

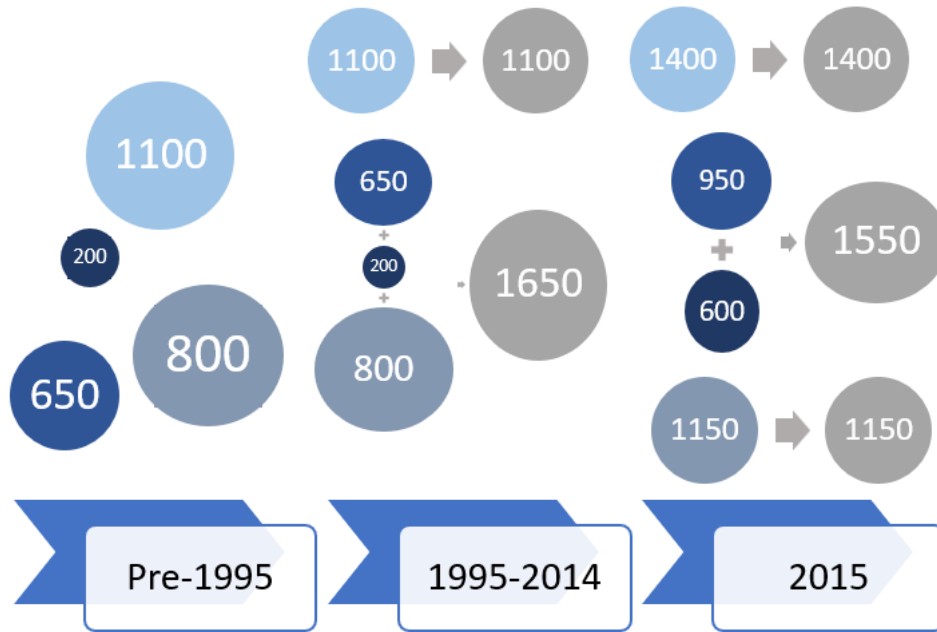
Appendix: For Online Publication

Polity size and local government performance: evidence from India

Veda Narasimhan and Jeffrey Weaver

A Appendix Figures

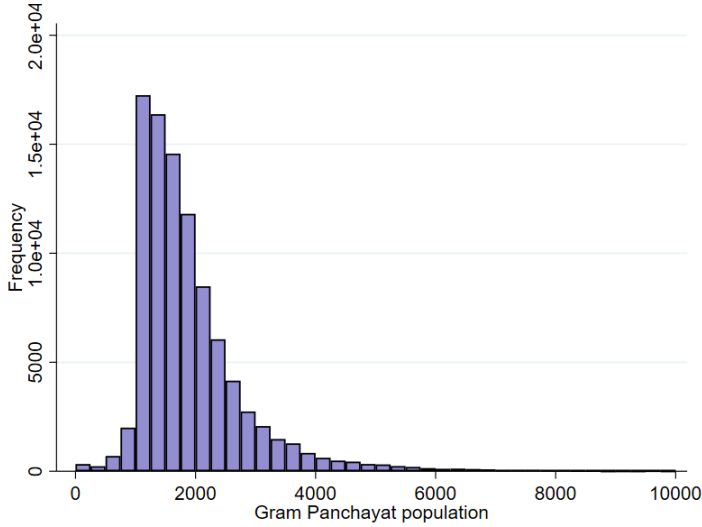
Figure A1: Village Distribution Timeline



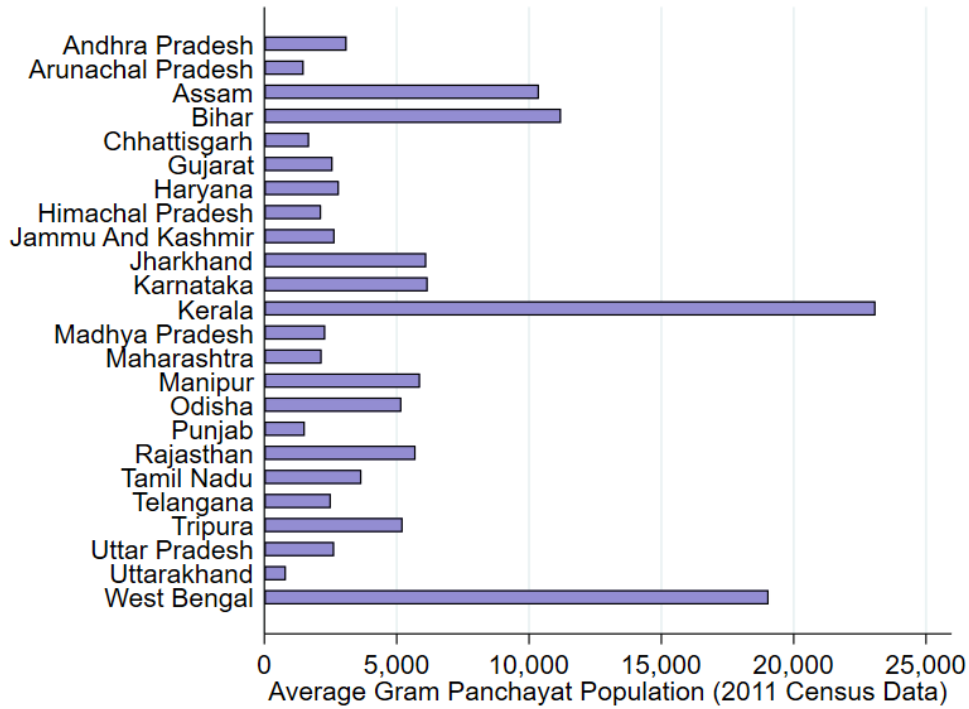
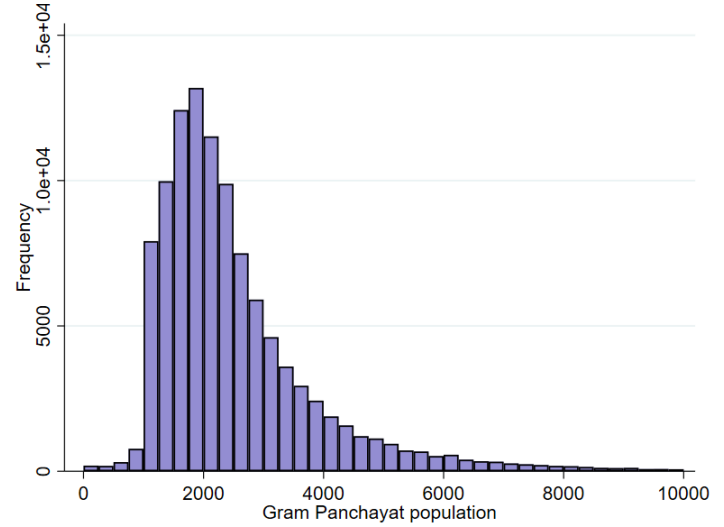
Notes: This figure provides an example of how villages might be divided into GPs following the delimitation rules. The first section shows four villages prior to 1995 along with their populations. Following the implementation of the reform in the 1995 GP elections, these villages were divided into two GPs based on population size. Villages with a population exceeding 1000 were granted separate GPs, while those with fewer residents shared a common GP – in this case, the one village with a population of 1100 is split into its own GP, while the other three are joined in a second GP. In 2015, a second administrative demarcation occurred, following the same initial guidelines – because the fourth village was above the threshold, it is split into its own GP, while the second and third villages remain together.

Figure A2: Gram Panchayat population distribution

(a) Uttar Pradesh GP population distribution (1995-2015)



(b) Uttar Pradesh GP population distribution (2015-2021)

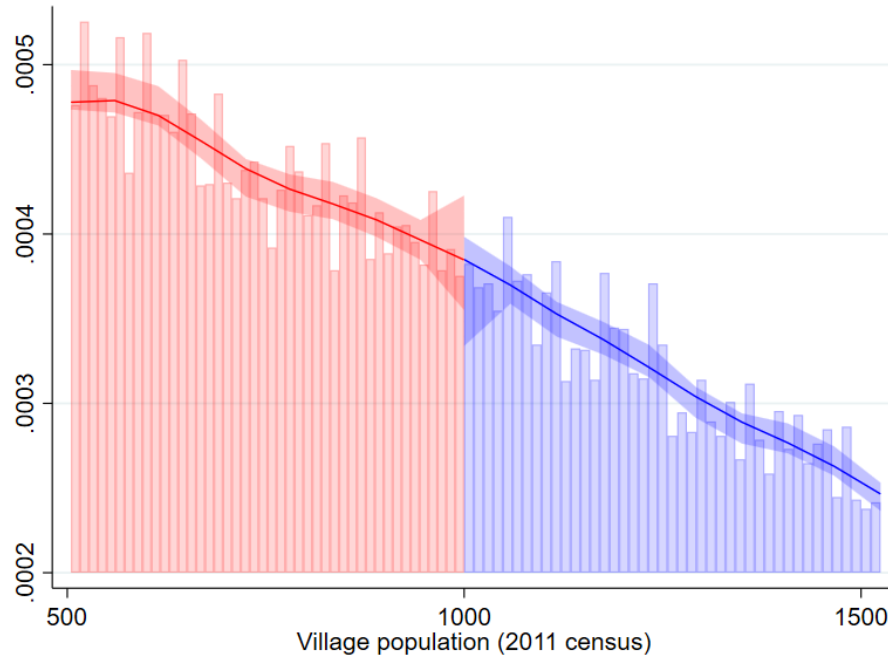
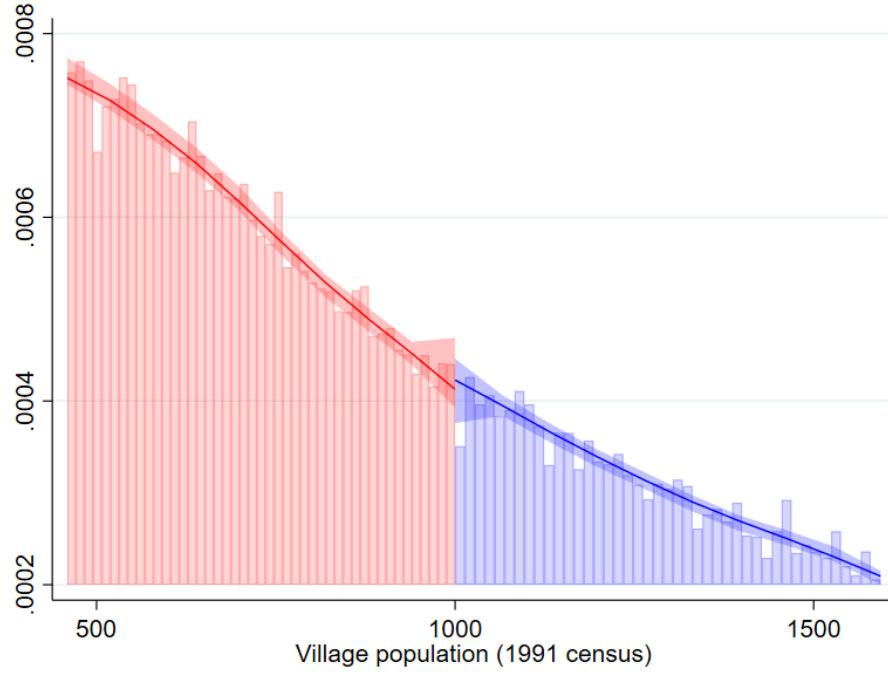


(c) Average size of gram panchayats across states

Notes: Panels (a) and (b) plot the distribution of gram panchayat populations in Uttar Pradesh in 1995 and 2015. Figure (a) uses the 1991 India census data for gram panchayat population for the gram panchayats that existed between 1995 and 2015. Figure (b) uses the 2011 Indian census population data for the gram panchayats that came into existence in 2015. Large GPs of populations significantly above 1000 typically occur because there is a large village with a population well above 1000 in that GP. Figure (c) plots the average population of gram panchayats by state. For figure (c), gram panchayats are defined using the 2021 Local Government Directory, and population is based on the 2011 Indian census (unadjusted for population growth between 2011 and 2021).

Figure A3: Distribution of the running variable

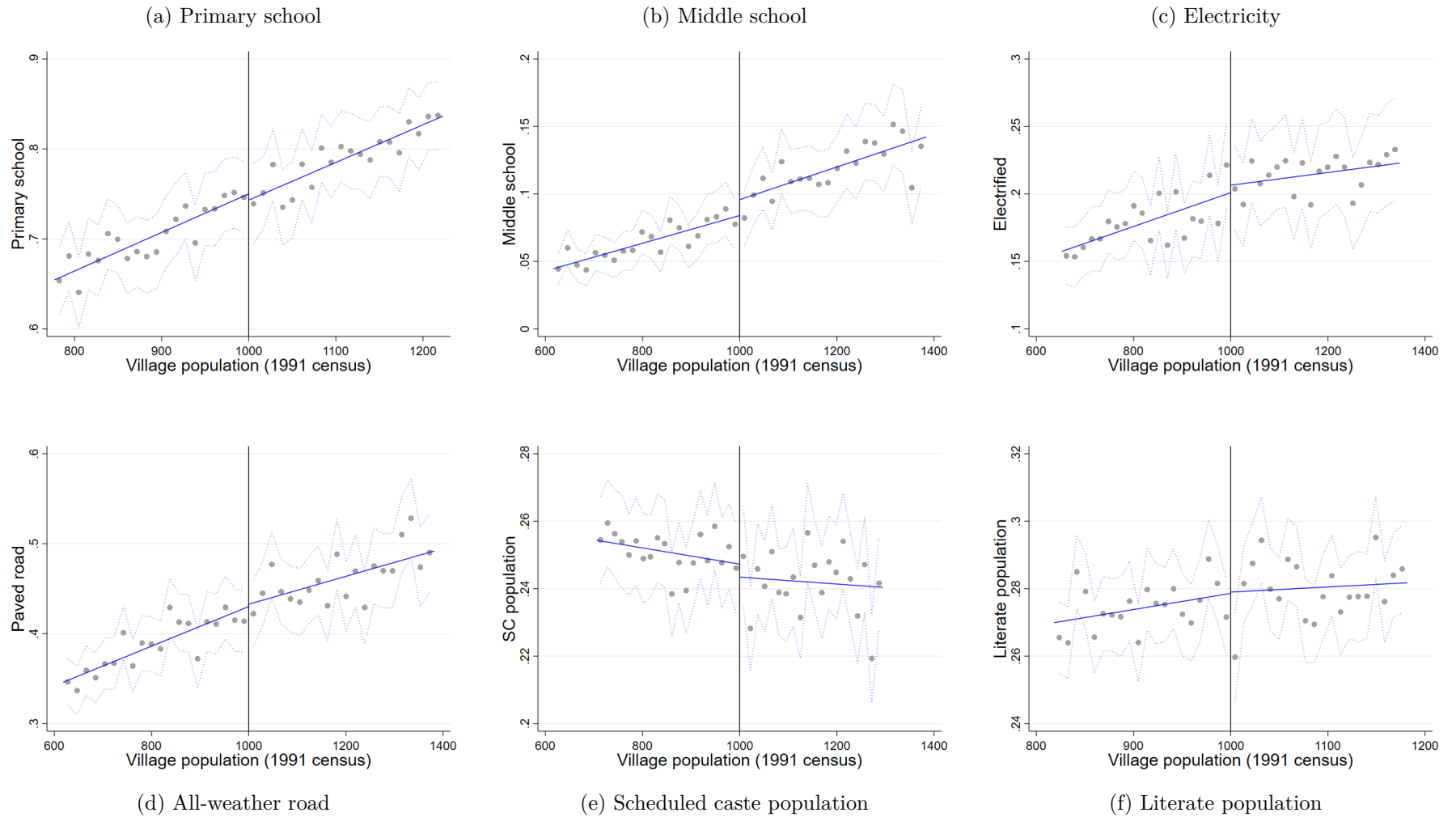
(a) First stage (1991 census)



(b) First stage (2011 census)

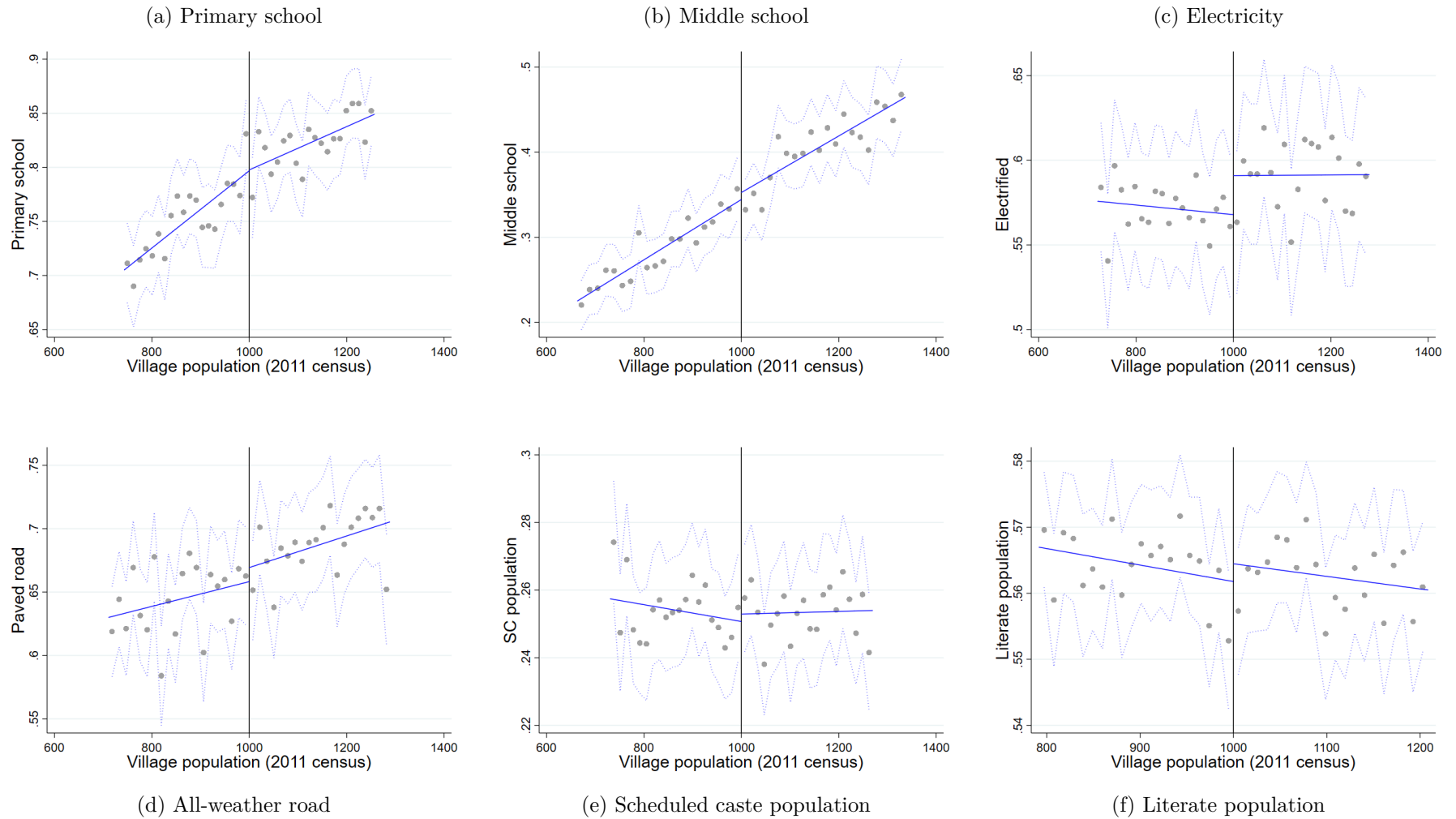
Notes: These figures plot the distribution of village population around the relevant population thresholds and a non-parametric regression for each half of the distribution in order to test for a discontinuity in population density around the threshold (Cattaneo et al., 2020).

Figure A4: Balance on baseline village characteristics (1991 discontinuity)



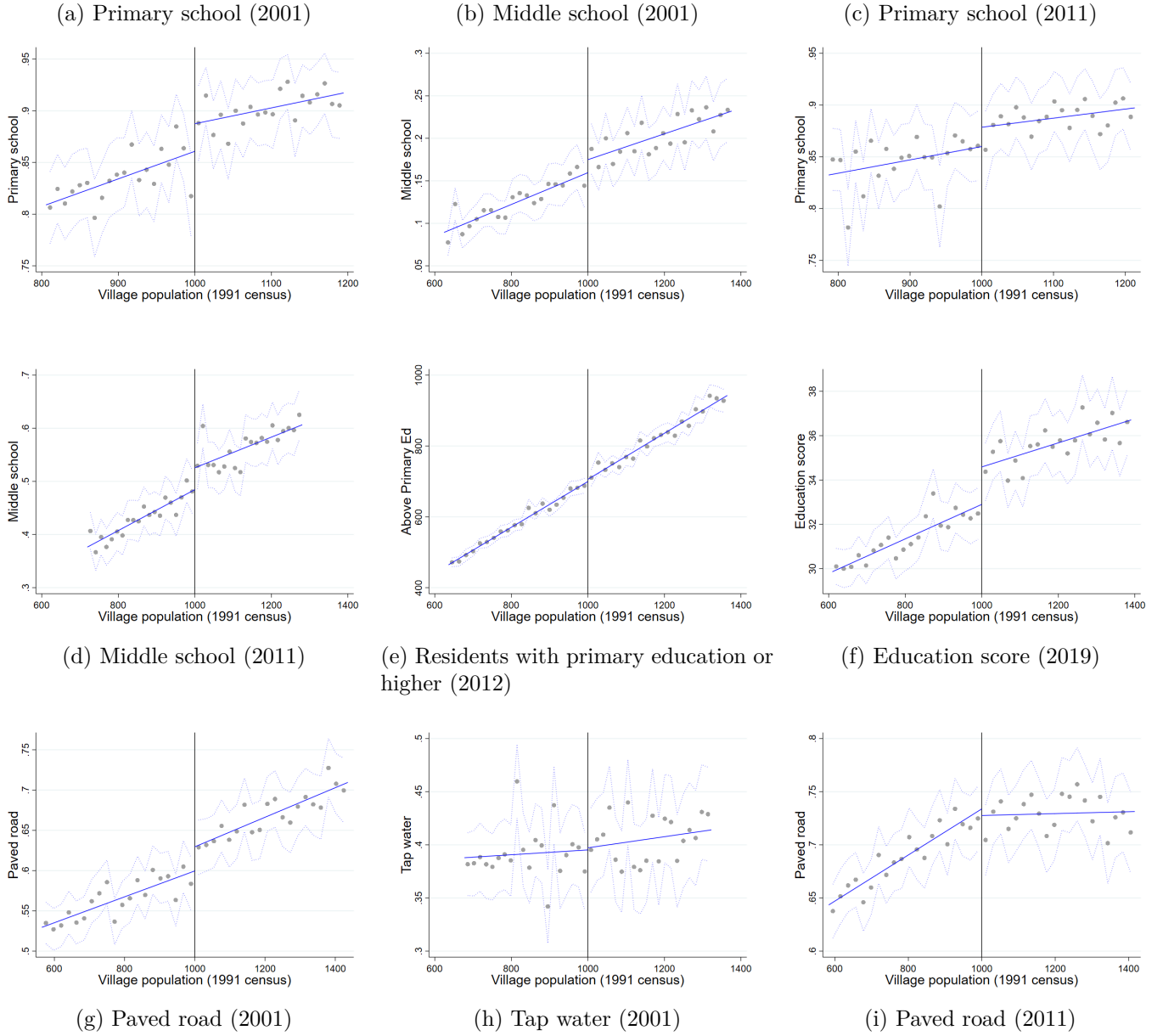
Notes: These figures plot the relationship between various characteristics of the village in the 1991 census and the village's population in the 1991 census. Points to the right of 1000 are above the treatment threshold, while points to the left are below the threshold. Each point represents approximately 480 villages. See Table 1 for the corresponding statistical tests.

Figure A5: Balance on baseline village characteristics (2011 discontinuity)



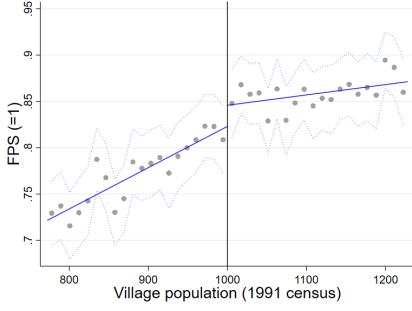
Notes: These figures plot the relationship between various characteristics of the village in the 2011 census and the village's population in the 2011 census. Points to the right of 1000 are above the treatment threshold, while points to the left are below the threshold. Each point represents approximately 480 villages. See Table 1 for the corresponding statistical tests.

Figure A6: Effects on village-level amenities (1991 discontinuity)

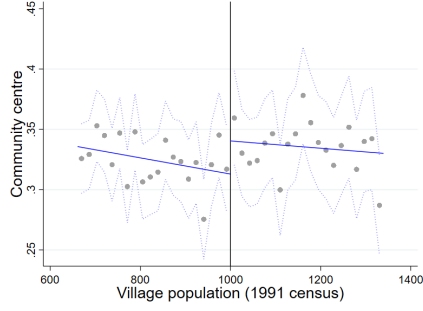


Notes: These figures plot the relationship between various characteristics of the village in the 2001 and 2011 censuses and the village's population in the 1991 census. Points to the right of 1000 are above the treatment threshold, while points to the left are below the threshold. Each point represents approximately 480 villages, and the bandwidth is based on the optimal bandwidth selection from Cattaneo et al. (2020). Panels (a) and (b) plot the relationship between the existence of primary and middle schools (the outcome is a dummy variable taking a value 0 when there is no school, and a value 1 when there is a school) in 2001 against the village population in 1991, while panels (c) and (d) plot the same outcomes in 2011. Panel (e) uses information from the Socio-economic and Caste Census, which measures the number of individuals who have a primary education or higher in 2012, while panel (f) uses the Mission Antyodaya data to generate an education measure in 2019. Panels (g) and (i) present the results with having a paved road as the main outcome for the years 2001 and 2011. Finally, panel (h) examines the presence of tap water in the village in 2001.

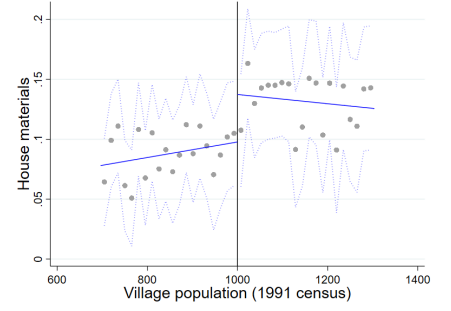
Figure A7: Effects on village-level amenities (1991 discontinuity, continued)



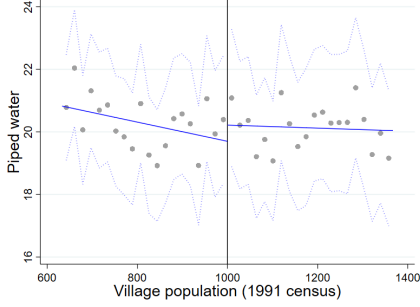
(a) Fair price shop (2011)



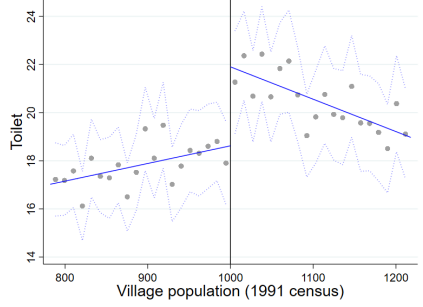
(b) Community centre (2011)



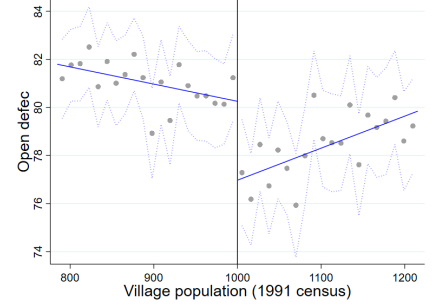
(c) House quality index (2011)



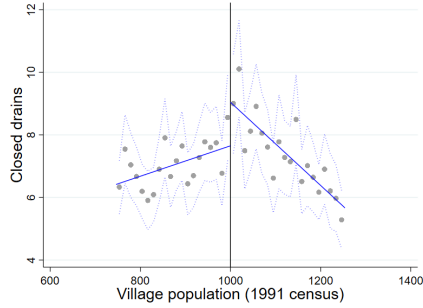
(d) In-home piped water (2011)



(e) Households with toilets (2011)



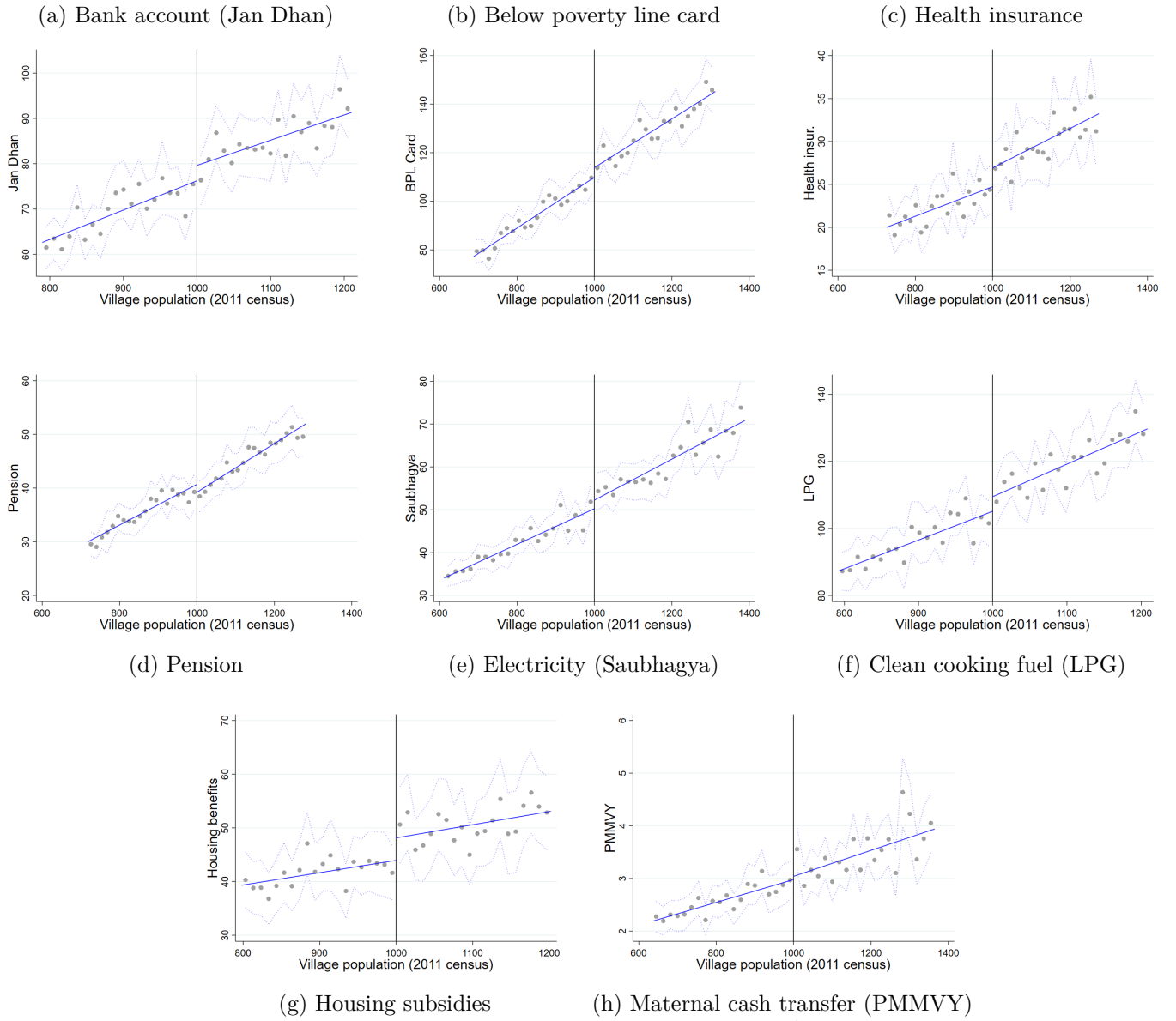
(f) Households who openly defecate



(g) Households with closed drains (2011)

Notes: These figures plot the relationship between various characteristics of the village in the 2001 and 2011 censuses and the village's population in the 1991 census. Points to the right of 1000 are above the treatment threshold, while points to the left are below the threshold. Each point represents approximately 480 villages, and the bandwidth is based on the optimal bandwidth selection from Cattaneo et al. (2020). Panel (a) plots the relationship between whether there is a fair price shop in the village against the village population in 1991. Panel (b) plots the relationship for whether there is a community centre in the village. Panel (c) presents the results for an index of house quality, based on the materials from which the wall, roof, and floor of the village was made. Panel (d) presents the figure for the outcome of piped water, measured by the share of households that have in-house piped water. Panel (e) follows this with a figure presenting the share of households that have a toilet within the village, while panel (f) presents the results for the complementary outcome of the share of households that openly defecate, and panel (g) shows the share of households with closed drains.

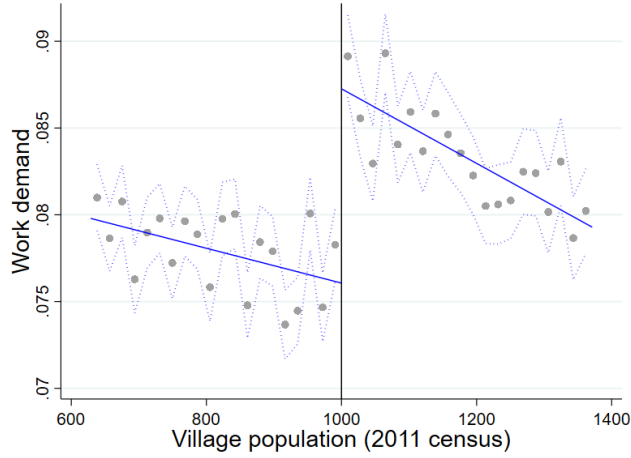
Figure A8: Effects on individual-level programs (2011 discontinuity)



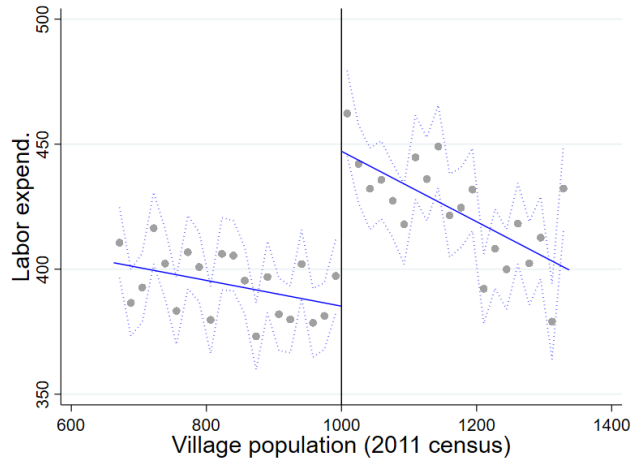
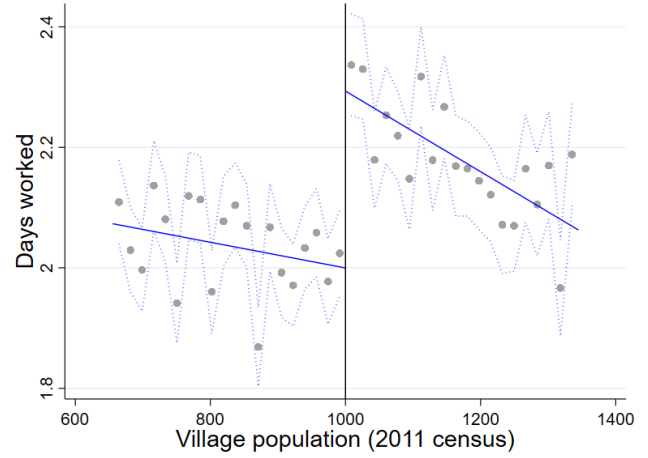
Notes: These figures plot the relationship between various service delivery outcomes in the Mission Antyodaya 2019 data and the village's population in the 2011 census. Points to the right of 1000 are above the treatment threshold, while points to the left are below the threshold. Each point represents approximately 480 villages, and the bandwidth is based on the optimal bandwidth selection from Cattaneo et al. (2020).

Figure A9: Effects on workfare performance (2011 discontinuity)

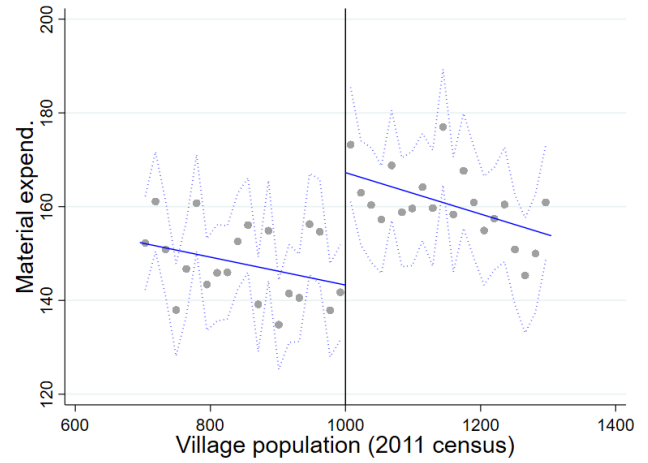
(a) Persons demanding NREGS work



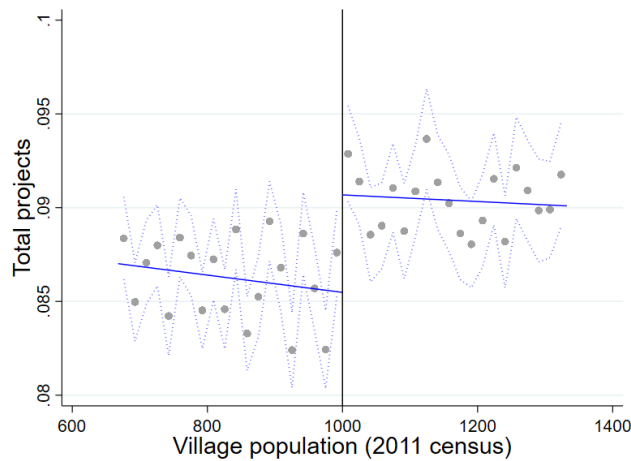
(b) Person-days worked



(c) Labor expenditure



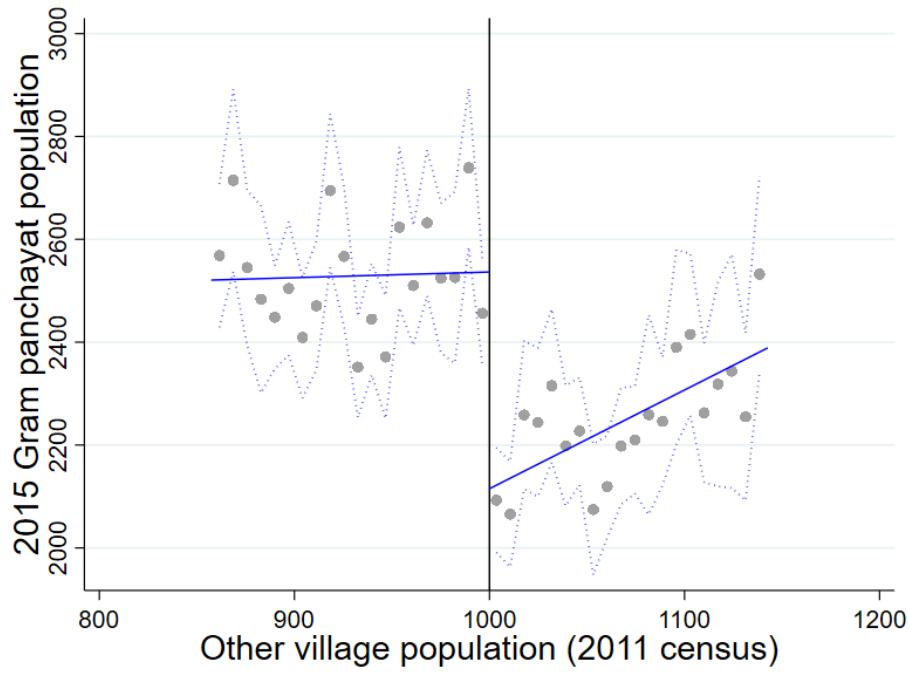
(d) Expenditure on materials



(e) Number of NREGS projects

Notes: These figures plot the relationship between outcomes related to the National Rural Employment Guarantee Scheme and the village's population in the 2011 census. Points to the right of 1000 are above the treatment threshold, while points to the left are below the threshold.

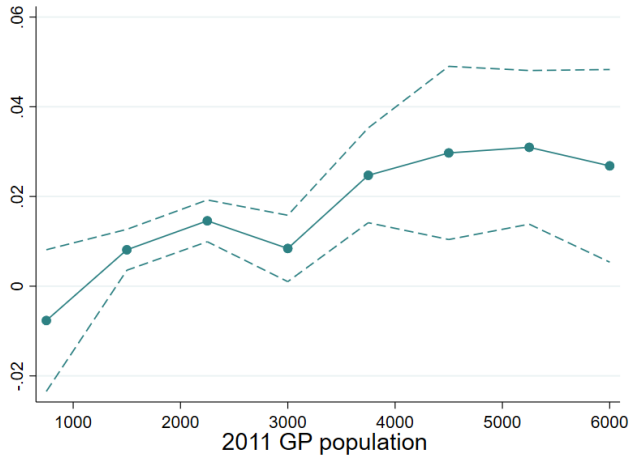
Figure A10: First stage (alternative RD)



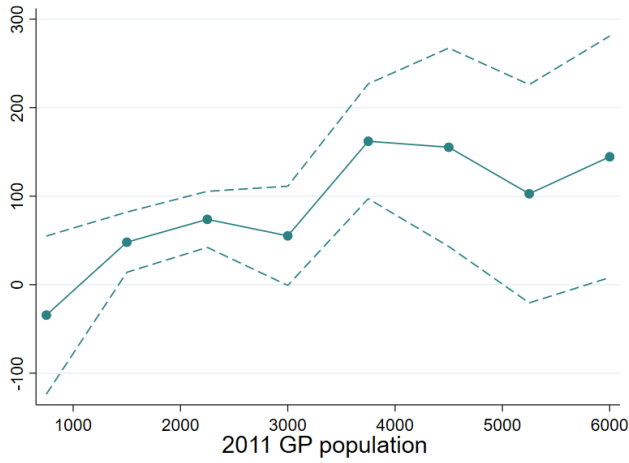
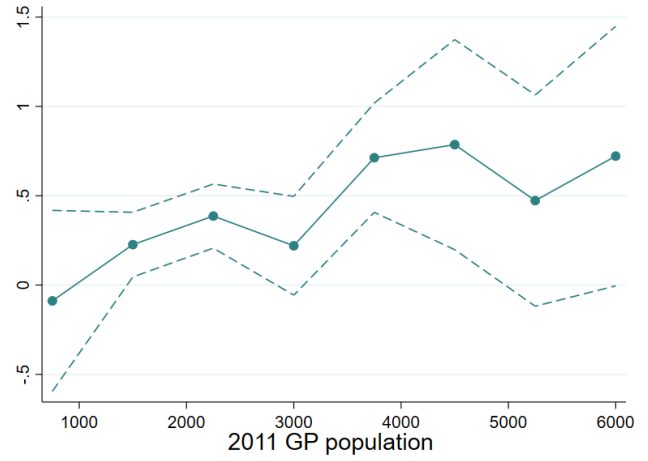
Notes: This figure uses an alternative regression discontinuity approach to measure the effect of polity size. For each gram panchayat in existence prior to the 2015 delimitation, we determine the village whose 2011 census population was closest to 1000; these are villages who are on the margin of being delimited into smaller GPs in the 2015 delimitation. We drop those villages and then use their 2011 census population as the running variable for the other villages in their pre-2015 GP. If the first village has a population to the right of the discontinuity, this will influence the post-2015 GP population for those other villages because the other villages will almost always stay together in a now smaller GP. This identifies a different local average treatment effect of the effect of post-2015 GP size as both the treatment and population of compliers are different. The figure plots the first stage relationship between the population of the omitted village and the 2015 gram panchayat population for the village of interest. Since the gram panchayat is more likely to be split if the omitted village has more than 1000 people, we observe that those panchayats with a village above the discontinuity have lower gram panchayat populations after delimitation. The sample includes only villages in GPs with greater than 2000 people at the time of the 2015 delimitation, as GPs with fewer than that were ineligible to be split.

Figure A11: Heterogeneity in effects on workfare performance by previous GP population

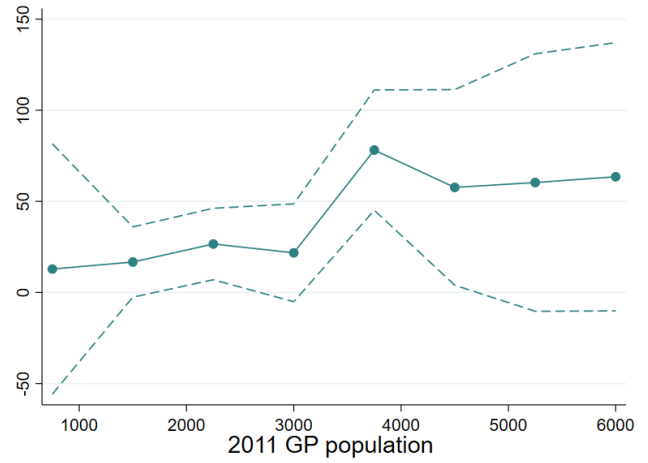
(a) Persons demanding NREGS work



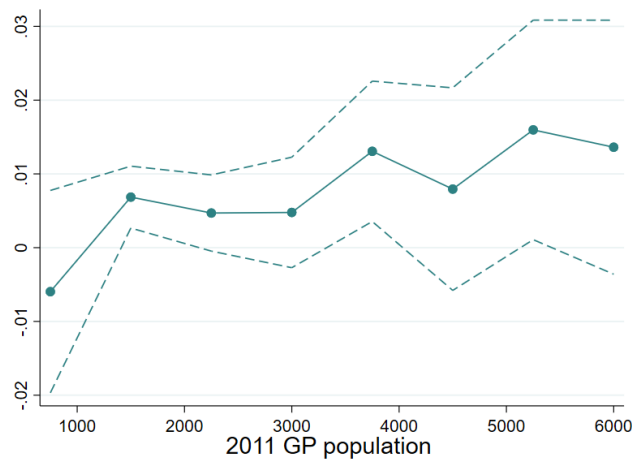
(b) Person-days worked



(c) Labor expenditure



(d) Expenditure on materials



(e) Number of NREGS projects

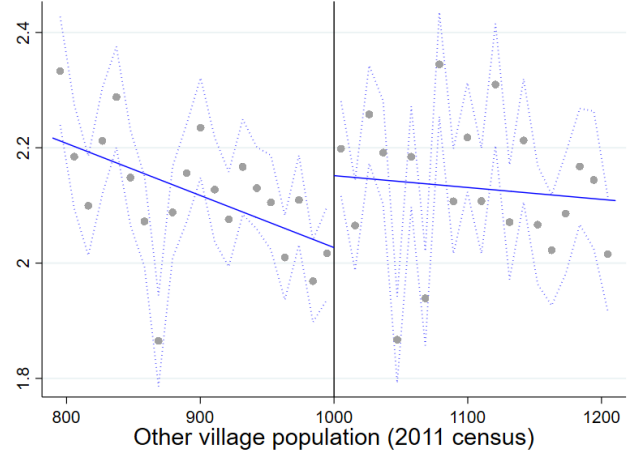
Notes: These figures examine heterogeneous treatment effects on outcomes from the National Rural Employment Guarantee Scheme. For these figures, we divide all of the villages in our sample based on the population of their GP prior to the 2015 delimitation. We run the main RD specification within each of these bins to look at effect of being above the treatment population threshold for villages within the bin. We then plot the coefficients and confidence intervals from each of these regressions in the figure. Since a village will stay with its pre-2015 GP if its population is below the threshold, these estimates can be interpreted as telling us about the treatment effect of moving from that population range into a GP with a population of around 1000 (the village itself).

Figure A12: Effects on workfare performance (alternative RD)

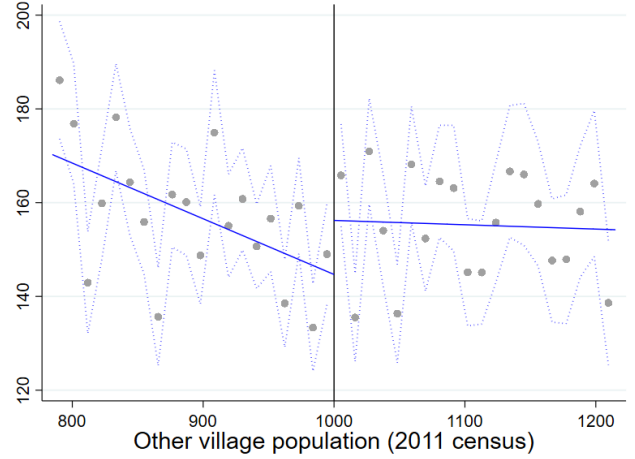
(a) Persons demanding NREGS work



(b) Person-days worked



(c) Labor expenditure



(d) Expenditure on materials

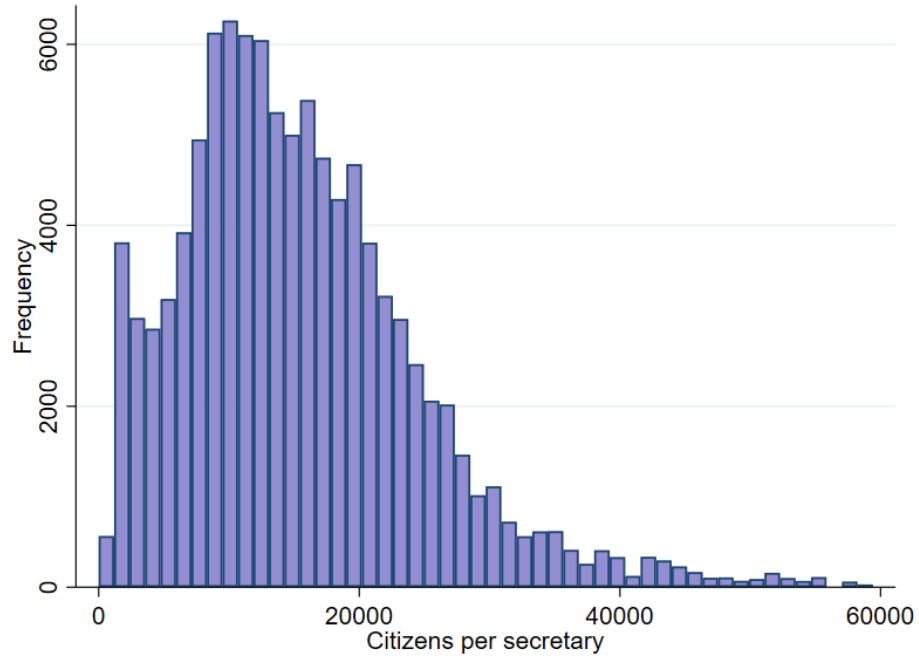


(e) Number of NREGS projects

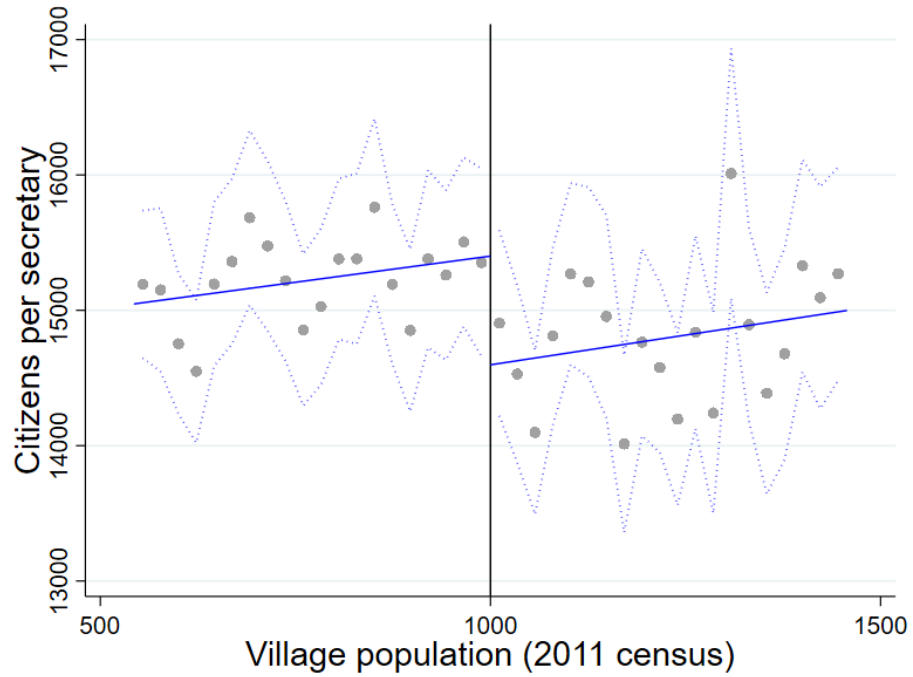
Notes: This figure uses an alternative regression discontinuity approach to measure the effect of polity size. For each gram panchayat in existence prior to the 2015 delimitation, we determine the village whose 2011 census population was closest to 1000; these are villages who are on the margin of being delimited into smaller GPs in the 2015 delimitation. We drop those villages and then use their 2011 census population as the running variable for the other villages in their pre-2015 GP. If the first village has a population to the right of the discontinuity, this will influence the post-2015 GP population for those other villages because the other villages will almost always stay together in a now smaller GP. This identifies a different local average treatment effect of the effect of post-2015 GP size as both the treatment and population of compliers are different.

Figure A13: Panchayat secretary allocation

(a) Citizens per panchayat secretary



(b) RD plot on ratio of citizens to panchayat secretary



Notes: These figures show the allocation of panchayat secretaries in the GP. Panel (a) shows the distribution of the number of citizens represented by a single panchayat secretary for each GP in Uttar Pradesh. Panel (b) reproduces the main RDD in the paper, but examines the outcome of citizens per panchayat secretary. The running variable is the 2011 census population of the village.

B Appendix Tables

Table A1: The Eleventh Schedule and All-India GP expenditure overview

<i>Panel A: Breakdown of all subjects in the 11th Schedule</i>		
• Drinking water • <i>Fuel and fodder</i> • <i>Non-conventional energy sources</i> • Poverty alleviation program • Libraries • Public Distribution System • Maintenance of community assets • <i>Cultural activities</i> • <i>Market and fairs</i> • <i>Fisheries</i> • Adult and non-formal education • Welfare of the weaker sections, and in particular, of the Scheduled Castes and the Scheduled Tribes	• <i>Small scale industries, including food processing industries</i> • <i>Land improvement, implementation of land reforms, land consolidation and soil conservation</i> • Social welfare, including welfare of the handicapped • Roads, culverts, bridges, ferries, waterways and other means of communication • <i>Rural electrification, including distribution of electricity</i> • Education, including primary and secondary schools • <i>Technical training and vocational education</i>	• Women and child development • <i>Minor irrigation, water management and watershed development</i> • Animal husbandry, dairying and poultry • Family welfare • <i>Social forestry and farm forestry</i> • <i>Minor forest produce</i> • <i>Khadi, village and cottage industries</i> • Health and sanitation, including hospitals, primary health centers and dispensaries • <i>Agriculture, including agricultural extension</i> • <i>Rural housing</i>
<i>Panel B: Breakdown of average GP expenditure during the 13th and 14th Finance Commissions</i>		
Expenditure Category	13 th FC (2013-2015)	14 th FC (2015-2018)
1 Building/Community Assets	17.9	24.26
2 Parks, Playgrounds, Burial And Cremation Grounds	4.9	4.14
3 Construction Local Body Roads and Footpaths	27	18.43
4 Maintenance of Local Body Roads and Footpaths	2.3	1.99
5 Water Supply	12.4	11.7
6 Sanitation	10.1	10.74
7 Street Lighting	7.7	10.39
8 Productive Sector	3.1	0.71
9 Welfare	0.9	1.64
10 Salaries, Wages & Employee Pensions	1.9	3.02
11 Office Expenses	2.5	2.6
12 Scheme related expenditure - Central	0.6	0.95
13 Scheme related expenditure - State	0.7	5.29
14 Others	8	4.13

Panel A of this table lists the sectors that the Central government highlighted as potential areas for devolution to the gram panchayat. State governments selected areas from within this list to delegate to GP responsibility. The items that were allocated to GPs in Uttar Pradesh in practice are indicated in bold (Ministry of the Panchayati Raj, 2016; Indian Institute of Public Administration, 2019). Those were legally devolved, but not actually taken up by the GP in practice are indicated in italics (Ibid). Panel B presents the average expenditure shares (as a %) for GPs for key areas of spending. This information is taken from Centre for Policy Research (2019), a report on fund flows in GPs, which uses a national-level sample of GPs to analyse their expenditure patterns.

Table A2: Match rates

	Match rate		Match probability	
	Full sample	In sample	RD estimate	p-value
<i>Panel A: Match to 1991 census data</i>				
DDWS GP directory	96.28%	97.35%	.00507 (.00384)	.187
Socioecon. and Caste Census	95.11%	96.65%	.00663 (.0053)	.21
Mission Antyodaya	98.27%	98.55%	.00254 (.00312)	.415
GP election (2010)	87.18%	89.85%	.01064 (.00823)	.196
<i>Panel B: Match to 2011 census data</i>				
Local government directory	99.13%	99.43%	.00049 (.0017)	.774
NREGS	86.35%	87.55%	.00887 (.00837)	.289
eGramSwaraj	85.54%	93.66%	-.01364 (.00943)	.148
Mission Antyodaya	95.89%	98.74%	.00011 (.00286)	.968
GP elections (2015/2020)	93.91%	96.57%	-.00216 (.00525)	.681
Gram sabha meetings	93.53%	96.62%	-.00264 (.00491)	.591

Columns (1-2) show match rate between the census data and listed data set for all villages and the villages with populations between 800 to 1200. Columns (3-4) test whether the match rate differs around the discontinuity. Column (3) reports regression discontinuity estimates from the main estimating equation of the effect a village of being above the treatment population threshold on whether the village successfully matched to the listed data set: the first number is the coefficient from the RD equation, while the standard errors are reported below that in parentheses. Column (4) reports the p-values from that regression. The running variable is the population of the village in the relevant census round.

Table A3: Effects on additional village amenities

Panel A: Additional educational outcomes								
	Index	2001		2011				
		Secon school	Sen Secon school	Secon school	Sen Secon school			
RD_Estimate	0.032 (0.021) [0.122]	0.004 (0.005) [0.381]	0.001 (0.003) [0.638]	0.012 (0.010) [0.236]	0.016 (0.008) [0.055]			
Dep var mean	-0.072	0.023	0.013	0.093	0.055			
Bandwidth	246	370	448	226	212			
Effective Obs	19514	30198	36904	17871	16822			
Panel B: Incompletely devolved functions (2001)								
	Index	Electrified	Irrigated frac	Maternity- child centre	Community health worker			
RD_Estimate	-0.005 (0.015) [0.767]	0.005 (0.014) [0.724]	0.005 (0.009) [0.580]	-0.008 (0.008) [0.314]	-0.022 (0.011) [0.044]			
Dep var mean	-0.026	0.284	0.794	0.069	0.163			
Bandwidth	315	278	305	303	385			
Effective Obs	25324	22263	20136	23575	30554			
Panel C: Incompletely devolved functions (2011)								
	Index	Electricity	Irrigated frac	Anganwadi centre	Nutrition centre	Weekly market	Ag mar- keting society	ASHA
RD_Estimate	0.023 (0.013) [0.090]	0.818 (0.562) [0.146]	0.004 (0.005) [0.476]	-0.001 (0.011) [0.917]	0.011 (0.014) [0.445]	0.018 (0.011) [0.082]	0.024 (0.011) [0.029]	-0.008 (0.009) [0.362]
Dep var mean	0.033	21.584	0.875	0.887	0.476	0.166	0.249	0.918
Bandwidth	270	360	427	196	325	321	389	236
Effective Obs	21597	29284	34809	15496	26291	25860	31750	18774

This table reports regression discontinuity estimates from the main estimating equation of the effect of being above the treatment population threshold on different village amenities. Panel A presents additional outcomes related to education: whether the village had a secondary or senior secondary school present in 2001 or 2011. The outcomes in panel B are outcomes in 2001 related to functions that are legally devolved to the gram panchayat, but are not actually devolved in practice (Indian Institute of Public Administration, 2019): whether the village has electricity, the fraction of village land that is irrigated, whether the village has a community health worker, and whether the village has a maternity-child welfare centre. in 2011 (all measured in the respective Census round). The outcomes in panel C are outcomes in 2011 related to functions that are legally devolved to the gram panchayat, but are not actually devolved in practice: the fraction of households with electricity, fraction of village land that is irrigated, and whether the village has an anganwadi centre, dedicated ICDS nutrition centre, regular market (mandi), agricultural marketing society, or Accredited Social Health Activist (ASHA). All variables are measured in either the 2001 or 2011 census of India. We create an index of the variables in each panel following Kling et al. (2001). The running variable is the population of the village in the 1991 census. Each specification uses a linear polynomial, triangular kernel, and MSE-optimal bandwidth estimated following Calonico et al. (2017). Standard errors are clustered at the gram panchayat level and reported below the point estimates. p-values are reported within brackets.

Table A4: Effects on housing materials

Panel A: Wall materials						
	Aggregate		Individual materials			
	Low	High	Brick	Mud	Oth. high quality	Oth. low quality
RD_Estimate	-1.146 (0.515) [0.026]	1.079 (0.565) [0.056]	1.371 (0.563) [0.015]	-0.788 (0.587) [0.180]	-0.463 (0.260) [0.075]	-0.490 (0.286) [0.087]
Dep var mean	31.285	66.405	63.601	24.795	2.827	6.262
Bandwidth	404	345	377	299	210	247
Effective Obs	33162	27982	30783	23932	16577	19602
Panel B: Roof materials						
	Aggregate		Individual materials			
	Low	High	Brick	Grass	Oth. high quality	Oth. low quality
RD_Estimate	-1.575 (0.545) [0.004]	1.710 (0.551) [0.002]	2.072 (0.832) [0.013]	-1.180 (0.474) [0.013]	-0.256 (0.905) [0.777]	-0.184 (0.103) [0.075]
Dep var mean	29.588	69.910	32.147	26.387	37.738	1.428
Bandwidth	327	321	274	442	308	259
Effective Obs	26363	25855	21926	36444	24802	20618
Panel C: Floor materials						
	Aggregate		Individual materials			
	Low	High	Mud	Cement	Oth. high quality	Oth. low quality
RD_Estimate	-0.235 (0.427) [0.582]	0.222 (0.426) [0.602]	-0.206 (0.429) [0.632]	0.383 (0.360) [0.288]	-0.139 (0.227) [0.540]	-0.033 (0.023) [0.145]
Dep var mean	86.475	13.459	86.211	9.257	4.230	0.263
Bandwidth	244	245	245	226	242	348
Effective Obs	19426	19510	19426	17945	19246	28150
Panel D: Condition of the house						
	Condition			Avg Rooms		
	Bad	Livable	Good			
RD_Estimate	-0.360 (0.196) [0.067]	1.317 (0.675) [0.051]	-0.816 (0.701) [0.244]	-0.007 (0.014) [0.626]		
Dep var mean	7.423	50.750	36.980	2.250		
Bandwidth	405	277	282	338		
Effective Obs	33162	22086	22490	27322		

This table reports regression discontinuity estimates from the main estimating equation of the effect of being above the treatment population threshold on different public-service related outcomes. Panels A through C present outcomes related to the fraction of village residents living in houses made out of particular materials. The first two columns aggregate these into either high or low quality materials, based on whether the variable is positively or negatively related to the condition of the housing stock in the village. The remaining columns disaggregate the effects for two

Table A5: Additional amenities results: related government programs

	(1) IAY (=1)	(2) TSC (=1)
RD_Estimate	0.039 (0.017) [0.021]	0.015 (0.009) [0.077]
Dep var mean	0.440	0.066
Bandwidth	219	208
Effective Obs	17708	16713

This table reports RD estimates of the effect of delimitation on delivery of two key government programs: IAY and TSC. The running variable is the population of the village in 1991. Standard errors are clustered at the gram panchayat level. The outcome variables are measured as dummy variables. The first is a binary variable for whether a village had any construction under the Indira Awaas Yojana Program in the 2014-2015 fiscal year, as measured by administrative data from the Ministry of Housing and Urban Affairs on IAY construction. Unfortunately, this is the only year of data available prior to the 2015 delimitation exercise. The second comes from a variable in the 2011 census of India that records whether the village has participated in the Total Sanitation Campaign; however, it is unclear from survey documentation exactly how this was defined.

Table A6: Effects on income

	HHs in income range (Rs)			SHRUG
	(1) <5,000	(2) 5-10,000	(3) > 10,000	(4) Poverty rate
RD_Estimate	-0.005 (0.007) [0.446]	0.003 (0.005) [0.560]	0.002 (0.003) [0.356]	-0.004 (0.005) [0.456]
Dep Mean	0.728	0.194	0.078	0.360
Bandwidth	267	267	300	237
Effective Obs	20669	20741	23464	17914

This table reports RD estimates of the effect of delimitation on household income using outcome data from the 2011-2012 Socioeconomic and Caste Census. Columns (1)-(3) use village-level SECC data that measures the fraction of households in different monthly income categories: less than Rs. 5000 (100USD at the time of SECC), Rs. 5,000-10,000 (100-200USD), and more than Rs 10,000 (200USD). Column (4) uses data from the SECC made available in the Socioeconomic High-resolution Rural-Urban Geographic dataset on India (SHRUG) (Asher et al., 2021). Using household-level SECC microdata, they estimate the fraction of households in each village below the poverty line. Standard errors are clustered at the gram panchayat level.

Table A7: Robustness check on post-2015 outcomes (dropping villages in split GPs below threshold)

<i>Panel A: Individual-level programs</i>									
	Programs Index	BPL Card	Health insur.	Pension	Saubhagya	LPG	Housing benefits	PMMVY	Jan Dhan
RD_Estimate	0.083 (0.038) [0.029]	7.129 (2.442) [0.004]	1.692 (1.201) [0.159]	-2.167 (1.191) [0.069]	1.890 (1.510) [0.211]	5.220 (2.588) [0.044]	4.513 (2.191) [0.039]	0.018 (0.152) [0.903]	4.611 (2.279) [0.043]
Dep var mean	0.107	87.062	22.498	34.742	41.444	94.673	40.545	2.546	69.961
Bandwidth	152	369	263	290	377	226	241	369	210
Effective Obs	9924	26253	17986	20057	26882	14987	16262	26186	13724
<i>Panel B: Workfare program implementation</i>									
	NREGS Index	Work demand	Days worked	Labor expend.	Material expend.	Total projects			
RD_Estimate	0.217 (0.027) [0.000]	0.016 (0.001) [0.000]	0.440 (0.052) [0.000]	91.566 (9.893) [0.000]	36.655 (5.808) [0.000]	0.004 (0.002) [0.004]			
Dep var mean	0.066	0.077	1.995	385.158	142.603	0.087			
Bandwidth	370	361	350	349	331	366			
Effective Obs	116525	113322	109454	109316	102600	114980			

This table reports regression discontinuity estimates from the main estimating equation of the effect of being above the treatment population threshold on different welfare program and NREGS outcomes. See Table 3 for details. We restrict the sample to remove villages that are just below the threshold but paired with a village that is above the threshold. The running variable is the population of the village in the 2011 census. Each specification uses a linear polynomial, triangular kernel, and MSE-optimal bandwidth estimated following Calonico et al. (2017). Standard errors are clustered at the gram panchayat level and reported below the point estimates. p-values are reported within brackets.

Table A8: Correlation between poverty and BPL card ownership

	BPL frac
Above X Poverty Rate	0.830 (0.355) [0.019]
Above (=1)	-0.198 (0.141) [0.160]
Dep var mean	0.516
Bandwidth	300
Observations	22650

This table reports estimates of how targeting of below poverty line ration cards responds to being allocated into a smaller polity. It reports estimates of β_4 and β_7 from the regression specification:

$$BPL = \beta_0 + \beta_1 \text{poverty} + \beta_2 \text{poverty} \times \text{population} + \beta_3 \text{poverty} \times \text{population} \times \text{above} \\ + \beta_4 \text{poverty} \times \text{above} + \beta_5 \text{population} + \beta_6 \text{population} \times \text{above} + \beta_7 \text{above} + u \quad (3)$$

where *poverty* is the poverty rate of the village as measured in the SECC, *population* is the village's population in the 2011 census, and *above* is a dummy variable equal to one if that population is above 1000. β_1 measures the relationship between poverty and BPL card ownership, the second term allows this relationship to vary with the village population, and the third term allows this relationship to vary above and below the discontinuity. The coefficient of interest is β_4 , which measures whether the relationship between poverty and BPL card ownership becomes discontinuously stronger at the threshold. The remaining terms saturate the model and correspond to a standard RD equation in population; β_7 can be interpreted as the change in BPL cards at the discontinuity if the poverty rate of the village is equal to zero. Standard errors are clustered at the gram panchayat level.

Table A9: Effect on delivery of workfare (by identity group)

	Job cards		HHs worked		Person-days		
	(1) SC	(2) Non-SC	(3) SC	(4) Non-SC	(5) SC	(6) Non-SC	(7) Female
RD_Estimate	0.017 (0.008) [0.029]	0.010 (0.002) [0.000]	0.010 (0.002) [0.000]	0.008 (0.001) [0.000]	0.390 (0.115) [0.001]	0.347 (0.050) [0.000]	0.101 (0.023) [0.000]
Equality test (p-value)		.397		.581		.727	
Dep var mean	0.208	0.129	0.077	0.047	3.107	1.876	0.704
Bandwidth	292	384	388	376	357	371	335
Effective Obs	101301	135735	134390	132851	123500	131008	117879

This table reports RD estimates of the effect of delimitation on delivery of workfare programs through NREGS. It looks at three NREGS outcomes for which the data breaks down receipt by caste identity of the recipient in columns (1)-(6). In column 7, it examines the only variable disaggregated by gender, which is person-days worked. The running variable is the population of the village in 2011 census. Standard errors are clustered at the gram panchayat level.

Table A10: Effects on village amenities (bandwidth of 200)

<i>Panel A: Educational outcomes</i>							
	Index	2001		2011			2019
	Education	Primary school	Middle school	Primary school	Middle school	Above Primary Ed	Education score
RD_Estimate	0.091 (0.022) [0.000]	0.027 (0.012) [0.022]	0.019 (0.013) [0.161]	0.018 (0.012) [0.124]	0.059 (0.018) [0.001]	20.320 (10.983) [0.064]	2.229 (0.615) [0.000]
Dep var mean	-0.108	0.833	0.141	0.846	0.446	632.422	32.175
Bandwidth	200	200	200	200	200	200	200
Effective Obs	15761	15760	15760	15761	15761	15231	15603
<i>Panel B: Village-level infrastructure</i>							
	Index	2001		2011			
	Amenities (village)	Paved road	Tap water	Paved road	FPS (=1)	Community centre	
RD_Estimate	0.060 (0.025) [0.017]	0.033 (0.017) [0.057]	0.025 (0.017) [0.142]	-0.010 (0.016) [0.514]	0.029 (0.013) [0.024]	0.019 (0.017) [0.257]	
Dep var mean	0.287	0.584	0.396	0.711	0.775	0.318	
Bandwidth	200	200	200	200	200	200	
Effective Obs	15761	15760	15761	15706	15761	15761	
<i>Panel C: Household-level infrastructure</i>							
	Index	House		Sanitation			
	Amenities (HH)	House materials	Piped water	Toilet	Open defec	Closed drains	
RD_Estimate	0.093 (0.026) [0.000]	0.036 (0.020) [0.067]	0.045 (1.000) [0.964]	3.266 (0.680) [0.000]	-3.388 (0.706) [0.000]	1.175 (0.540) [0.030]	
Dep var mean	-0.040	0.092	20.035	17.848	81.031	7.042	
Bandwidth	200	200	200	200	200	200	
Effective Obs	15759	15759	15759	15759	15759	15759	

See the table note from Table 2 for further details. The running variable is the population of the village in the 1991 census. Each specification uses a linear polynomial, uniform kernel, and MSE-optimal bandwidth estimated following Calonico et al. (2017). Standard errors are clustered at the gram panchayat level and reported below the point estimates. p-values are reported within brackets.

Table A11: Effect on delivery of services (bandwidth of 200)

<i>Panel A: Individual-level programs</i>									
	Programs Index	Below Poverty Line Card	Health insur- ance	Pension	Saubhagya electr.	LPG	Housing benefits	PMMVY	Jan Dhan
RD_Estimate	0.080 (0.030) [0.007]	5.636 (3.011) [0.061]	2.451 (1.241) [0.048]	-1.319 (1.287) [0.305]	3.818 (1.911) [0.046]	5.831 (2.505) [0.020]	5.597 (2.202) [0.011]	0.247 (0.179) [0.169]	5.253 (2.107) [0.013]
Dep var mean	0.104	99.290	22.941	36.996	46.126	96.394	41.595	2.767	69.765
Bandwidth	200	200	200	200	200	200	200	200	200
Effective Obs	16005	16005	16005	16005	16005	16005	16005	16005	16005
<i>Panel B: Workfare program implementation</i>									
	NREGS Index	Work demand	Days worked	Labor expend.	Material expend.	Total projects			
RD_Estimate	0.157 (0.034) [0.000]	0.011 (0.002) [0.000]	0.301 (0.065) [0.000]	60.351 (12.339) [0.000]	21.749 (6.854) [0.002]	0.006 (0.002) [0.004]			
Dep var mean	0.071	0.077	2.023	390.602	146.261	0.086			
Bandwidth	200	200	200	200	200	200			
Effective Obs	71033	71033	71033	70923	70923	70923			

This table reports regression discontinuity estimates from the main estimating equation of the effect of being above the treatment population threshold on different welfare program and NREGS outcomes. Panel A presents outcomes related to welfare programs as measured in the Mission Antyodaya data in 2019. See the table notes for 3 for details of the outcomes. The running variable is the population of the village in the 2011 census. Each specification uses a linear polynomial, triangular kernel, and fixed bandwidth of 200. Standard errors are clustered at the gram panchayat level and reported below the point estimates. p-values are reported within brackets.

Table A12: Effects on GP election outcomes in unreserved GPs

<i>Panel A: Political participation</i>						
	Gram Sabha			Election		
	Total partic.	SC partic.	Female partic.	Voter turnout	Candidates per 1000	Eff. can- didates per 1000
RD_Estimate	0.022 (0.005) [0.002]	0.022 (0.008) [0.002]	0.003 (0.001) [0.002]	1.906 (0.301) [0.002]	1.057 (0.101) [0.002]	0.859 (0.048) [0.002]
Dep var mean	0.080	0.134	0.012	75.736	4.964	2.870
Bandwidth	262	374	292	307	287	371
Effective Obs	7013	9172	7827	17877	16792	21586
<i>Panel B: Candidate characteristics</i>						
	Avg Educ	Avg criminal record	Avg assets (asinh)	General caste perc	Female perc	Avg age
RD_Estimate	0.011 (0.037) [0.958]	0.001 (0.016) [0.978]	-0.093 (0.111) [0.691]	-0.005 (0.014) [0.958]	-0.010 (0.007) [0.255]	-0.192 (0.142) [0.255]
Dep var mean	2.759	0.128	12.874	0.481	0.194	41.549
Bandwidth	345	277	270	266	315	446
Effective Obs	19983	7952	7768	15656	18371	25852
<i>Panel C: Winner characteristics</i>						
	Education	Criminal Record	Total Assets (Asinh)	General Caste	Female	Age
RD_Estimate	0.009 (0.055) [0.976]	-0.008 (0.018) [0.902]	-0.293 (0.114) [0.008]	-0.020 (0.013) [0.166]	-0.022 (0.011) [0.064]	0.038 (0.290) [0.976]
Dep var mean	3.069	0.129	13.203	0.493	0.184	41.612
Bandwidth	304	262	435	351	292	286
Effective Obs	25602	7529	12250	29908	24951	24380

This table reports regression discontinuity estimates from the main estimating equation of the effect of being above the treatment population threshold on electoral outcomes in that village's GP. The outcomes in Panel A are related to the overall political participation in the GP, panel B relates to the characteristics of candidates contesting for the position of pradhan, and panel C focused on the characteristics of the candidate elected as pradhan. The data is from meetings notes of the Gram Panchayat Development Programme (GPD) from the Ministry of the Panchayati Raj and the Uttar Pradesh State Election Commission for the 2015 and 2020 elections, including aggregate voting data and candidate affidavits. The first three outcomes in panel A are attendance rates at GPD Gram Sabha meetings for the whole village, female residents, and scheduled caste residents. The second three are voter turnout, number of candidates per 1000 residents of the GP, and effective candidates (receiving more than 5 percent of votes) per 1000 residents in the 2015 and 2020 GP elections. The outcomes in Panel B are the average education level, criminal record, age, and asset holdings (inverse hyperbolic sine) of candidates, as well as fraction of candidates who are general caste and female. The outcomes in Panel C are the education level, criminal record, age, asset holdings (inverse hyperbolic sine), caste identity and gender of the person elected pradhan. The running variable is the population of the village in the 2011 census. Each specification uses a linear polynomial, triangular kernel, and MSE-optimal bandwidth estimated following Calonico et al. (2017). Standard errors are clustered at the gram panchayat level and reported below the point estimates. p-values are reported within brackets.

Table A13: Political competition in GP elections

					Margin of victory		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Candidates	Effective candi- dates	Herfindahl index (votes)	Margin of victory	< 5 pp	5-15 pp	15+ pp
RD_Estimate	-0.931 (0.09) [0.000]	-0.229 (0.03) [0.000]	0.019 (0.00) [0.000]	0.457 (0.58) [0.432]	-0.022 (0.01) [0.027]	-0.022 (0.01) [0.011]	0.048 (0.01) [0.000]
Dep var mean	7.840	4.280	0.291	11.577	0.353	0.408	0.234
Bandwidth	329	402	305	311	299	405	330
Effective Obs	51091	62652	47607	48511	46540	62969	51203

This table reports regression discontinuity estimates from the main estimating equation of the effect of being above the treatment population threshold on electoral outcomes in that village's GP. This uses data from the 2015 and 2020 GP elections. Column (1) reports effects on the total number of candidates, while column (2) reports effects on the number of candidates who garner at least 5% of the vote. Column (3) reports a Herfindahl index of the vote shares in the election. Column (4) reports effects on the margin of victory for the winning candidate, while the remain columns use binary variables for whether the margin of victory was within particular ranges. Standard errors are clustered at the gram panchayat level.

Table A14: Effects on number of representatives

	Citizens per rep.		Effect of ward councilors	
	Secretary	Ward Councilor	Programs Index	NREGS Index
RD_Estimate	-690.636 (236.133) [0.003]	-40.246 (1.502) [0.000]	-0.018 (0.025) [0.475]	-0.032 (0.028) [0.264]
Dep var mean	15221.645	178.121	0.103	0.080
Bandwidth	463	345	423	334
Effective Obs	36309	26991	56313	201715

Columns (1) and (2) of this table reports regression discontinuity estimates from the main estimating equation of the effect of being above the treatment population threshold on the ratio of citizens to representatives (panchayat secretary and ward councilor) for the village. Columns (3) and (4) take a separate RD approach to test the effect of additional ward councilors for access to programs and workfare. The population of the GP is the running variable and there are thresholds at values of 2000 and 3000. See section B for details. Standard errors are clustered at the gram panchayat level.

Table A15: Effect of delimitation on GP budgets

<i>Panel A: Per capita budget</i>			
	Total	GP taxation	Additional schemes
RD_Estimate	3.345 (6.618) [0.613]	0.068 (0.054) [0.213]	-0.186 (0.745) [0.803]
Dep var mean	541.284	0.186	5.371
Bandwidth	153	100	297
Effective Obs	109156	70573	210441
<i>Panel B: Total Budget (in '000 rupees)</i>			
	Total	GP taxation	Additional schemes
RD_Estimate	-257.047 (12.239) [0.000]	-0.034 (0.099) [0.732]	-3.330 (1.283) [0.009]
Dep var mean	916.292	0.421	11.953
Bandwidth	282	107	306
Effective Obs	199918	76873	217948

This table reports regression discontinuity estimates from the main estimating equation of the effect a village of being above the treatment population threshold on financial outcomes in their GP. The outcomes in Panel A are measured on a per capita basis, while those in panel B are the total amount for the GP (in thousands of rupees). Column (1) is the total GP budget across all funding categories, Column (2) is the local taxation revenues for the GP. Column (3) is funding from additional government programs. We use the discontinuity based on 1991 population for 2013-2014 and on 2011 population for 2016-2020. We exclude the 2015-16 financial year since that is the year of delimitation. The running variable is the population of the village in the 1991 census for 2013-14 and 2011 census for years after that. Each specification uses a linear polynomial, triangular kernel, and MSE-optimal bandwidth estimated following Calonico et al. (2017). Standard errors are clustered at the gram panchayat level and reported below the point estimates. p-values are reported within brackets.

Table A16: Heterogenous effects based on whether a village was single hamlet or multi-hamlet

<i>Panel A: Post-1995 delimitation outcomes</i>						
	Education		Village Amenities		HH Amenities	
	Single	Multiple	Single	Multiple	Single	Multiple
RD_Estimate	0.093 (0.031) [0.003]	0.055 (0.024) [0.021]	0.058 (0.026) [0.028]	0.044 (0.027) [0.107]	0.107 (0.036) [0.003]	0.075 (0.026) [0.004]
Equality test (p-value)		.323		.718		.47
Dep var mean	-0.115	-0.140	0.245	0.267	0.064	-0.137
Bandwidth	232	294	427	288	263	269
Effective Obs	7945	13714	15954	13473	9073	12567

This table reports regression discontinuity estimates from the main estimating equation of the effect of being above the treatment population threshold on different welfare program and NREGS outcomes. It splits the sample based on whether the village is composed of multiple or a single hamlet. It examines the three main indices of outcomes following the 1995 delimitation process. See the online Appendix for details of each variable. The running variable is the population of the village in the 2011 census. Each specification uses a linear polynomial, triangular kernel, and MSE-optimal bandwidth estimated following Calonico et al. (2017). Standard errors are clustered at the gram panchayat level and reported below the point estimates. p-values are reported within brackets.

C Checking discontinuity usage by other programs

As with any regression discontinuity design, a major concern is that other variables may vary discontinuously across the cutoff, and so the estimates do not solely capture the effect of the treatment of interest. That concern is particularly salient here, where round population figures such as 1000 could plausibly be used by policymakers for other state or national-level programs. For example, Asher and Novosad (2020) study a national road-building program in which hamlets with populations of more than 500 or 1000 persons in the 2001 census were more likely to receive all-weather roads, Burlig and Preonas (2021) examine a national electrification program that targeted hamlets with populations greater than 300 people (2001 census), and Spears and Lamba (2016) evaluate a sanitation program with incentives to build toilets based on the gram panchayat population in the 2001 census.

Although we are unaware of any other programs that use the population threshold of 1000 persons in the 1991 and 2011 censuses, we implement five robustness checks as a test. The first addresses the concern of all-India programs that use the threshold. Note that the three papers listed in the previous paragraph studied national-level programs, so national programs are a particular concern. Panels (a) and (b) of Figure A14 plot the relationship between village population and GP population across all of the other states of India (similar to Figure 1). These figures demonstrate that there is not a discontinuity at a village population of 1000 persons in either the 1991 or 2011 censuses, confirming that other states did not use this cutoff for delimitation of GPs. We can thus check for the presence of national programs using this cutoff by checking to see if there is a discontinuity in the outcome variables in other states. If so, this would indicate that our results could be explained by a national program rather than GP delimitation. We run our main RD specification but now use outcome data from other states: we focus on the indices that compose our main outcome variables (education, village infrastructure, household infrastructure, workfare, and individual-level programs). Figure 4 plots the point estimates and 95% confidence intervals (orange dashed lines). Of the four outcomes, one (village infrastructure) is statistically significant at the 5% level, but the rest are not.⁴⁸ We will return to this at the end to consider whether this one statistically significant result may be due to chance since we are running many regressions.

While this points against a national-level program explaining our results, Uttar Pradesh may have used this discontinuity for state-specific programs. Our second check uses a similar approach in checking whether there is a discontinuity in the outcomes in areas of Uttar Pradesh that did not follow the delimitation rules. In both 1995 and 2015, approximately 15% of Uttar Pradesh districts did not adhere to the rules. For these districts, there is no first stage relationship between village population and GP population in both years (panels (c) and (d) of Figure A14).⁴⁹ While

⁴⁸We cannot test the workfare index for other states as we only have NREGS data for Uttar Pradesh.

⁴⁹We identify these districts by running the first-stage regression in each of them individually and selecting those with estimated t-statistics less than 1.5 or a first stage coefficient of less than 0.15

it could be the case that these districts did not follow population-based rules for other programs, the non-compliant districts are mostly different in 1995 and 2015 – with a correlation of only 0.15 – suggesting that this non-compliance is idiosyncratic rather than reflecting fixed characteristics that make them less likely to follow rules for other programs. All five of the estimates are statistically insignificant (brown dashed lines in Figure 4).

As a third check, we examine the relationship between the outcomes and the village population in the 2001 census. This population was not used for delimitation of GPs (panels (e) of Figure A14), but might have been used for other programs if Uttar Pradesh or national-level programs have a propensity to use a population cutoff of 1000 for program eligibility. None of these estimates (silver dashed lines in Figure 4) are positive and statistically significant, suggesting against this being a typical threshold for resource allocation in Uttar Pradesh.

The fourth check leverages the fact that prior to the 2015 delimitation, some villages were either (i) the lone village in their GP but still had a population around the cutoff;⁵⁰ or (ii) their GP had fewer than 2000 people in it (including their village). In the first case, since their GP could not be divided any further, the delimitation process did not affect them. In the second case, Uttar Pradesh guidelines were for GPs to have a minimum of 1000 people. Thus even if the focal village were above the threshold, it could not be split off because the other villages in its GP would be too small to form their own independent GP. This lack of a first stage relationship can be seen in panel (f) of Figure 1. However, in both cases, villages that are to the right of the cutoff would still have benefited from any state or national program using an 1000 person cutoff for eligibility, and so they form an ideal placebo sample. As seen in the red dashed lines in Figure 4, the estimates are close to zero and non-statistically significant.

The final check takes advantage of how the delimitation of one village affects the other villages that were in the same panchayat prior to delimitation. Suppose that between 1995-2014, villages a and b were in a gram panchayat with village c , but village c 's population is just above the population cutoff for the 2015 delimitation. Village c is more likely to be split into a GP of its own, meaning that villages a and b will be left in a GP with 1000 fewer people. However, villages a and b need not themselves have populations near the 1000-person threshold, so any effects for them cannot be the result of another state-specific program using that cut-off. To implement this design, for each GP that existed in 2011, we determine the village whose 2011 census population was closest to a population of 1000. We then drop those villages from the sample, but use their population as a running variable for the remaining villages in their GP. As seen in Figure A10, there is a strong discontinuity, albeit with a smaller first stage than the main design. Section D runs and discusses this analysis, showing that the estimated treatment effects are similar with this approach.⁵¹

⁵⁰This might occur because the village is distant from other villages or there are geographical barriers (e.g., rivers) so there are not other villages to agglomerate it with.

⁵¹The other analysis in section D, which estimates heterogeneous treatment effects related to counterfactual GP size, also points against other programs explaining our results. We find the strongest effects for workfare when

Out of the 16 placebo indices that we test, only one estimate is positive and statistically significant at the 5% level. Given the large number of placebo tests we ran, this could easily be due to chance. To better assess this, we conduct a joint test of the placebo coefficients: we generate a 95% confidence interval around the average of the 16 placebo test coefficients (see "Average placebo" in Figure 4).⁵² The average value for the placebo coefficient estimates is small (0.016), and we cannot reject that the average effect in the placebo samples is equal to zero ($p = 0.228$). By contrast, of the 7 tests of polity size (main RD; alternative RD), which we denote using solid bars, 6 out of the 7 are statistically significant at the 1% level, and we are likely underpowered to detect effects on the last one.⁵³ These coefficients have an average value of 0.098 (see "Average RD effect" in figure 4), and we easily reject that there is no average effect of polity size on these outcomes ($p < 0.001$).⁵⁴ We interpret these tests as robust evidence for the role of polity size in our results rather than other programs that use this discontinuity.

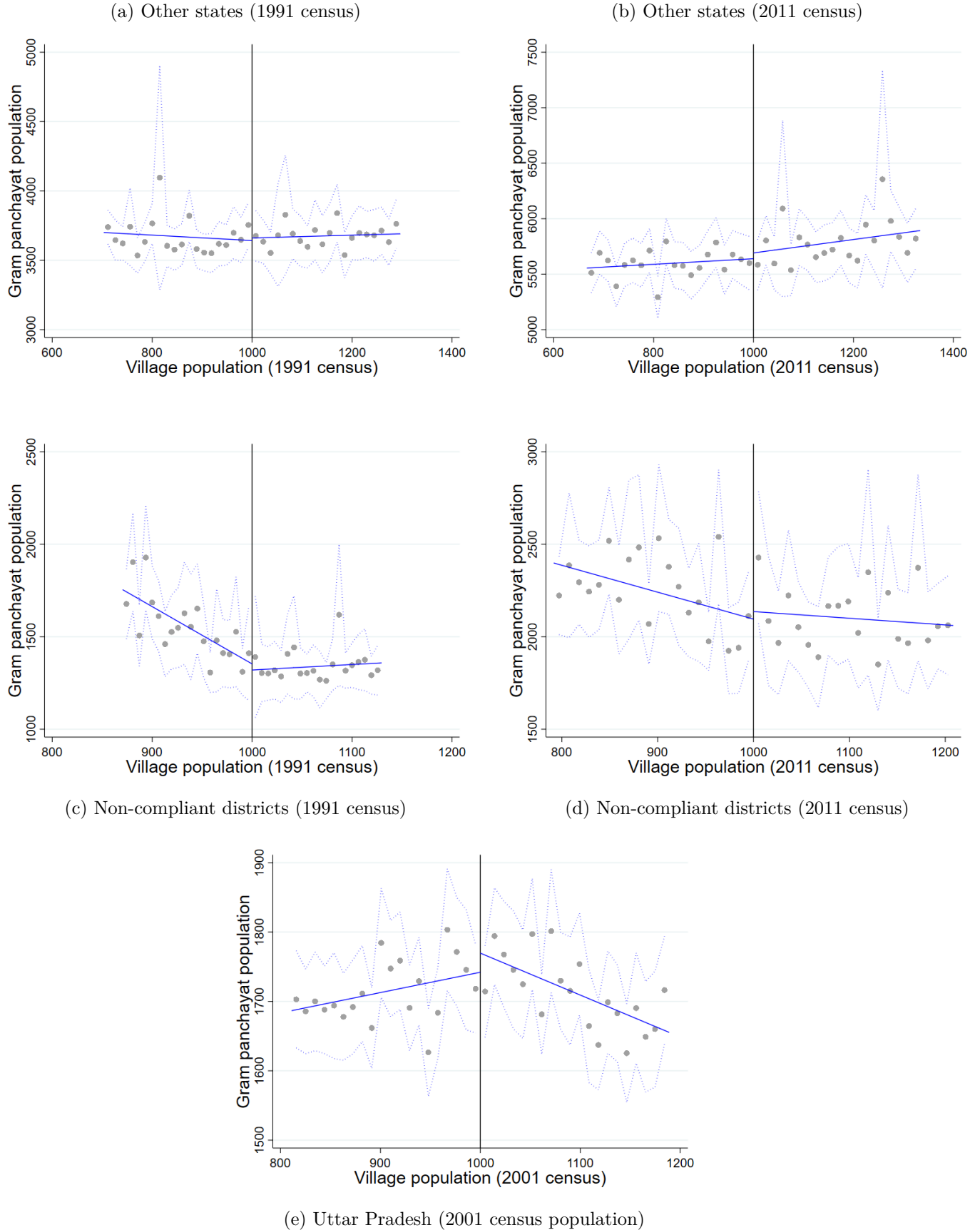
a village shifts from being in a relatively large GP to a solo-village GP. If this effect were simply about other programs that the village gets as a result of being above the discontinuity, we would not expect to see this pattern of heterogeneity.

⁵²To generate confidence intervals, we block bootstrap at the 1995 GP level since that is the unit of variation in polity size for the 1995 delimitation. The 1995 GP boundaries also contain any child GPs in 2015, so this is at worst conservative for the post-2015 delimitation estimates. In the bootstrap samples, we re-estimate all of the 7 main and 16 placebo coefficients and take the averages. We do this 500 times to generate confidence intervals.

⁵³The estimate for this outcome using the main RD is 0.088. However, the first stage effect of the main RD on polity size is twice that for the alternative RD. Even if there were an effect on the programs index from the alternative RD that were half the main RD estimate (due to the smaller first stage for the alternative RD), that would be too small to reliably detect given the magnitude of the standard errors. See section D for more details.

⁵⁴For the purposes of this robustness check, we restrict the main and alternative RD to the samples of villages that are eligible for delimitation, i.e., are in districts that follow the rules, are not solo-village GPs prior to 2015 (for the post-2015 outcomes), and had a sufficiently large GP population prior to the 2015 to be split in 2015 (for the post-2015 outcomes). This produces a clearer comparison between the placebo sample and the sample in which polity size changes take place.

Figure A14: First stage of village population on gram panchayat population in placebo samples



Notes: These figures plot the first stage relationship between the population of the gram panchayat that a village is in and the village's population in the relevant census. Sub-figures (a) and (b) are for states other than Uttar Pradesh. Sub-figures (c) and (d) are restricted to districts in Uttar Pradesh in which the delimitation was not implemented according to the official rules. Sub-figure (e) uses the 2001 village population in Uttar Pradesh as the running variables.

D Spillovers

A concern with the interpretation of these results is the possibility of financial spillovers between communities, and in particular, whether the increase in benefits in villages allocated to a smaller GPs comes at the expense of less spending in untreated villages. In such a zero-sum world, our findings would be just as informative about the mechanisms underlying the effects of polity size, showing that leaders in smaller polities exert greater effort in service delivery. However, the policy implications would be more ambiguous and require consideration of distributional effects. A similar concern is present with any policy executed at scale, such as business training/ loans/ individual job training that may take business from untreated businesses/job seekers rather than expand overall output (Crépon et al., 2013) or other papers measuring government performance as citizen access to public goods (e.g. Martinez-Bravo, 2017). A nice of feature of our context is that it is possible to test for spillovers since the boundaries of spillovers are well-defined.

A key to whether or not negative spillovers occur is whether there are hard budget caps on funding for the relevant item and whether those caps bind in practice. If either condition does not hold – budget constraints are soft or additional spending does not push spending over the cap (and so does not require a loss for others) – then budgetary gains for one GP do not imply losses for another. This may be plausible in this environment, especially given that even an upper bound on the increased costs from the delimitation policy are small in the broader context. For example, an increase in the number of GPs by 10% due to the delimitation rule multiplied by the largest treatment effect we estimate would only induce an increase in spending in that outcome category of around 1%. For most other items, it would be even smaller. We take two broad approaches to understanding spillovers: first, examining budget allocations to assess whether conditions exist for spillovers; and second, testing empirically for spillovers.

We first examine whether budget caps bind by examining whether annual budgets are typically fully spent. Between 2004 and 2009, the Comptroller Audit General of India (CAG) conducted audits on a sub-sample of district, block and village panchayats to determine how effectively the funds were being utilized. This survey was intended to be representative of the state, and although the exact numbers varied year to year, it covered an average of 20 district panchayats, 60 blocks, and 1000-2000 GPs. At the district (*zila*) level, Figure A15 finds a notable level of under-utilization of funding in each year. Significant portions of the budgets between 2006-2009 also remain under-utilized at the block (*kshetra*) and GP levels. If budgets are not being fully spent down each year, then it appears feasible to absorb the additional funding induced by the treatment we study within the slack capacity. However, this might not be true for a substantively larger push for more/smaller polities, a point that we will return to below.

Figure A15: Utilization of funds across the three tiers

(₹ in crore)

Year	Number of PRIs checked	Opening balances	Funds received	Total funds received	Expenditure (per cent in bracket)	Closing balances
Zila Panchayats						
2006-07	52	338.56	476.91	815.47	497.80(61)	317.67
2007-08	52	319.41	589.80	909.21	484.00(53)	425.21
2008-09	55	439.04	993.15	1432.19	1022.87(71)	409.32
Kshetra Panchayats						
2006-07	139	51.19	160.57	211.76	151.53(72)	60.23
2007-08	130	53.33	282.39	335.72	274.59(82)	61.13
2008-09	300 [†]	156.36	532.09	688.45	503.09(73)	185.36
Gram Panchayats						
2006-07	2430	39.18	135.36	174.54	132.32(76)	42.22
2007-08	4525	87.28	376.92	464.20	346.73(75)	117.47
2008-09	3003 [†]	71.85	363.89	435.74	307.84(71)	127.90

Notes: This table is taken from the Comptroller and Auditor General of India's audit of Gram Panchayats in Uttar Pradesh in 2009. The table is in crore rupees (10 million). The data comes from the audit of a representative sample of GPs in the state of Uttar Pradesh.

While the conditions may be present for a lack of negative spillovers, we now directly test for spillovers in the outcomes we study. The concern that we wish to test for is that if a particular local government has a larger number of other local governments with which it is potentially competing for resources, does this worsen access to services for its population? We assess this at the block level, as many development related programs are administered at this level and so this is a likely place at which negative spillovers could emerge.

The ideal experiment would be one in which the average GP size is randomized at the block level, producing cross-sectional variation in the density of political units across blocks to account for spillovers or other general equilibrium effects at the block-level. We approximate this experiment in our data by again leveraging the discontinuity in GP size. We use variation coming from whether the block has a larger fraction of villages just above the threshold as opposed to below: blocks with a higher proportion of villages above the cutoff will have a larger number of GPs, but are plausibly otherwise similar to those with a higher proportion just below. More precisely, for each village, we count the villages in their block with a population of between 800 and 1200 people. Leaving out the village itself, we calculate the fraction within the same block and this bandwidth that are above the population cutoff of 1000. We use this as an instrument for the total number of GPs in the block, and then examine how the instrumented total number of the GPs in the block is related to the main outcomes of interest.

The first column of panel A of table A17 tests the first stage of whether having a higher fraction of villages above the cut-off is indeed related to the total number of GPs in the block for the 1995 delimitation exercise.⁵⁵ We find that it is: on average, there would be 16.5 more GPs among the blocks where all the villages were above the threshold as compared to those with zero. Since the average block has a total of 69.3 GPs, this shock could have a moderate impact on intra-GP

⁵⁵In all these regressions, we also control for the total number of villages in the block and village population.

competition for resources, and the first stage F-statistic is strong (Kleibergen-Paap F-statistic of 31.5). Column (1) of panel B does the same thing with the 2015 instrument. The first-stage is weaker, with an associated increase of 10.0 GPs among blocks where all the villages were above the threshold as opposed to none and a Kleibergen-Paap F-statistic of 7.7.

One concern is that the fraction of villages above the cutoff could plausibly be related to other characteristics of the block that determine public goods delivery and so the estimates in the above equation could be biased. The remaining columns of panels A and B conduct placebo checks for whether the fraction of villages above the cutoff at the time of delimitation is related to village characteristics in the previous census round. We do not find any evidence of this, suggesting that this approach can be used to test for spillovers. This is particularly encouraging for the 1995 delimitation exercise given that many of the pre-1995 variables are the same as our outcome variables, indicating a lack of pre-existing imbalances on those outcomes.

In panel C, we run the IV regression for the five indices that are the main study outcomes. We instrument for the total number of GPs in the block using the fraction above the cutoff as compared to below, with the first stage Kleibergen-Paap F-statistic displayed below each of the regression output along with the p-value on the coefficient. We do not observe statistically significant negative spillovers for the post-1995 outcomes. The coefficient on household infrastructure is positive and statistically significant, which is driven by toilet construction. This could potentially be due to positive knowledge spillovers or yardstick competition, where villages see elevated public goods delivery for their neighbors and so demand greater services; we are testing for this in a separate project.

For the post-2015 outcomes, the coefficient on the index of individual programs is not statistically significant, but the one for workfare is negative and barely statistically significant at the 10% level ($p = 0.098$). It could be that NREGS fund allocations have tighter budget caps, and so are more subject to negative spillovers. But given how marginal this statistical significance is and the weakness of the first stage (KP F-stat = 7.7), we interpret this as at best weak evidence of negative spillovers. In general, the post-2015 analysis may be underpowered, with much larger standard errors than the post-1995 outcomes due to the weaker first stage.

The lack of large negative spillovers is consistent with the modest variation in the number of GPs induced by delimitation. In most cases, the delimitation rules caused a difference of only 2-4 GPs per block. If this were multiplied by the estimated treatment effects on the outcomes we study, this would have a modest effect on block-level expenditures that could be accommodated with modest slack; thus we would not expect large negative spillovers. If the government were to undertake a delimitation exercise that produced significantly more local government units such as doubling the total number of GPs, then negative spillovers will be more likely to occur. However, for meaningful but more limited exercises, like the one we study, the overall effect appears positive.

Table A17: Block-level spillovers

<i>Panel A: First stage and placebo checks (Post-1995)</i>						
	First stage	Placebo checks (1991)				
	Total GPs	Primary school	Middle school	Paved road	Literate pop.	SC pop
Frac above cutoff	16.482 (2.939)	0.018 (0.029)	-0.005 (0.016)	0.046 (0.044)	-16.997 (23.079)	-0.031 (0.022)
Dep var mean	69.273	0.588	0.118	0.421	307.199	0.244
Observations	95170	95170	95170	95170	95170	95170
<i>Panel B: First stage and placebo checks (Post-2015)</i>						
	First stage	Placebo checks (2011)				
	Total GPs	Primary school	Middle school	Paved road	Literate pop.	SC pop
Frac above cutoff	9.991 (3.607)	-0.027 (0.030)	-0.008 (0.030)	0.041 (0.057)	-15.530 (31.316)	-0.012 (0.021)
Dep var mean	79.323	0.718	0.409	0.660	871.474	0.238
Observations	93855	95102	95102	94759	95103	95103
<i>Panel C: Outcomes</i>						
	Post-1995 outcomes			Post-2015 outcomes		
	Education	Village Infra.	HH Infra.	Indiv. programs	Workfare	
Total GPs	0.000 (0.003)	0.001 (0.006)	0.012 (0.006)	-0.013 (0.010)	-0.029 (0.018)	
K-P F-Stat	31.45	31.45	31.54	7.67	7.96	.
p-value	0.888	0.834	0.030	0.202	0.098	.
Dep var mean	-0.206	0.110	-0.037	0.000	0.000	
Observations	95170	95170	95110	93855	413018	

This table reports reduced form estimates of the spillovers of delimitation on nearby villages. We include controls for the total number of villages in the block and the village's own population. Panel A estimates the first stage between the total number of GPs and the instrument for the 1995 delimitation episode, while the remaining columns run placebo checks of the instrument on pre-existing covariates. Panel B does the same for the 2015 delimitation episode, while panel C reports the reduced form relationship between the instrument and the outcomes. Standard errors are clustered at the gram panchayat level.

E Financing

Section IV showed that there is an increase in provision of infrastructure and public services. It is important to understand the financing of these goods, and in particular, whether this is the result of mechanical funding rules that may favor smaller polities. In that case, even though our estimates would be internally valid for India, they would not generalize to contexts with different fiscal rules or necessarily demonstrate generalizable mechanisms. If this is not about mechanical funding rules, we can interpret our results as reflecting mechanisms such as greater representative effort or reductions in corruption, which generalize more easily to other contexts.

Section C noted that fiscal formulas should cause larger and smaller GPs to have nearly equal budgets on a per capita basis, but we confirm this with data from the Panchayati Raj accounting system. Table A15 uses our main RD approach to measure whether polity size affects local government budgets between 2013 and 2020. Panel A examines spending on a per capita basis, while panel B looks at the total budget. Consistent with the stated fiscal funding formula, smaller GPs do not have more spending on a per capita basis (panel A, column 1). As a result, they have smaller *total* spending (panel B, column 1), which may limit larger projects. However, this would go in the opposite direction of what we find, and so cannot explain our results.⁵⁶

In addition to testing the government budgets, another potential source of funding is through local taxation. Even though local taxation is limited, taxes levied by local government may be related to local government size; smaller local governments could be more empowered to impose taxes, but also may have poorer capacity to do so. Applying the same discontinuity and using the same data, column 2 of Table A15 finds no evidence of a relationship between polity size and per capita local taxation. The relatively limited scope of local taxation is both a strength and limitation of our setting. Results may differ in contexts in which local governments have more substantial taxation powers, but limited local taxation is common in low capacity states, and so our results likely generalize quite broadly.

We next consider whether there is a state or national funding formula that causes smaller local governments to be *mechanically* favored for the set of outcomes that are not funded by the local government budget. We showed previously that the smaller local governments receive more spending per-capita for programs like NREGS. However, the question is whether funding rules mechanically produce this effect, which would not generalize to other contexts, or the higher spending reflects generalizable aspects of polity size, such as how it can affect local politician effort and performance.

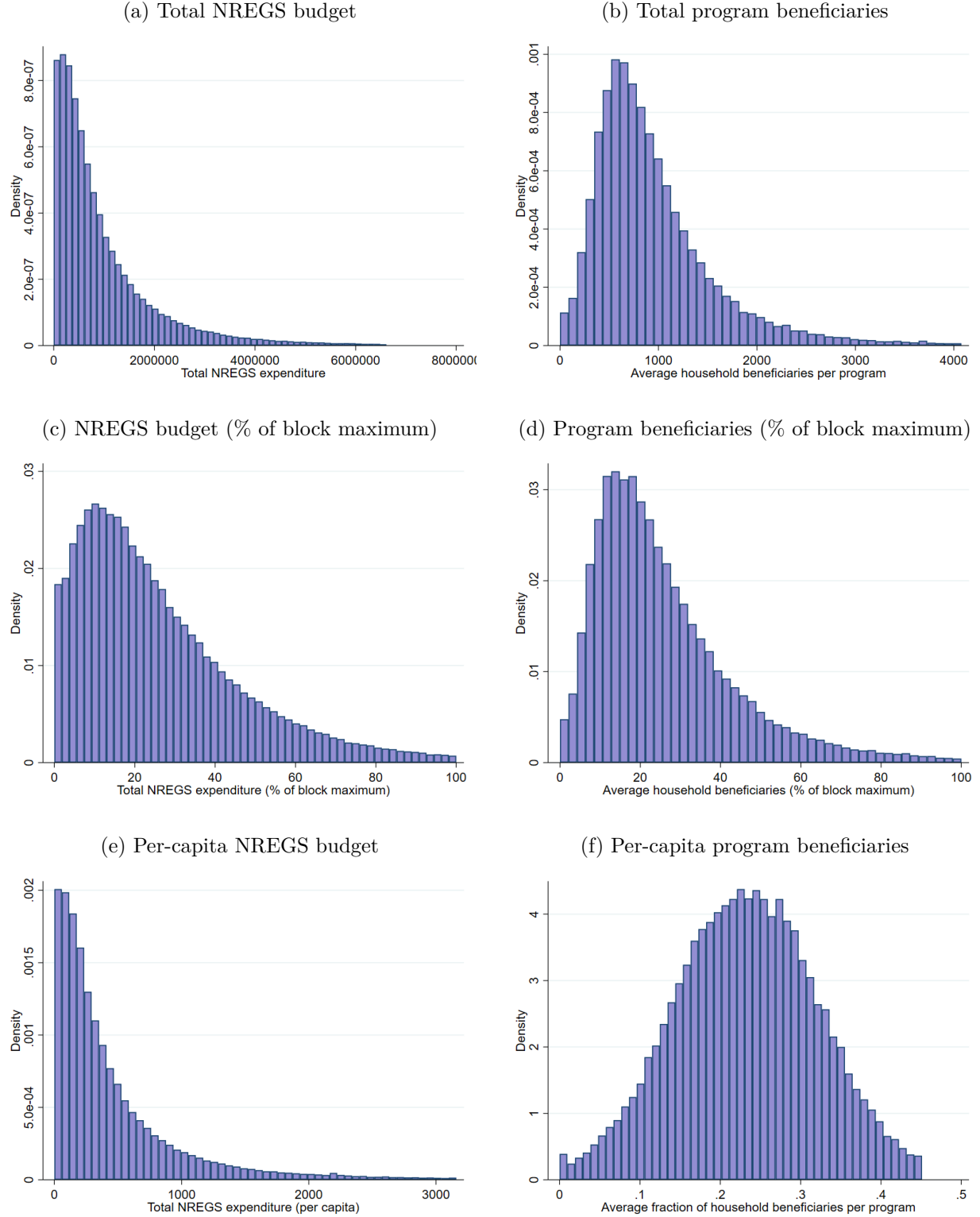
To test this, we consider the sets of rules that could mechanically lead to smaller polities receiving more funding on a per capita basis, in particular: (1) GPs receive fixed allocations that are invariant to population, so smaller GPs get more per capita; or (2) there are spending caps or spending floors defined in absolute rather than per capita terms, such that GPs at the cap or floor will spend more

⁵⁶Column 3 of Table A15 also finds no effects on spending for specific tied projects through the local government budget.

per capita if they are smaller. These possibilities can be tested with visualizations of program-specific allocations. We focus on NREGS and the set of eight individually-targeted programs since those are our best measures of funding allocations. For NREGS, we measure the total amount spent in a given year. For the individually-targeted programs, we approximate spending with the total number of beneficiaries since the per-person spending within a program will be similar.

Panels A and B of Figure A16 plot the total GP-level spending and number of beneficiaries respectively; the histograms are similar if we instead look at disaggregated outcomes within workfare (e.g., number of NREGS job cards) or the number of beneficiaries for each program individually. There is substantial variation, demonstrating that the first hypothesis of population invariant budgets is incorrect. Panels C and D test for caps or floors. For each GP, we calculate the maximum amount of spending (beneficiaries) among other GPs within the same block and the GP's spending (beneficiaries) as a fraction of that amount. If there were binding spending caps, then there should be bunching at a value of 100%; instead, there is no more mass there than would expected by chance. Similarly, if there were floors, we would expect bunching at non-zero floor amounts, but instead there are many zeros. Finally, panels E and F plot per capita spending. The dispersion across GPs is meaningful, where per capita NREGS spending amount is five times higher in the top quarter than the bottom. This points to non-mechanical funding mechanisms explaining this dispersion.

Figure A16: Gram Panchayat-level program allocations



Notes: These figures plot the distribution of allocations to each gram panchayat in Uttar Pradesh between 2013 and 2020 for the NREGS workfare program. Panels (a) and (b) plot the total annual NREGS spending, while panel (b) plots the total number of beneficiaries of the eight programs we study. Panels (c) and (d) measure spending as a fraction of the maximum within the block. For each GP, we calculate the maximum amount of spending (beneficiaries) among other GPs within the same block and then estimate the spending (beneficiaries) in the GP as a fraction of that amount. Panels (e) and (f) plot per capita NREGS spending and number of beneficiaries.

F Cost-Benefit Analysis

Our analysis has highlighted that the policy generated increases in many services, but it is not clear whether the benefits of these services exceeds their costs. To understand these trade-offs better, we conduct some back-of-the-envelope cost-benefit calculations, anchoring our calculations over a 10-year period. This is divided into two parts. The first estimates the direct administrative costs resulting from the creation of a new polity; personnel, physical infrastructure, and elections. The second is more speculative, and relates to whether the additional services being provided themselves provide benefits that exceed their costs.⁵⁷

Direct costs For personnel costs, we use information on the wages/salaries of ward councilors, the pradhan (leader of the GP), and the panchayat secretary, as the new polity will lead to an increased number of such representatives. First, ward councilors are traditionally not paid for their work.⁵⁸ An allowance for ward councilors to attend meetings was only introduced in 2021, and so this does not impose additional costs over our study period. Pradhans however, were given a monthly honorarium of around \$42 per month (3500 rupees per month). The amount was fixed over an extended period of time. In order to discount the gains and costs, we use the government of India interest rate of 5%. This totals to about \$4101 accounting for discounting over a 10-year period.

The last personnel cost is payments for the panchayat secretary. From recent online job postings, we find that the cost of salary and benefits can range from roughly Rs. 30,000 to Rs. 60,000, and so we take the midpoint of that range for our estimate. However, personnel records show that each GP does not have its own secretary: on average, a panchayat secretary in Uttar Pradesh is responsible for an average of over ten GPs and around fifteen thousand citizens. Furthermore, when a new GP is created, this often is just assigned to the same secretary as the previous GP, and so the addition of a new GP leads to only a 4% reduction in the ratio of secretary to citizens (table A14). We therefore multiply the payments to secretaries by 4% to arrive at an annual additional salary expenditure of roughly Rs. 21,600 (\$260). This totals to \$2108 over a 10-year period using a 5% discounting rate.

Next, there are associated physical infrastructure costs from the construction of a physical building that serves as the local government headquarters. However, in many cases, there is not a physical building, and the pradhan simply uses their house as the headquarters for interfacing with citizens (Bussell, 2019). This is particularly true in Uttar Pradesh, where only 56% of GPs

⁵⁷Note that we do not compare the aggregate costs and benefits of the introduction of the decentralized local government system in general; rather, we measure the marginal costs and benefits stemming from additional local government units within this system.

⁵⁸<https://www.hindustantimes.com/cities/lucknow-news/yogi-hikes-honorarium-of-panchayat-representatives-101639591660651.html>

have their own buildings.⁵⁹ To assess the costs of construction, we take information from a Ministry of Panchayati Raj press release discussing a recent push for construction of these buildings (Ibid), which budgeted Rs. 20 lakh (\$25,000) for each building. We estimate the expected additional cost from physical infrastructure to be 56% of the total cost of constructing a panchayat building, or \$14,000.

Finally, election costs are relatively straightforward to assess as they should not be any different as a result of the polity size reforms. Since delimitation happened simultaneously with new elections, elections would anyway have been conducted at the same time. Furthermore, polling booths are set up at the village level (typically within schools or other existing government buildings), with an average of one booth for every 900 voters following Election Commission of India guidance that citizens should not have to travel more than 2km to vote. While we do not observe polling stations in our data, this strongly suggests that the total number of polling stations and polling station attendants will remain unchanged. This is also a once in every five years expense, so even if there were a small increase, this should be quite small relative to the other cost categories, so we ignore this in the calculation.

Summing these together, we thus would interpret the marginal cost to government as around twenty thousand dollars for each additional polity created over a ten year time horizon (or about \$2 annually for each affected person).

The second part of this exercise is related to the costs and benefits of the outcomes we measure. We elect not to conduct a full cost-benefit calculation in which we aggregate costs and benefits across goods to arrive at a single cost-benefit number. We made this choice for a few reasons. First, many benefits are very difficult to calculate credibly. For example, NREGS wages or possession of a below poverty line ration card entails a transfer from the government to a citizen, and so putting this in cost benefit terms requires making assumptions about the value of a marginal dollar in the hands of a citizen rather than the government. In other cases, our data does not allow us to fully measure the benefits – for example, much of the school construction occurred between 2001 and 2011, but we only can measure educational attainment in 2011. That doesn't give enough of a time gap to see how attainment is affected in steady state. Given these and other related concerns, we instead provide estimates of costs of delivery and some of the benefits, but without making a final determination of cost effectiveness based on these numbers.

Workfare and Individually Targeted Benefits The paper finds relatively large increases in the provision of workfare and individually targeted benefits. As noted, these involve a transfer from the state to the individual, and so are hard to incorporate into cost-benefit calculations without

⁵⁹<https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1654695> ; in the rest of India, 78% of GPs have their own building.

making assumptions about the relative returns to either party holding the transfer amount. This is further complicated by the fact that many of these transfers are targeted to extremely poor households, for whom the marginal utility of additional income is likely to be high. We do not attempt to make this calculation here, but simply provide information on the monetary value of these transfers.

According to table 3, the main gains for citizens come from NREGS earnings, BPL cards, health insurance, the Saubhagya electrification scheme, LPG gas connections, housing benefits, and access to bank accounts. The NREGS earnings are easiest to quantify: the average resident of the villages earns Rs. 63 (\$0.77) more annually (panel B, column 4 of table 3). For BPL cards, taking current rice and wheat prices, we estimate the average value to a household as around Rs. 800 monthly (\$10) for a family of five, so with an increase of 5.1 cards per village, that is an average of around \$0.60 per citizen.⁶⁰ The value of health insurance, electrification, and having a bank account are harder to measure, so we leave these out of our calculations. However, the average value of a new LPG connection is around Rs. 1600 (<https://www.pmuy.gov.in/>), so with the gain of 5.64 households per village, this comes to \$0.11 per citizen. The value of housing benefits is much larger, at around Rs. 70,000 per household,⁶¹ and so with an increase of 5.55 beneficiaries per GP, this comes to \$4.85 per resident. This suggests that transfers increased by over \$6 per resident annually, even if we do not include individual-level benefits in which there were gains like health insurance. This value already exceeds the direct costs of the policy.

Education We next examine the field of education. To assess the benefits derived from improved educational outcomes, we combine our results on the change in educational attainment generated by the reform with estimates provided in Khanna (2023), focusing our cost calculations on primary schooling improvements. Using the 2011-2012 Socioeconomic and Caste Census (SECC), we can measure the total number of individuals at different levels of education attainment in the village (illiterate, primary, middle, etc.). Our estimates imply that the reform led to approximately 14 more children completing primary schooling.⁶² As a baseline, we assume that completion of primary schooling conservatively implies at least one more additional year of schooling. Khanna (2023) finds that the presence of primary schooling corresponds to a 13.4% increase in return per year of education in 2010. As a baseline return, we use the agricultural day wage of \$1.21, obtained using daily wage data from the CEIC⁶³ for agricultural workers in 2011 to calculate the increase in

⁶⁰For all of these calculations, we divide the total gain to citizens in the new polity by 1000 people, the number of residents of the polity around the discontinuity.

⁶¹<https://prsindia.org/budgets/parliament/demand-for-grants-analysis-rural-development>

⁶²Our main analysis focuses on two sets of educational outcomes - the first is the existence of primary and middle schooling, and the second is educational attainment. Given that educational attainment is a direct measure of the effect, we use this estimate in our analysis.

⁶³<https://www.ceicdata.com/datapage/en/india/minimum-daily-wage-rate-uttar-pradesh/minimum-daily-wage-rate-uttar-pradesh-agriculture>

returns. An increase of 13.4% leads to a wage of \$1.37 per day. Assuming 300 days worked in a year, this is an increase of \$48 per year per person, which gives a benefit of \$4670 over a 10 year period (including a discount rate of 5%) across 14 extra children per school.

In order to calculate costs of education, we use expenses incurred due to teacher's wages and infrastructure costs. Teacher wage data is taken from Kingdon and Muzammil (2009).⁶⁴ In 2006, the average per month wage for a grade school teacher was \$54.9 (4500 rupees) in Uttar Pradesh under the Fifth Commission Pay Scale (1994 pay rates) Kingdon and Muzammil (2009). Assuming a full 12-month payment cycle, this gives a per-year salary of \$658. Accounting for a 5% discount rate, this totals to \$5341.4 per teacher over a 10-year period. According to CEIC data, on average there are 2.78 teachers per school⁶⁵. This takes the costs up to about \$14,849 per school for the 10-year period.

Next, there is a one-off cost for the construction of the school. While there may be long-term maintenance costs, our time period is a relatively short horizon in terms of building maintenance, therefore we focus on just the initial investment covering a 10-year period. We take the estimates of classroom unit costs from the World Bank (2003)⁶⁶, which finds that in India on average the per unit cost is approximately \$3400⁶⁷. Analysis on Indian education from National University of Educational Planning and Administration (NUEPA) (2010)⁶⁸ has found that a little more than 40% of Uttar Pradesh rural elementary schools (primary and upper primary) in 2008 had fewer than 4 classrooms per school. Given that we focus on villages around the 1000 person threshold, and to be consistent with the teacher numbers, we assume there are three classrooms on average. This gives us a total infrastructure cost of \$10,200.

Just looking at the raw numbers would suggest that the costs of education exceed the benefits from educational infrastructure construction. However, the benefit calculations almost certainly understate the actual benefits since the time between the construction of schools (typically over the 2001 to 2011 period) leaves little time for the improvements in educational attainment to occur and be measured in the 2011-12 SECC. In particular, we might expect to see more responses in terms of middle or secondary school completion if we had educational data from a later date. We also assume that completion of primary school corresponds to only one additional year of education and don't have a good way of valuing the reduction in travel time for children from having a school in their own village, which is likely to be a significant benefit beyond effects on educational attainment.

⁶⁴This paper additionally finds that wages make up 94.7% of total expenditures on average in 2006, so we use this as the main expense.

⁶⁵This is calculated using the total number of primary school teachers in 2006 divided by the number of primary schools in 2006

⁶⁶ The data is taken from a World Bank review report on educational costs that can be found here: <https://documents1.worldbank.org/curated/pt/984571468765318102/pdf/268890PAPER0Education0Notes0Construction.pdf>

⁶⁷It was \$3700 in 1993 and \$3100 in 2001. We take a simple average of these two numbers)

⁶⁸https://ras.org.in/statistics_on_elementary_school_education_in_rural_india#bib7

Roads To evaluate road provision, we begin by calculating the average infrastructure cost of constructing a road in a village. We utilize data from villages in Chhattisgarh, Gujarat, Madhya Pradesh, Maharashtra, Orissa, and Rajasthan provided in Asher and Novosad (2020) since there are no specific estimates available for Uttar Pradesh. The paper finds that the average cost of building a road in a village amounts to approximately \$150,000, while the gains in terms of additional consumption per year per village are estimated at \$3,945. We average these out over a 10 year period, as there may be depreciation of road quality after this point, which would require further investments. Discounting the consumption gains using the interest rate of 5% gives us a total gain over the 10 year period of approximately \$31,985 per village. In terms of pure numbers, the cost of investment in road improvements seem to significantly outweigh the potential consumption gains, as we compare a cost of \$150,000 to a gain of \$31,985. This calculation however is measured in terms of consumption gains relatively soon after the roads were built. This does not account for long-term gains over the next few decades.

Sanitation Finally, for sanitation, we take benefits from improved sanitation from a UNICEF report in India, which breaks them down into four categories.⁶⁹ These aspects include averted medical expenses, time saved, reduced mortality rates, and increased property value resulting from access to a toilet. Collectively, these benefits amount to \$808.40 per year⁷⁰ Since on average around 6 more households in treated villages have access to toilets, we calculate a per village per year benefit of \$4850. In order to generate a cost measure, we use the cost of construction of a toilet in a household, based on calculations by Gautam (2023). This paper finds that on average constructions costs \$112 inclusive of raw materials and labor costs per household, so the benefits appear to far exceed the costs.

G Supplementary Analysis

A Correlation of outcomes with political mechanisms

In the mechanisms section, we provide evidence that smaller polities experience increases in civic engagement and modest decreases in criminality of elected leaders. However, it could be that civic engagement/criminality are not actually related to the public good outcomes that are the focus of the paper. While demonstrating a causal link between these mechanisms and the outcomes is outside of the scope of this paper, this appendix examines correlations.

⁶⁹<https://www.unicef.org/india/media/1206/file/Financial-and-Economic-Impacts.pdf>

⁷⁰The breakdown of benefits is based on a per year per household calculation. To break it down further, this includes \$124 per year in reduced medical costs, \$382 per year in the value of time saved, \$273 in the value derived from lower mortality rates, and \$294 in added property value over a 10-year period, which we simplify to calculate the yearly benefit per household.

First, we look at the relationship between our measures of civic engagement in section A – participation in village meetings and voter turnout – and the two key post-2015 outcome indices (workfare and individual-level programs). We begin by looking at attendance at Gram Sabha meetings in the year 2018 and how this is related to these outcomes in the following years. The Gram Sabha functions as a place where residents can make their voices heard in local government. Panels A and B of Figure A17 show the correlation between meeting attendance and these outcomes. This relationship is strong and positive, holding even when we reduce our bandwidth to look just at villages that fall within 300 people of the 1000 person cutoff (panels (c) and (d)).⁷¹

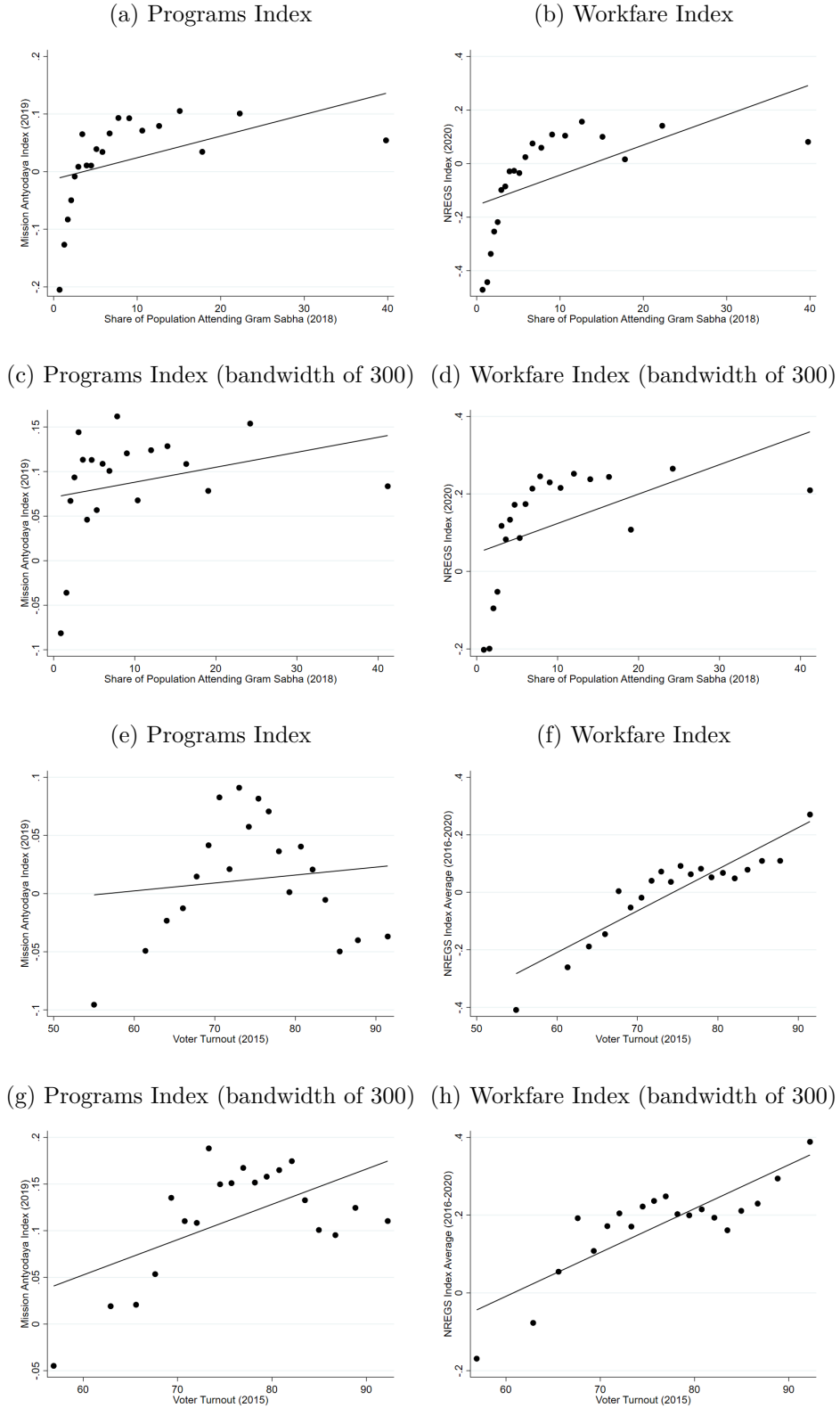
We next explore the connection between voter turnout and the provision of public goods. To gauge voter turnout, we rely on data from the 2015 GP election, which was the nearest election round preceding our available data. Panels (e) and (f) find strong correlations with voter turnout for both outcomes, particularly for levels of voter turnout below 80%. This relationship holds even when we reduce our bandwidth to look just at villages that fall within 300 people of the 1000 person cutoff (see Panel (g) and (h)). These are simply correlations, but they suggest that there could be a relationship between civic engagement and public good provision at the GP-level.

The paper also examined the effect of polity size on criminality of elected candidates, where the relationship between criminality and public good provision is not necessarily clear (Vaishnav, 2017). As an additional analysis, we look at how criminality of winning candidates was correlated with public good provision. Our criminality data comes from election affidavits for the 2020 election, and we correlate this with the same set of outcomes highlighted previously.⁷² Table A18 presents the results of this regression, run at the GP-level, controlling for GP population. The Programs Index has a significant negative correlation with the election winner’s criminal history, indicating that GPs with an elected leader who had a criminal record tend to have a worse service delivery, but there is no significant relationship with the workfare index. Vaishnav (2017) demonstrates that the relationship between criminality and politician performance is theoretically unclear, so it could be that this relationship is only present for certain types of public goods. Regardless, the correlations we present are at best suggestive, although consistent with what has been found in other papers on politician criminality in India (Prakash et al., 2019)

⁷¹We also run these specifications with additional controls such as population size, average village income, and literacy rates to try to control for other factors that may be related to both civic participation and public goods provision, and find no significant differences in the relationship.

⁷²A major and unfortunate limitation of the criminality analysis is that our outcome measures pre-date or are concurrent with the only election in which we can measure criminality (2020). However, some of the candidates are likely to be incumbents, or in cases where they cannot run themselves due to gender reservation, may run one of their family members as candidates. Thus criminality of the village leadership in 2020 is likely correlated with criminality in 2015.

Figure A17: Political participation and public service provision



Notes: These figures plot the relationship between civic participation and public good provision. Civic participation is measured by two variables: (1) Percentage of the population in attendance at the Gram Sabha in 2018 (panels a to d); and (2) voter turnout in the 2015 GP elections (panels e to h). Public good provision is measured by two indices - the workfare index and individual-level programs index.

Table A18: Correlation between Criminality and Public Service Delivery

	Mission Antyo- daya Index (2019)	NREGS Index (2020)
Winner Criminal History (2020)	-0.043 (0.02) [0.014]	-0.017 (0.02) [0.437]
Dep var mean	0.091	0.202
Observations	16872	16878

B Ratio of representatives to citizens

Another potential mechanism for the effect of smaller polities is a higher ratio of representatives to citizens. When a representative serves a smaller group, it is the same cost to provide a higher per-person level of service, so they may provide more. In the context of the gram panchayat, there are three types of representatives who may be relevant: the panchayat secretary, who is the bureaucratic counterpart to the pradhan and carries out village-level bureaucratic tasks; ward councilors, who each represent a jurisdiction of a few hundred people; and the village head/pradhan.

To test this, we first examine the panchayat secretary. In Uttar Pradesh, a single panchayat secretary typically serves many GPs due to staffing shortages so it is not clear that being in a smaller GP would meaningfully improve the ratio of panchayat secretary to citizens for a village. To test this, we collected personnel records from the Ministry of the Panchayati Raj that give the current allocation of panchayat secretaries for all the GPs in Uttar Pradesh. The average panchayat secretary serves many different GPs and an associated population of around fifteen thousand people (panel A of Figure A13). Panel B of Figure A13 and column 1 of Table A14 show how the number of citizens to secretary varies across the discontinuity. There is a modest drop of around 4% when moving across the discontinuity, which certainly could improve service delivery, but seems unlikely to account for most of the effects.

As a back of the envelope calculation of how much this drop could affect service delivery, we take the cross-sectional relationship between number of citizens per secretary and the indices of individually-targeted programs and workfare outcomes. This is plausibly an upper bound on the causal relationship between workload and performance on those outcomes given that areas with a worse ratio are likely to also be disadvantaged in other ways, such as being more rural. If we multiply these cross-sectional estimates by the effect on secretary workload induced by the discontinuity (ratio of citizens to secretary; Table A14), we find that the better ratio of panchayat secretaries to citizens

would predict improvements of 0.0012 SD and .0036 SD in those outcomes respectively. That is only 1.3% and 2.4% of the effects found in the paper, suggesting that workload changes for the panchayat secretary do not explain much of the findings.

Next, we investigate the role of ward councilors, as the ratio of ward councilors to citizens improves by around a quarter across the discontinuity (column 2 of Table A14). In Uttar Pradesh, the number of representatives depends on the population of the GP: a GP will have 9 councilors if its population is less than 1000, 11 if the population is between 1000-2000, 13 if between 2000-3000, and 15 if more than 3000. Since the number of representatives increases less than linearly with GP population, smaller units have a higher ratio of ward councilors to citizens. This higher ratio may make it easier for households to make claims on the state.

We test this by exploiting the discontinuities in number of elected representatives at cutoffs in GP population. Note that this concerns the population of the *gram panchayat* rather than the *village*, so the cutoff of 1000 people will often differ from that for village population used in the rest of the paper. Nonetheless, we focus on the cutoffs of 2000 and 3000 to avoid the cases of overlap. We combine the two discontinuities in columns 3 and 4 of Table A14 and find no evidence that a higher ratio of councilors to citizens affects either the index of NREGS outcomes or individual-level programs. We are currently investigating this across all the states in India in a separate paper (Narasimhan and Weaver, 2023), but this demonstrates that this channel is not relevant here.

Finally, there is a higher ratio of pradhans to citizens in villages to the right of the discontinuity. We don't have a way of measuring the importance of this mechanism as there is no independent variation across units. However, this mechanism will always be present in local government units that have some sort of executive head (which is standard globally), so is a highly generalizable part of the creation of smaller polities even if we cannot quantify its importance.

C Economies of Scale

A commonly cited reason for a preference for larger polities is that they can take advantage of economies of scale. In particular, there may be some public goods with high fixed costs and low variable costs such that the average cost of provision is lower in more populous jurisdictions. In our context, outcomes are typically better in smaller polities, but this may not be true for a subset of outcomes in which economies of scale are more relevant. To assess this, we divide the services affected by local governments in India into three groups – (1) physical infrastructure, (2) individual benefits (e.g. possession of a below poverty line ration card), and (3) workfare. For the individually targeted benefits, GP leaders carry out tasks such as registering individuals for programs and helping them access their entitlements. These do not have significant fixed costs and the variable cost should be similar regardless of the size of population served. Thus there is not a strong reason to expect economies of scale there. Similarly, for providing workfare opportunities, the total costs

should scale along with population such that economies of scale would not be expected.

Economies of scale are most plausible for physical infrastructure. A good example of this would be a paved road that connects a village to the state highway network. The cost of building a connector road does not depend on how many people live in the village, but the distance between the village and the state highway network. This cost will be amortized over a larger number of people, generating economies of scale. Another example would be building a school, for which there are fixed costs of construction or employing a teacher, although here the variable costs are more pronounced than in the case of a road (e.g. building a physically larger school or employing more teachers to accommodate more students).

However, in considering the relevance of economies of scale in the context of polity size, it is important to note how polity size reforms work. They do not change the underlying geographical distribution of the population or increase the population at a particular point in space. Instead, the spatial distribution of the population is fixed and simply grouped into different local government units. This is important because physical infrastructure in this context has a spatial component. For many types of physical infrastructure, an individual only benefits if this is constructed in their village. For example, suppose that village A and B are located near one another and may be agglomerated together into a local government. If A is electrified and B is not, then residents of B get the same amount of benefit (i.e., no benefit) regardless of whether they are in the same local government as A.⁷³ As a result, the cost-benefit ratio of these types of goods does not change as the population of the polity expands because the spatial distribution of the population is the same on either side of the discontinuity.⁷⁴

For other types of physical infrastructure, an individual can benefit even if it is not located physically in their village – e.g., if village A has a road built to connect it to the state highway network, individuals in B can travel to A in order to travel on that road. However, note that use of the road is not excludable based on one's local government – an individual in B can travel on A's road regardless of if they are in the same local government unit. That is also true for other types of physical infrastructure constructed by the government in the rural Indian context: e.g., if village A and B are located near one another, and there is a middle school in village A but not village B, children from B can still attend school in A even if they are not in the same GP. As a result, the cost-benefit calculation does not change with polity size or across the discontinuity. Economies of scale in population would be present if the increases in polity population came from an increase in the population at a particular point in space (e.g. the population of village A increases), but this

⁷³There may be some benefit from having the electricity network run closer to one's village in terms of the potential for future network expansion, but that will not depend on local government boundaries since electricity is not managed by the local government.

⁷⁴A classic economies of scale public good like national defense – the main example cited in Alesina and Spolaore (2005) – does not have this spatial feature, where spending on national security provides similar benefits across citizens regardless of their exact location in the country.

is not what polity size reforms do.

Is there another type of infrastructure that local governments in our context provide and for which there may be economies of scale? In a recent meta-review, Gómez-Reino et al. (2021) summarize the types of local public goods for which researchers have found evidence of economies of scale: these include power grid services, provision of transit, police/fire services, water and sewage, and waste collection. For each of these services, the local governments that we study in this paper are either too small to independently administer it (power grid, provision of transit, police/fire services; all of these are managed at a higher level) or the technology used is relatively rudimentary in this context so that we would not expect the relevant economies of scale to be triggered (water and sewage, waste collection).⁷⁵ Our take-away is therefore that for local governments of this size, the disadvantages from loss of economies of scale do not appear to be a major factor that reduce the benefits of smaller polities. Note however that the change induced by this discontinuity is the right estimate for policy – when policymakers draw polity boundaries, they are organizing the existing population rather than shifting its location around. Whereas it would be interesting to know what happens if polity size increases as a result of the population in a given location increasing, forced migration or similar policies that accomplish such a goal are atypical policies in democratic contexts.

Given this, classic economies of scale turn out not to be so relevant here. There are, however, related mechanisms that could be relevant. Since budgets are fixed on a per-capita basis, polities with a larger population have a larger *total* budget. As a result, they may be able to produce public goods that are costly and would not be possible in smaller polities with smaller budgets. If this mechanism is relevant, we would expect a higher number of more costly services and a smaller number of lower cost benefits in those polities. We do not see this. This is potentially because the GP budget is insufficient to produce really big ticket items – for example, the average cost of a paved road to the external paved road network is much larger than the average annual budget of the typical GP in our sample (and so is paid by a higher level of government; see Asher and Novosad (2020) for a discussion of roads in rural India). It could also be that countervailing forces related to coordination negate the benefit of a larger budget: if village A and B are united in a larger polity, but the physical infrastructure must be located in either A or B, contestation over location may mute the benefits of a larger budget.

D Polity size and preference heterogeneity

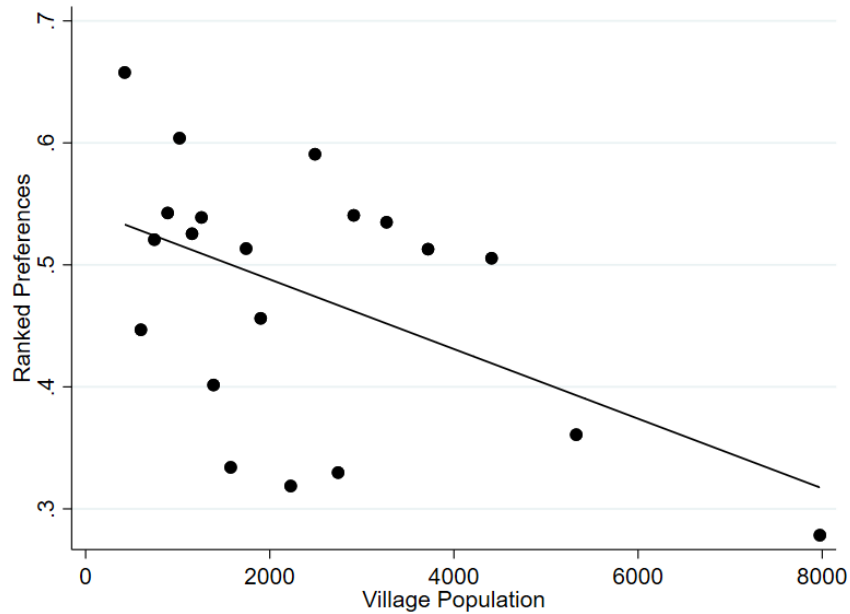
A common rationale for the creation of smaller polities is the creation of units with more homogeneous preferences. This section tests whether preferences are indeed more homogeneous in smaller

⁷⁵For example, economies of scale are certainly present for the management of the technologically advanced sewage system of a large, developed country city. In the rural Indian context, where sewage systems are covered or uncovered drains, fixed costs are low relative to variable costs.

jurisdictions in our context. We use survey data on individual preferences towards public goods from the Rural Economic & Demographic Survey (REDS), which includes 259 villages across India. We rely on a question that asked individuals to rank their preferences regarding 14 categories of public goods (e.g., drinking water, education). To measure homogeneity of preferences, we generate pairwise comparisons among all individuals within each village – a value of 1 was assigned if their rankings matched and 0 otherwise. We then calculate the village-level average across these pairs, where as this approaches 1, it implies that preferences are more homogeneous. Finally, we examine how these are related to the population of the village.⁷⁶

Figure A18 shows that there is indeed a negative correlation between the homogeneity of preferences and population size.⁷⁷ While this is a correlation, that is not a major issue here: if preferences are spatially correlated, as this analysis indicates, then splitting up populations into smaller groups will generate more homogeneous groups in terms of their preferences. However, we cannot causally establish that more homogeneous preferences are causally linked to public goods provision; demonstrating this would require data that we do not have here such as pre-delineation survey data.

Figure A18: Correlation between heterogeneity of preferences and population



⁷⁶The survey is representative at the level of the village rather than the GP, so we use village size to proxy more broadly for the relationship between population size and homogeneity of preferences. Since villages may have multiple hamlets similar to how a GP may have multiple villages, this seems like a reasonable assumption.

⁷⁷See online appendix section OA.4 for two additional methods for testing this relationship; results are similar.