount disbursed divided by household income of each borrower. For each borrower, we compute the net loan amount disbursed as loan amount disbursed minus any repayment made by the end of the calendar year 2014. We measure loan performance using two measures: (1) % Overdue amount captures the fraction of loan amount disbursed that was overdue (2) Total loan amount that was overdue. The dependent variable in columns 1 and 4 of panel C is natural logarithm of loan maturity. In columns 2 and 5, the dependent variable is Total Overdue amount while in columns 3 and 6 it is % Overdue amount. The dependent variable in panel D is the average interest rate across loans for each borrower. Our sample consists of individuals who had a loan with the bank by the end of the calendar year 2014. We include villages in Odisha and Uttarakhand that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. All specifications include state and threshold fixed effects and baseline borrower-level controls for age, land and asset ownership, education and gender. Panel A reports the odds ratio from a logit framework for estimation while all the remaining panels use Ordinary Least Squares (OLS) estimation and reports the coefficient estimates. For each regression, the outcome mean for the control group (villages with population below the threshold) is also reported. Bootstrap samples are taken independently within each village and bootstrapped standard errors are reported below point estimates.

	Panel A: Extensive margin	
Bandwidth	± 200	± 250
	(1)	(2)
Above cutoff	2.003***	1.733***
	(0.000)	(0.000)
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	93	116

	Panel B: Lending activity	
Bandwidth	±200	±250
	(1)	(2)
Above cutoff	0.025***	0.030***
	(0.010)	(0.009)
Control group mean	0.083	0.085
Controls	Yes	Yes
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	1,032	1,084

Continued...

	Panel C: Loan maturity and quality					
Bandwidth	±200			±250		
	(1)	(2)	(3)	(4)	(5)	(6)
	Ln(Maturity)	ODAmount	%OD Amount	Ln(Maturity)	ODAmount	%OD Amount
Above cutoff	-0.009	-164.872	0.057	-0.028*	-184.306	-0.066
	(0.017)	(184.396)	(0.349)	(0.016)	(175.312)	(0.347)
Control group mean	1.11	104.7	0.12	1.11	100.6	0.12
Loanpurpose fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Threshold fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	630	630	630	665	665	665

	Panel D: Interest rates	
Bandwidth	± 200	± 250
	(1)	(2)
Above cutoff	-0.002	-0.005
	(0.005)	(0.005)
Control group mean	0.15	0.15
Loanpurpose fixed effects	Yes	Yes
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	630	665

	Panel E: Lending quantity by loan type			
	Productive Loans		Non-Productive Loans	
	(1)	(2)	(3)	(4)
	±200	±250	±200	±250
Above cutoff	0.043***	0.045***	-0.044***	-0.038***
	(0.010)	(0.009)	(0.008)	(0.007)
Control group mean	0.047	0.047	0.066	0.067
State fixed effects	Yes	Yes	Yes	Yes
Threshold fixed effects	Yes	Yes	Yes	Yes
Observations	1,032	1,084	1,032	1,084

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

Table IA8: Impact of new roads on lending activity, robustness

The table presents robustness on the reduced form estimates from Equation 2 of the effect of new rural roads on lending activity within these villages. Panel A present results wherein we drop four villages in Uttarakhand with habitations. Panel B presents results with same slopes on either side of the population threshold and different intercept around the cutoff while panel C allows for same slope and intercept around the population threshold. Panel D presents results from baseline Table 4 without winsorizing our dependent variable while panel E presents results restricting the sample to new borrowers. In all panels, the dependent variable, *NetDisburse/Inc*, is the net loan amount disbursed divided by household income of each borrower. For each borrower, we compute the net loan amount disbursed as loan amount disbursed minus any repayment made by the end of the calendar year 2014. Our sample consists of individuals who had a loan with the bank by the end of the calendar year 2014. We include villages in Odisha and Uttarakhand that did not have paved roads at the start of our sample as recorded in the 2001 Population Census. All specifications include state and threshold fixed effects and baseline borrower-level controls for age, land ownership, household assets, education and gender. For each regression, the outcome mean for the control group (villages with population below the threshold) is also reported. We report bootstrapped standard errors below point estimates.

	Panel A: Drop villages with habitations	
Bandwidth	±200	±250
	(1)	(2)
Above cutoff	0.024**	0.029***
	(0.012)	(0.011)
Control group mean	0.082	0.084
Controls	Yes	Yes
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	1,019	1,071

	Panel B: Same slope and different intercept around the cutoff	
Bandwidth	± 200	± 250
	(1)	(2)
Above Cutoff	0.028**	0.030***
	(0.012)	(0.011)
Control group mean	0.083	0.085
Controls	Yes	Yes
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	1,032	1,084

Continued...

	Panel C: Same slope and intercept around the cutoff	
Bandwidth	± 200	± 250
	(1)	(2)
Above Cutoff	0.028**	0.030***
	(0.012)	(0.011)
Control group mean	0.083	0.085
Controls	Yes	Yes
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	1,032	1,084

	Panel D: No winsorization	
Bandwidth	±200	±250
	(1)	(2)
Above cutoff	0.025**	0.030***
	(0.012)	(0.011)
Control group mean	0.084	0.086
Controls	Yes	Yes
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	1,032	1,084

	Panel E: Ruling out evergreening	
Bandwidth	±200	±250
	(1)	(2)
Above cutoff	0.023*	0.028**
	(0.013)	(0.011)
Control group mean	0.079	0.081
Controls	Yes	Yes
State fixed effects	Yes	Yes
Threshold fixed effects	Yes	Yes
Observations	959	1,005

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

Table IA9: Impact of new roads by borrower characteristics, robustness

The table presents reduced form estimates of the heterogeneous effects of new rural roads by borrower characteristics for the sample of villages. Columns 1 through 3 present reduced form estimates for villages within 200 of the population threshold (300-700 for the 500 threshold and 800-1200 for the 1000 threshold) while columns 4 through 6 present reduced form estimates expanding the sample to include villages within 250 of the population threshold. The dependent variable in column 1 and 4 is the net loan amount disbursed divided by the household income for each borrower. The dependent variable in columns 2 and 5 is the fraction of the loan amount disbursed that was overdue while in columns 3 and 6 it is the average interest rate across loans for each borrower. For each borrower, we compute the net loan amount disbursed as the loan amount disbursed minus any repayment made by the end of the calendar year 2014. We interact Above cutoff with the following characteristics: *Age (years)* a continuous variable that captures the age of the borrower in years at the time of opening the bank account, *Low wealth* a dummy variable that takes a value of one if a borrower has below-median assets as well as at least one other independent piece of confirming evidence on their financial status, i.e., either below-median landholdings (in acres) or below-median jewellery (in grams) and zero otherwise, *School education* a dummy variable that takes the value of one if the borrower has ever attended any school class at the time of opening a bank account or zero otherwise, *SC/ST/OBC* an indicator for the whether the borrower belongs to any of the minority sub-groups (Scheduled caste, Scheduled Tribe, or Other Backward Castes), and *Female* an indicator for whether the gender of the borrower is female. Our sample consists of individuals from the sample of villages in Odisha and Uttarakhand who had a loan with the bank by the end of the calendar year 2014. All specifications include loan purpose, state, and threshold fixed effects. For each regression, the outcome mean for the control group (villages with population below the threshold) is also reported. We report bootstrapped standard errors below point estimates.

	± 200			± 250		
	(1) Loan Amount	(2) %ODAmount	(3) AvgIntRate	(4) Loan Amount	(5) %ODAmount	(6) AvgIntRate
Above cutoff	-0.016	0.687	0.010	-0.009	0.576	0.006
	(0.026)	(1.295)	(0.017)	(0.025)	(1.200)	(0.016)
Age (years)	-0.001**	0.000	-0.000	-0.000	0.002	-0.000
	(0.000)	(0.004)	(0.000)	(0.000)	(0.004)	(0.000)
Low wealth	0.006	0.252	-0.000	0.005	0.237	-0.001
	(0.007)	(0.256)	(0.003)	(0.007)	(0.245)	(0.003)
School education	-0.006	-0.922	0.000	-0.004	-0.809	-0.000
	(0.007)	(0.812)	(0.004)	(0.006)	(0.712)	(0.004)
SC/ST/OBC	-0.002	0.160	0.000	-0.002	0.142	0.000
	(0.008)	(0.185)	(0.003)	(0.008)	(0.166)	(0.003)
Female	-0.015***	0.286	0.006*	-0.016***	0.254	0.005*
	(0.005)	(0.295)	(0.003)	(0.005)	(0.261)	(0.003)
Above cutoff x Age (Years)	0.001	-0.019	-0.000	0.000	-0.020	-0.000
	(0.000)	(0.021)	(0.000)	(0.000)	(0.020)	(0.000)
Above cutoff × Low wealth	0.019*	-0.172	-0.001	0.018*	-0.184	-0.001
	(0.010)	(0.205)	(0.004)	(0.010)	(0.191)	(0.004)
Above cutoff x School education	0.014	0.892	0.007	0.012	0.741	0.008
	(0.011)	(0.774)	(0.009)	(0.010)	(0.678)	(0.008)
Above cutoff x SC/ST/OBC	-0.011	-0.556	-0.006	-0.008	-0.504	-0.007
	(0.011)	(0.388)	(0.005)	(0.010)	(0.357)	(0.005)
Above cutoff x Female	0.000	-0.464	-0.009*	0.001	-0.441	-0.008
	(0.010)	(0.347)	(0.005)	(0.010)	(0.316)	(0.005)
Control group mean	0.083	0.12	0.15	0.085	0.12	0.15
Loanpurpose fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Threshold fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,032	630	630	1,084	665	665

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