Supplementary material

District-level analysis

In this section, we provide a district-level analysis of the relationship between NREGA participation and wage payment delay using nationally representative administrative data from the period 2014/15–2017/18 for 657 districts in India. We first discuss the empirical strategy, followed by a discussion of the sources of the data, and then present the results.

Empirical strategy

The generic form of the regression model that we estimate using the district-level data takes the following form:

$$\mathrm{WP}_{it} = \beta_0 + \beta_1 \mathrm{DELAY}_{it} + \sum_{k>1} \gamma_k X_{it} + \prod_{m>1} \alpha_m Z_i + \theta_j + \varepsilon_i \quad (1)$$

where WP is the proportion of households seeking NREGA work among total households registered for the scheme, and DELAY is the delay in NREGA wage payments. As discussed previously, the sign of the coefficient of DELAY (β_1) is indeterminate. If there is a 'discouraged worker' effect, then the coefficient would be negative—that is, a higher wage payment delay will lead to lower NREGA participation. On the other hand, if the wage payment delay leads to a lower present value of labour earnings, and a lower value of labour as an asset, the coefficient is expected to be positive. In this case, a higher wage payment delay will lead to higher NREGA participation.

We also control for the variables that are likely to influence NREGA participation at the district level. Our control variables include both time-variant and time-invariant variables. The time-variant control variables are represented by the vector X_i in Equation (1). The most important among them is rural poverty. Districts with higher levels of rural poverty would see a higher number of households seeking NREGA work. Further, not controlling for rural poverty could lead to omitted variable bias in our estimates as poorer districts may have less capable state administrations to implement the NREGA, leading to payment delays. As reliable district-level data on rural poverty are not available for India, we use district-level agricultural wages (AWAGE) for manual labour as a proxy for rural poverty as well as a measure of the outside option for agricultural labourers in the private rural labour market. In addition to agricultural wages, we also control for the impact of rainfall variation (RAINFALL) across districts on NREGA participation. NREGA participation is likely to be higher in rainfall-deficit districts of India, where agricultural activities are not profitable, leading to low private demand for agricultural labour.

Z_i represents the vector of time-invariant control variables constructed at the district level and captures the level of village infrastructure and social backwardness that would be correlated with district-level economic activity and local demand, as well as the ease with which poor agricultural households may be able to find outside work in the private labour market. These time-invariant variables are the proportion of villages in total inhabited villages on a bus route (TRANSPORT), villages with electricity (POWER), villages with a post and telegraph office (POST), villages with paved approach roads (ROAD), and villages with a primary school (SCHOOL). Social backwardness variables are the proportion of households which are Scheduled Castes (SC) and

the proportion of households which are Scheduled Tribes (ST). Scheduled Castes and Scheduled Tribes are among the poorest in India (Gang et al. 2009), and the Government of India specifically targeted these households for NREGA work (Breitkreuz et al. 2017; Government of India 2013; Vij 2013). Finally, we include year dummies (θ_j) to control for the impact of economy-wide macro shocks on NREGA participation.

Data

For the district-level analysis, we focus on the period 2014/15–2017/18, for which we have annual data. We have data for 657 districts in India. Data for the district-level analysis are drawn from different sources. Data on NREGA demand and wage payment delay come from the MGNREGA data portal of the Government of India. For NREGA demand, we use households' willingness to participate in NREGA work (which is equivalent to household notional labour supply for the NREGA) rather than the actual number of households receiving NREGA work. To be specific, our dependent variable is the ratio of households demanding NREGA work to the total number of job card holders and not the ratio of households receiving NREGA work to total number of job card holders. The actual number of households receiving NREGA work is determined by the short side of the market depending on the availability of NREGA work from local governments.[1] In any case, our results do not change if we use the number of households receiving actual NREGA work rather than total households demanding NREGA work. For each district, the MGNREGA data portal provides the number of transactions delayed and days of delay in the interval classes of payment delays of: 15–30 days, 30–45 days, 45–60 days, 60–90 days, and over 90 days. We compute the average expected delay in payment by taking the average of the midpoints of the interval classes, weighted by the proportion of transactions in each interval class in total NREGA transactions in the district.[2]

Data on agricultural wages (AGRWAGE) come from the report 'Agricultural Wages in India', published by the Directorate of Economics and Statistics, Ministry of Agriculture and Farmers Welfare, Government of India. The report presents the data on wages for different types of rural manual labourers—ploughman, reaper/harvester, sower, and weeder—by centre, of which there could be one or more in a district. The data in each centre is also separate for male and female labourers. We have averaged the wages across types of labour, gender and centres (in case of more than one centre in a district) to arrive at an average for each district. The data on RAINFALL are taken from the *Rainfall Statistics of India*, published by the India Meteorological Department of the Ministry of Earth Sciences, Government of India. This report prepares rainfall statistics for various administrative zones, including districts on a seasonal and annual basis. We compute average monthly rainfall for each district from this report and use it in the empirical analysis. Our time-invariant district-level controls—TRANSPORT, POWER, POST, ROAD, and SCHOOL—are obtained from the 2001 Census of India, while the social backwardness variables (SC and ST) are obtained from the 2011 Census of India. For data on the variable capturing rural bank density in the district (RBANK), which we use as an instrument for DELAY in IV estimations, we rely on the Reserve Bank of India (RBI) publication, *Basic Statistical Returns of Scheduled Commercial Banks in India*. These reports provide comprehensive data

[1] As Dutta et al. (2014) find using primary data from the state of Bihar, there is significant unmet demand for NREGA. Therefore, in principle NREGA is a demand-determined scheme, so there should be no systematic unmet demand for NREGA, However, in practice, due to several institutional rigidities and supply-side bottlenecks, there is widespread rationing of NREGA work, especially in states with weaker state capacity (Himanshu et al. 2015).

[2] We also use the value of transactions in each interval class instead of the number of transactions as the weights to calculate expected delay in payment, and get no difference in our results.

on distribution of branch offices, number of deposits, and amount deposited, as well as outstanding credit of scheduled commercial banks by location (rural/urban) for all districts of India. Data on district-wise rural population figures are drawn from the Census of India 2011. The construction of the variables used in the analysis is presented in Table 1.

Table 1: Variables and their construction

Variable	Definition	Type	Source
Dependent variable			
WP	Ratio of households demanding NREGA work to total job card holders	Ratio	NREGA portal
Independent variables			
DELAY	Average expected delay in wage payment weighted using number of transactions in each interval class of delay	Continuous	NREGA portal
AGRWAGE	Agricultural wages for manual labour	Continuous	Agricultural Wages in India
RAINFALL	Average monthly rainfall	Continuous	Rainfall Statistics of India
TRANSPORT	Proportion of villages in total inhabited villages on a bus route	Ratio	Census of India 2001
POWER	Proportion of villages with electricity in total inhabited villages	Ratio	Census of India 2001
POST	Proportion of villages with a post and telegraph office in total inhabited villages	Ratio	Census of India 2001
ROAD	Proportion of villages with paved approach road in total inhabited villages	Ratio	Census of India 2001
SCHOOL	Proportion of villages with a primary school in total inhabited villages	Ratio	Census of India 2001
SC	Proportion of Scheduled Caste households in total households	Ratio	Census of India 2011
ST	Proportion of Scheduled Tribe households in total households	Ratio	Census of India 2011
Instruments			
RBANK	Ratio of rural bank offices to total offices in a district	Ratio	Basic Statistical Returns of Scheduled Commercial Banks in India

Source: authors´ estimates.

Table 2 reports the summary statistics of the district-level variables included in our analysis. Our dependent variable is the ratio of households demanding NREGA work to total registered households. We find that fewer than half of the registered households (45 per cent) sought work under NREGA. With regard to our main variable of interest, *DELAY*, we find that about 28 per cent of the wage payments are delayed by 15 days or more. Table 2 also reports the summary statistics for the control variables and instruments used in our estimations. On average, Indian districts received a monthly rainfall of 98 mm. Our computations reveal that the average daily wage rate for a rural manual labourer stood at Rs. 236 per day. Scheduled Castes (SCs) and Scheduled Tribes (STs) constituted 28 per cent of the total population at the district level. Our average estimates at the district level also show that 86 per cent of the villages had a primary school, nearly half of the villages had better road connectivity, 85 per cent of the villages had electricity, 53 per cent had a post and telegraph office, and around 60 per cent had a paved approach road.

Table 2: Summary statistics of variables: district-level analysis

Variables	Number of observations	Mean	Standard deviation	Minimum	Maximum
WP	2,460	0.450	0.234	0.01	1
DELAY	2,460	28.334	23.100	0.0019	91
Control variables					
RAINFALL	2,460	98.594	70.180	1.733	658.225
AWAGE	798	235.875	93.474	88.875	765.84
SC	657	14.972	6.352	0.49	32.35
ST	657	13.162	18.428	0	93.76
SCHOOL	657	0.864	0.123	0.369	1
TRANSPORT	657	0.499	0.459	0.051	1
POWER	657	0.852	0.213	0.102	1
POST	657	0.536	0.280	0.104	1
ROAD	657	0.598	0.262	0.191	1
Instruments					
RBANK	657	0.059	0.033	0.0004	0.272

Source: authors' estimates based on field survey data.

Empirical results

We present the results based on the district-level data in Table 3. Five specifications of Equation (1) are estimated using OLS. In column 1, we report the OLS estimates of Equation (1) with just the DELAY variable. We introduce RAINFALL in column 2. Year effects are controlled for in column 3. In column 4 we also introduce AWAGE. We present the OLS estimates of Equation (1) with DELAY, the control variables (both time-variant and time-invariant) and year effects in column 5. In all regressions, standard errors are corrected for clustering at the district level.

Our district-level results also confirm the existence of a positive relationship between payment delay and worker participation. The coefficient on the expected delay in wage payment is positive and significant at the 1 per cent level in all specifications. Among the control variables, the measures of social backwardness of the district—the proportion of SC households in total households (SC) and the proportion of ST households in total households (ST)—are significant at the 1 per cent level and have the right sign, suggesting that households in socially backward regions are more likely to demand NREGA work. On the other hand, we obtain positive and significant coefficients for RAINFALL, POWER, and POST, indicating that participation is higher among households who live in largely rain-fed districts and districts with better village infrastructure. These results are contrary to our expectations, as one would expect higher demand for NREGA work from rainfall-deficit districts and districts with low levels of village infrastructure. These findings possibly reinforce our earlier claim of programme-capture by elite groups.

One limitation of the agricultural wage data that we use in our regressions as an important control variable is that they are only available for three years and for 266 of the 615 districts for which we have NREGA demand and wage payment delay data. To see whether our results change if we use the entire sample for which we have data, we estimate Equation (1) for all 498

districts without the agricultural wage variable. We do not find any difference in our key finding that higher delay of wage payments leads to greater demand for NREGA work.[3]

Table 3: Payment delay and worker participation: district-level results

Variables	(1)	(2)	(3)	(4)	(5)
DELAY	0.0013***	0.0011***	0.002***	0.001***	0.002***
	(0.0002)	(0.0002)	(0.0004)	(0.0004)	(0.0005)
Control variables					
RAINFALL		0.0008***	0.001***	0.0003	0.0005***
		(0.0001)	(0.0001)	(0.0002)	(0.0002)
AWAGE				0.026	0.040
				(0.026)	(0.035)
SC					0.012***
					(0.002)
ST					0.004***
					(0.0008)
SCHOOL					0.019
					(0.111)
TRANSPORT					0.0003
					(0.015)
POWER					0.142**
					(0.058)
POST					0.173**
					(0.075)
ROAD					0.042
					(0.083)
Constant	0.414***	0.341***	0.238***	0.180	–0.487***
	(0.008)	(0.009)	(0.017)	(0.136)	(0.188)
Year effect?	No	No	Yes	Yes	Yes
F	28.41	74.63	146.47	71.46	43.35
N	2460	2460	2460	798	657
Regression	OLS	OLS	OLS	OLS	OLS
R^2	0.016	0.07	0.124	0.059	0.278

Notes: (a) our dependent variable in all estimations is the proportion of households demanding NREGA work in total registered households. (b) District is the unit of analysis. (c) Our dataset corresponds to the four-year period, 2014–15 to 2017–18. (d) Control variables: RAINFALL: average monthly rainfall; AGRWAGE: annual average agricultural wage for respective years; SC: proportion of SC households in total households; ST: proportion of households who are STs; SCHOOL: proportion of villages with a primary school in total inhabited villages; TRANSPORT: proportion of villages with a bus connection in total inhabited villages; POWER: proportion of villages with electricity in total inhabited villages; POST: proportion of villages with post and telegraph offices in total inhabited villages; ROAD: proportion of villages with paved roads in total inhabited villages. (e) ***, ** and * indicates significance at 1, 5, and 10 per cent levels respectively. (f) Figures in parentheses are standard errors, corrected for clustering at the district level.

Source: authors' estimates.

A key limitation of the district-level analysis is that we are unable to address reverse causality from WP to the wage payment delay, in contrast to the IV strategy we used in the household-level analysis. This is because we are unable to identify credible instruments for the wage payment delay variable that vary both across districts and over time. Further, the lack of time variation in our key variable of interest—wage payment delay—does not allow us to use district fixed effects to control for unobserved district-level factors that may affect both wage payment delay and household labour supply. For these reasons, we interpret the positive relationship

[3] Due to lack of space, we do not present these results here. However, they are available from the authors upon request.

between wage payment delay and the proportion of households demanding NREGA work as correlational, rather than causal.

References

Breitkreuz, R., C. Stanton, N. Brady, J. Pattison-Williams, E.D. King, C. Mishra, and B. Swallow (2017). 'The Mahatma Gandhi National Rural Employment Guarantee Scheme: A Policy Solution to Rural Poverty in India?'. *Development Policy Review*, 35(3): 397–417.

Dutta, P., R. Murgai, M. Ravallion, and D. van de Walle (2014). *Right-to-Work? Assessing India's Employment Guarantee Scheme in Bihar*. Washington, DC: World Bank.

Gang, I., K. Sen, and M. Yun (2009). 'Poverty in Rural India: Caste and Tribe'. *Review of Income and Wealth*, 54(1): 50–70.

Government of India (2013). *The Mahatma Gandhi National Rural Employment Guarantee Scheme: Report to the People*. New Delhi: Ministry of Rural Development.

Himanshu, A. Mukhopadhyay, and M.R. Sharan (2015). 'NREGS in Rajasthan: Rationed Funds and Their Allocation across Villages'. *Economic and Political Weekly*, 50(6): 52–60.

Vij, N. (2013). 'Empowering the Marginalized: Mahatma Gandhi National Rural Employment Guarantee Act in India'. *Human Affairs*, 23: 91–104.