Land Inequality and the Provision of Public Works

—Evidence from National Rural Employment Guarantee Scheme in India

Yanan Li *

December 29, 2018

Abstract

Does existing inequality hinder redistributive policies that aim to help the poor? This paper answers this question under a widely used redistributive policy in developing countries—public works schemes. Using district-level data on land ownership distributions and the implementations of the National Rural Employment Guarantee Scheme in India, I find robust evidence that the concentration of land ownership reduces public works provision. This relationship could be explained by the mechanism that public works schemes raise agricultural wages in the private labor market, thereby incentivizing big landlords to use their political power to oppose this program. To address the potential endogeneity due to unobservables and measurement error, I leverage a historical institution in India, the land revenue collection system established by British colonial rulers during 1750-1861, to construct an instrumental variable for land inequality. Due to the concentration of post-independence land reforms enacted in landlord-dominated areas, those areas have lower land inequality today than the previously non-landlord dominated areas. The IV estimates suggest that a 1 percent increase of land Gini coefficient would lead to a 3-5 percent decrease in public job provision. The results are robust to using the alternative measurements of land inequality and public works implementation. To exclude the possibility that the higher provision of public jobs in more equal areas is driven by a higher demand for public jobs, I show that more equal areas have higher agricultural wages in the private labor sector. This paper provides the first empirical evidence that the concentration of land ownership, a proxy for political power, is a hurdle to providing public employment to the poor, suggesting power asymmetries could hinder policies aimed at promoting equity.

Key Words: Inequality; Redistribution; Land Ownership; Land Reform; Public Works; National Rural Employment Guarantee Scheme

JEL Classification: D04, I38, J43, Q15

*Economics Department at Beijing Normal University, Beijing, China. Email: yananli@bnu.edu.cn. I am very grateful to Christ Barrett, Christopher Boone, Leah Bevis, Nancy Chau, Ritam Chaurey, John Hoddinott, Lawrence Kahn, Ravi Kanbur, Travis Lybbert, Victoria Prowse, Tanvi Rao, Naveen Sunder and Laura Zimmermann for their helpful comments. All errors are my own.

1 Introduction

Recent debate on inequality has focused on the trends of inequality and its influencing factors, but relatively silent on the consequences of inequality. The current paper fills in this literature by examining the potential effects of inequality on redistribution policies. Does inequality lead to less redistributive efforts to the poor? If the answer is yes, then inequality would lead to a vicious cycle – higher inequality, less redistribution efforts to the poor, then even higher inequality. Therefore, it is an important question to answer. Previous theoretical literature is still inconclusive on this question, with the earlier literature suggesting a positive association (e.g. Alesina and Rodrik, 1994; Persson and Tabellini, 1994) and the more recent suggesting a negative association (e.g. Benabou, 2000; Galor et al., 2009). The empirical evidence is relatively lacking in identifying the direction of the effect and the mechanisms through which inequality might affect redistributive policy, with a few exceptions (e.g. Boustan et al., 2013; Cinnirella and Hornung, 2016; Ramcharan, 2010).

This paper focuses on landownership inequality to study the consequence of inequality. The inequality of land ownership is an important form of inequality, as land is the main production factor before the industrial economy and still so today in many developing countries. Furthermore, the distribution of land is directly linked to the concentration of political power. This power gravitates towards landlords, who may either influence tenants' votes or directly influence the politicians in the direction beneficial to themselves. The literature has provided evidence that large landlord elites influence the political process to prevent economic reforms or redistributive policies, such as educational expenditure (Cinnirella and Hornung, 2016; Ramcharan, 2010), human-capital accumulation (Galor et al., 2009), general social assistance programs (Anderson et al., 2015) and public goods (Beg, 2016).

In line with these recent studies on inequality and redistribution, this paper answers the question under another widely used redistributive policy—public work schemes, which, due to its complexity in design and implementation, warrant special attention. A public works program is the provision of employment at a prescribed wage for those unable to find alternative employment by the creation of public infrastructure projects, such as transport infrastructure (e.g. roads, railroads and canals) and public services (e.g. sewage and dams). It is financed by the government and functions as a form of social safety net in many developing countries, such as India, Philippines, Bangladesh and Chile (Subbarao, 1997). The provision of public jobs raises agricultural wages in the private labor market (Imbert and Papp, 2015; Merfeld, 2019; Muralidharan et al., 2017), thereby incentivizing big landlords to use their political power to oppose this program. India is the perfect context studying the relation between land inequality and the provision of public works, because it has the world's largest public works program—the National Rural Employment Guarantee Scheme (thereafter, NREGA) and faced with a historical tension arising from land inequality.

Empirically, I compare district (within-state) variations of land ownership inequality and public works provision, using census data on district-level land distribution in 2005 and the implementation data of the NREGA program since its inauguration in 2006. Land inequality is measured by the Gini coefficient. The provision level of public employment is measured by four dimensions: the fraction of rural households provided with employment, the per capita labor expenditure, average days of employment provided per person in either Schedule Caste or Schedule Tribe (thereafter, SC/ST) and the total number of completed works per rural person. OLS estimates suggest that a 1 percent difference in land Gini coefficient leads to a 0.5-1.2 percent gap in NREGA provision. Using the method of selections on observables (Altonji et al., 2005), I show that the results are less likely to be fully driven by unobservables.

To further address the potential endogeneity issue arising from measurement errors and omitted variables in the OLS estimation, I use a historical institution as the instrumental variable for land inequality—the land revenue collection system established by British colonial rulers during 1750-1861. This variable derives from the study by Banerjee and Iyer (2005). Despite a higher Gini coefficient of land ownership inequality in landlord-dominated areas during 1885-1948, such areas experienced more frequent land reforms after Indian independence. Therefore, the first-stage conditional correlations suggest that landlord-dominated districts have significantly lower Gini landownership inequality in 2005. Under the assumption that the instrument is exogenous, the IV estimates confirm the negative effect of landownership inequality on public works schemes. 2SLS estimates suggest that a 1 percent difference of land Gini coefficient leads to a 3-5 percent gap in NREGA provision. Using the method proposed by Conley et al. (2012), I show that the negative effect still holds when relaxing the exclusion restriction of the instrumental variable by allowing a negative association between the instrumental variable and NREGA provision.

Both OLS and IV results are robust when using the alternative measurement of land inequality—the share of land owned by the top 10 percent largest farmers, which more directly captures the top distribution and hence large farmers' political power. I finally exclude the possibility that the higher provision of public jobs in more equal areas is due to a higher demand for public jobs, by showing that these more equal areas have higher agricultural wages in private sector.

This study adds to the understandings of the heterogeneity of the implementation of NREGA across different districts. NREGA claims to provide 100 days of working opportunity to each rural household in need of jobs. As a matter of fact, there is an un-met demand for jobs in almost all districts and the extent of un-met demand differs by districts. Existing literature has been trying to explain this heterogeneity of NREGA implementation mostly in terms of political incentives and administrative capacity (Gulzar and Pasquale, 2016; Niehaus and Sukhtankar, 2013; Nath, 2015; Gupta and Mukhopadhyay, 2016; Sheahan et al., 2016), and of the political reservation system (Dunning and Nilekani, 2013; Bose and Das, 2015). To the best of my knowledge, this paper is the first study to link district-level heterogeneity in the provision of NREGA jobs to the inequality of landownership distribution. Districts with more concentrated land distributions are expected to see a lower provision of NREGA employment, because in those districts big farmers have stronger economic incentives and political power to block wage-increasing public works schemes. Indeed, there is abundant anecdotal evidence showing that big farmers lobby to suspend the provision of NREGA employment (Maiorano, 2014), but broad-based quantitative testing of this notion has not been attempted previously. [1]

Investigating the question of land inequality and public works provision adds to the understanding of Indian land inequality which, as a legacy of British colonial institutions, has been a historically important and intricate issue. The relation between landlords and the landless affects different aspects of rural life and shapes the effectiveness of public policies. There has been a large number of land reforms since Indian independence, but most of them are through legislated ceilings on landholding (rather than direct land redistribution) and such reforms have been rarely implemented with any degree of seriousness (Besley and Burgess, 2000). As a result, after all those land reforms, the share of land occupied by the top 10 percent biggest farmers is still as large as 46 percent. This paper, by showing that the concentration of landownership hence political power is a hurdle to redistributive efforts and successful anti-poverty policies, offers a potential justification for further efforts at land reforms. Moreover, compared to the previous estimates of the effect of inequality on redistribution that were derived using soil or other geographical information as instrumental variables, the IV estimates in the current paper are particularly policy relevant because the lower levels of land inequality seem to be driven by land reforms (rather than natural conditions).

This paper speaks to the general discussion on inequality and public expenditures. The literature

[1] In studying clientilism between landlords and the landless in Indian villages, Anderson et al. (2015) show land-owning elites will prefer weak provision of centrally funded pro-poor prgrams such as Employment Guarantee Program. The current paper differs from their paper in at least three respects. First, their survey data is restricted to 3 regions in the state of Maharashtra, while the current study uses district-wise nationally representative data. Second, they proxy landlords' political power by the proportions of land in the village dominated by the upper caste, Maratha. I use the concentration of land ownership, which goes beyond the constraint of caste backgrounds and have more general implications. Third, one of the pro-poor policies in their paper, EGS, is a previous form of NREGA. It is believed that NREGA has incorporated the lessons and successes of EGS, with broader goals and better implementations.

finds a detrimental effect of early inequality on the emergence of human-capital accumulating and growth-promoting institutions (e.g. Persson and Tabellini, 1994; Sokoloff and Engerman, 2000; Galor et al., 2009). The main mechanism is that land concentration induces landowners to use political power to assure lower public expenditure in education, for fear that higher public education investment would raise up labor cost or generate migration from agricultural sector to industrial sector. This mechanism also applies in the context of public works schemes. Providing public employment to the landless and the marginal farmers will increase labor wages, and this wage effect will incentivize landlord elites to oppose the implementation of the public works schemes (Anderson et al., 2015; Maiorano, 2014).

This paper also broadly speaks to the literature on inequality, redistribution and economic growth. This literature initially argues that inequality is conducive to the adoption of growth-retarding redistributive policies (Alesina and Rodrik, 1994; Persson and Tabellini, 1994). This positive relationship is supported by some existing literature (Boustan et al., 2013). However, the current paper, coupled with other recent empirical evidence (e.g Galor et al., 2009; Ramcharan, 2010), casts doubt on this underlying mechanism. Instead, the recent evidence suggests that inequality is a hurdle for redistribution, provided that the landlords, or better-endowed agents, have sufficient political power to influence redistribution policies.

The remainder of this paper is organized as follows. In section 2, I discuss the background information of the NREGA, highlighting the necessary facts that make it possible for landlords to play a role in the provision of NREGA jobs. Section 3 discusses the mechanism of how land inequality affects public works provision. Section 4 discusses data issues. Section 5 presents empirical strategies and principal findings, followed by robustness checks. Section 7 concludes.

2 Background: The National Rural Employment Guarantee Act

2.1 Demand-Driven Nature of NREGA Employment

The Mahatma Gandhi National Rural Employment Guarantee Act of 2005 created the "right to work" for all households in rural India through the National Rural Employment Guarantee Scheme. It was a three-phased nation-wide rollout, with 199 districts in Phase 1 (Feb 2006), 128 districts in Phase 2 (April 2007) and the remaining 261 districts in Phase 3 (April 2008). By 2008, it reached all districts in India. It is the largest public works program in the world so far and asserts guaranteeing 100 days of working opportunity for each household per financial year (June in the current year to May in next year). Households need to obtain job cards from the local governments, which are used to record work done and payment. According to the Act, as long as an eligible household files applications for jobs, the local government must provide employment within 15 days and within 5 kilometers of the applicant's home. Otherwise, states are liable to pay unemployment allowances. In practice, however, there are still frictions in the implementation leading to some unmet demand, such that those wanting work do not get it in a timely manner.

More than half of the works are related to water conservation, with other types of works including irrigation provision, land development and rural connectivity. Wages are to be paid at the statutory minimum wage rates, which makes this program a means of enforcing minimum wage laws. Unlike in the private labor market where women earn a much lower wage rate than men, wage rates in NREGA are job specific rather than gender specific. Therefore, NREGA jobs are especially appealing to women. In addition, as a social insurance tool, NREGA has stronger demand under adverse agricultural conditions. Santangelo (2016) finds workers resort to NREGA to a larger extent when the local economy is hit by worse agricultural productivity shocks.

2.2 Financing NREGA and the Supply Constraint

The National Rural Employment Guarantee Act incentivises states to provide employment by stating that 100 percent of the unskilled labor cost and 75 percent of the material cost of the program is borne by the central government. The labor to material ratio could vary from 90:10 to 60:40.

The overall annual labor spending on NREGA at state/district/ block/ village level is a pre-determined cap. Labor budget for each financial year is determined in the previous year, following a "bottom-up" process from the village level to the state level and last to the central government (NREGA Operational Guidelines, 2013). This budget plan includes (i) the anticipated quantity of demand for jobs in the next year (ii) the precise timing of the demand for work and (iii) a shelf of projects to be prepared and prioritized to meet job demand. Table 1 presents the various steps involved in the preparation and finalization of annual labor budgets. Because labor budget is an estimation and NREGA is a demand driven program, the Act states that the States may, based on actual performance, any time during the year, come back to the Ministry requesting revision of their existing labor budget, following the procedures in Table 1. However, in fact, the flexibility is limited. Once the labor budget is finalized, the maximum supply of jobs in each state/district/block will not be changed for the next financial year.

Therefore, there will be a shortage of supply for NREGA jobs if any of the following cases occurs—(i) an exogenously fixed maximum level of spending on NREGA by the center government; (ii) an

Table 1: Timelines for various steps involved in preparation and finalization of annual labor budget.

Data	Action to be taken
August 15	Gram Sabha to approve GP Annual Plan and submit to PO
September 15	PO submits consolidated GP Plans to Block Panchayat
October 2	Block Panchayat to approve the Block Annual Plan and submit to DPC
November 15	DPC to present District Annual Plan and LB to District Panchayat
December 1	District Panchayat to approve District Annual Plan
December 15	DPC to ensure that shelf of projects for each GP is ready
December 31	Labour Budget is submitted to Central Govt.
January	Ministry scrutinizes the Labour Budget and requests for compliance for deficiencies, if any
February	Meetings of Empowered Committee are held and LB finalized
February, March	"Agreed to" LB communicated to States. States feed data of Month-wise and District-wise breakup of "Agreed to" LB in MIS and communicate the same to Districts/blocks GPS
Before 7th April	States to communicate OB, Center to release upfront/ 1st Tranche.

Source: Mahatma Gandhi National Rural Employment Guarantee Act 2005 – Operational Guidelines, 4th version. Chapter 6.10.

underestimation of job demand in the budget planning; (iii) a poor timing of job demand; and other cases. The actual implementation is further complicated by states' constraints in organizing projects and workers. Even if the budget planning is not an issue, accommodating supply to demand could still be a challenge because of the incapability to meet the relatively skilled labor requirements at the local level, such as panchayat technical assistants (Dutta et al., 2014). As a result, although the NREGA program is designed to be a demand-driven program, there is an un-met demand for jobs in almost all states (Dutta et al., 2014). On average, each household works roughly 35 person-days per financial year, far less than the claimed 100 days. The extent of the un-met demand differs by districts and by time.

2.3 Landlords and NREGA Employment

Landlords are an important interest group in the implementation of NREGA. Providing public jobs to the landless and marginal farmers will increase labor wages (Imbert and Papp, 2015; Merfeld, 2019; Muralidharan et al., 2017), which will potentially increase production costs for landlords who hire casual labors. Thus this wage effect brings landlords an economic incentive to oppose the implementation of NREGA (Anderson et al., 2015; Maiorano, 2014). There are at least two stages where big landlords can intervene the process of providing NREGA jobs.

First, at the stage of making the labor budget, landlords may lobby for a budget plan that does not provide enough jobs to the rural poor. As Table 1 shows, budget planning is a bottom-top decision making process. The demand for NREGA jobs and the shelf of projects are first identified at the

Gram Panchayat level, then the demand and supply are consolidated at the block level, and further aggregated at the district and state levels. The fact that lower level governments such as block and village have a substantial discretion in this process renders big landlords' influences very likely. It is after all easier for landlords to lobby village governments than state governments.

Second, even after labor budget is made, big farmers can still use their political power to block the implementation, such as delaying work assignment, payment and some complementary machinery (see Maiorano (2014) for anecdotal evidence of lobbying). As a result, as NREGA annual report shows, the final work completed is smaller than the original budget.

3 Mechanism

The political mechanism of inequality and redistribution has been established by the literature. Higher inequality lowers the level of awareness of the poor, decreasing the level of their political participation (e.g. Bardhan and Mookherjee, 2005; Ramcharan, 2010). Meanwhile, greater inequality can concentrate the benefits of political participation and simplify the collective action problem among the landed, which leads to a higher and more effective political participation among the landed elites. In the cases that the landlord elites are a net loser from redistribution, they would block redistribution. Therefore, a higher land inequality predicts lower redistributions to the poor.

As the primary interest of this paper lies in economic effects rather than political effects, I will impose a crude political mechanism under which landlords have sufficient political power against redistributive policies. Instead, I will focus on the economic incentives that lead big landlords to oppose the provision of public employment.

Providing public jobs to the poor introduces a competition for labor between the public works schemes and the rural private employers. The literature has found that the introduction of NREGA increases rural casual labor wages by at least 6 percent (Imbert and Papp, 2015; Merfeld, 2019; Muralidharan et al., 2017). This wage effect could potentially reduce landlords' profit, if they keep hiring casual labor. Moreover, the loss of profits from wage increases is greater for bigger farms. This idea is formulated in the following simple framework.

Following Galor et al. (2009), I assume landowners are a fraction $\lambda \in (0, 1)$ of all individuals in society who equally share the entire stock of land in the economy, X. Thus each landlord owns X/λ units of land. Assume agricultural production only needs two inputs, land X/λ and labor L_i. Assume a Cobb-Douglas production function, $F(X/\lambda, L_i) = (\frac{X}{\lambda})^{(1-\alpha)} L_i^{\alpha}$, where $0 < \alpha < 1$. Assume product price is normalized to be 1, and per unit labor cost is w. Then the profit for farm i is as follows,

$$\pi_i(L_i) = (\frac{X}{\lambda})^{(1-\alpha)} L_i^{\alpha} - wL_i.$$

By taking the first order derivative of the profit function $\pi_i(L_i)$ with respect to L_i, and setting it to zero, I derive the firm's optimal labor input choice is

$$L_i(w) = (\frac{X}{\lambda})(\frac{1-\alpha}{w})^{(1-\alpha)}.$$

Now plug $L_i(w)$ into the profit function. The optimal profit is a function of w,

$$\pi_i^*(w) = (\frac{X}{\lambda}) w^{(1-1/\alpha)} B, \tag{1}$$

where $B = [(1-\alpha)^{1/\alpha - 1} - (1-\alpha)^{1/\alpha}]$. As $0 < \alpha < 1, B > 0$.

Take the first order derivative of $\pi_i^*(w)$ with respect to w,

$$\frac{d\pi_i^*(w)}{dw} = (\frac{X}{\lambda})(1 - \frac{1}{\alpha}) w^{-1/\alpha} B < 0.$$

The negative sign of the derivative suggests that farm profits decrease with an increase of labor wages. As NREGA has been documented to lead to at least a 6-percent increase of private sector wages (Imbert and Papp, 2015; Merfeld, 2019; Muralidharan et al., 2017), this simple framework confirms the profit losses for farms. In addition,

$$\frac{\partial}{\partial \lambda}\left(\frac{\partial \pi_i^*(w)}{\partial w}\right) = -(\frac{X}{\lambda^2})(1 - \frac{1}{\alpha}) w^{-1/\alpha} B > 0. \tag{2}$$

A positive second-order derivative means that marginal profit loss from a wage incrase, $-\frac{d\pi_i^*(w)}{dw}$, decreases with λ. In other words, with a higher landownership concentration (a smaller λ) and hence a larger size per farm, the marginal profit loss from wage increases is even bigger. Therefore, the wage-increasing nature of public works schemes provides big landlords the economic incentives to oppose the program.

4 Data

4.1 Land Inequality

District-wise data on land distribution in 2005 come from Indian Agricultural Census (excluding Maharashtra), which is conducted at five yearly intervals. Although the information is collected on operational land holdings rather than owned land holdings, the wholly owned and self-operated holdings account for 97.14 percent (Page 29, Agriculture Census Report, 2005). Therefore, I use this dataset on operational land holdings to approximate the distributions of landownership in India.[2]

This dataset has information on the number and area of operational holdings across the following size bins (in 1000 hectares): below 0.5; 0.5-1; 1-2; 2-3; 3-4; 4-5; 5-7.5; 7.5-10; 10-20; 20 & above.[3] I use the average size of land holdings in each bin to construct landownership Gini coefficient. Overall, the average Gini coefficient in our sample districts is 0.47 (see Table 3). The largest 10 percent of operational holders operate about 46 percent of total land in India, as shown by Figure 1. Moreover, the state-wise average Gini coefficients have large variations (see Figure 2), suggesting different extents of landownership concentration across states.

4.2 NREGA Implementation

NREGA implementation data come from public data portal[4]. Table 2 presents summary statistics of NREGA implementation by financial year (starting from April in the current year and ends in March the next year) using alternative measurements. The first row tells that, among all working population in India, 12% of them worked for at least one day in public works in 2006, the first year that NREGA was introduced. This number increased to 18% in 2007, and 30% in 2010.

Labor expenses are deflated by state-level consumer price index, using 2006 as the base year. The average wages per rural person received (regardless of their work status in NREGA) increased from 68 Rupees in 2006 to 167 Rupees in 2010. When focusing only on the subpopulation that were provided with public employment, the average wages that each household received increased from 2667 Rupees in 2006 to 2867 Rupees in 2010.

[2] According to Agriculture Census in India, "an Operational holder is the person who has the responsibility for the operation of the agricultural holding and who exercises the technical initiative and is responsible for its operation." An operational unit could include multiple plots. The operated areas comprise of i) Land owned and self operated; ii) Land leased in; iii) Land otherwise operated.

[3] I use the information on "Sub-total" land holdings, including both individual holding and joint holdings, to measure district level land distribution. The ratio of joint holdings to individual holdings is, 1:6.5 in terms of numbers and 1:5 in terms of areas (Agriculture census report 2005, page 121). Land operated by institutions constituted less than 0.5% of the total area, hence excluded from the data.

[4] *http://nregarep2.nic.in/netnrega/dynamic2/dynamicreport_new4.aspx*

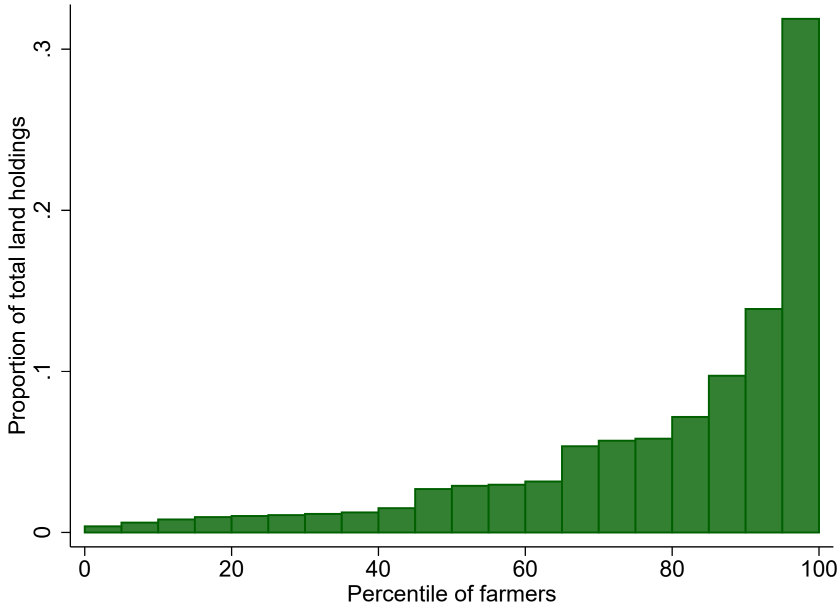

Figure 1: Shares of land area by percentiles of holdings, 2005

Note: Size-classes are as follows: below 0.5; 0.5-1; 1-2; 2-3; 3-4; 4-5; 5-7.5; 7.5-10; 10-20; 20 & above. The graph is derived by first ranking all land holdings by class size in India, then calculating the share of land operated at each five percentile.

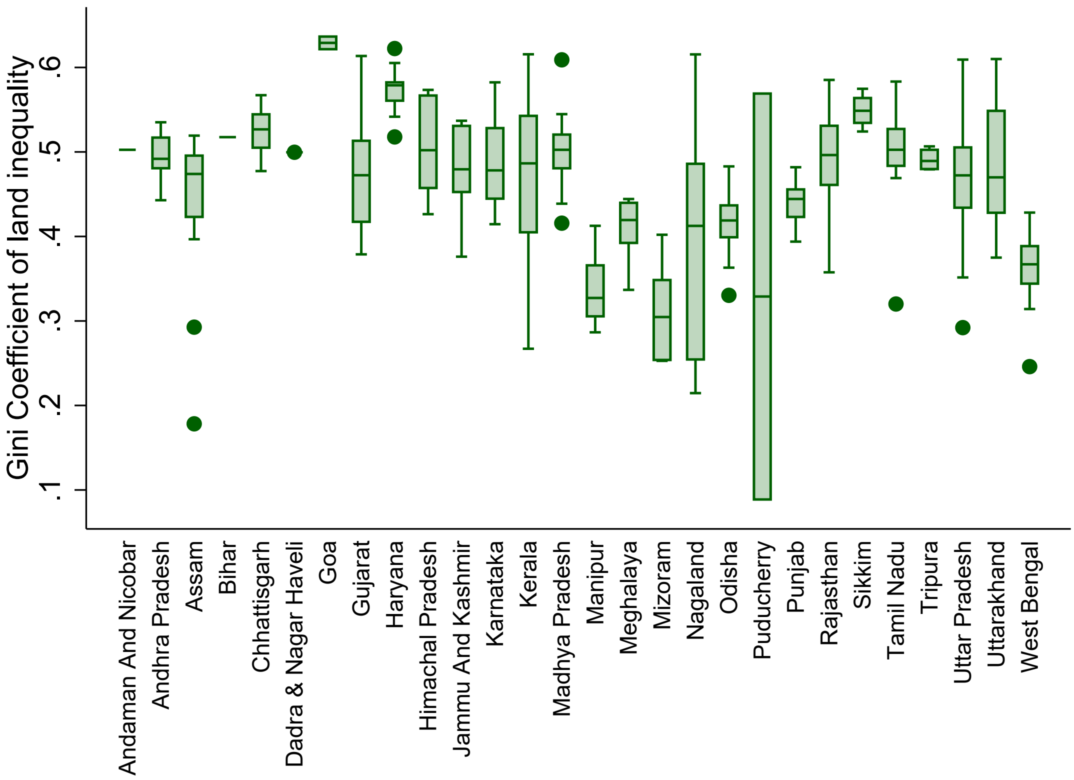

Figure 2: Land inequality (Gini coefficient) by state, 2005

Source: The author calculated Gini coefficient based on district-wise land distribution data from 2005 Indian Agricultural Census. Only states in the OLS regression sample are included.

Table 2: Summary statistics of NREGA implementation

	(1) 2006 mean/sd	(2) 2007 mean/sd	(3) 2008 mean/sd	(4) 2009 mean/sd	(5) 2010 mean/sd
% of households provided employment	12.40	17.69	27.32	29.63	31.09
	(23.01)	(25.16)	(26.11)	(24.22)	(23.81)
avg days of employment provided per rural SCST person	2.01	2.95	5.24	6.57	6.45
	(4.55)	(5.07)	(6.60)	(6.93)	(7.12)
avg days of employment provided per rural woman	1.25	1.79	3.48	4.32	4.11
	(3.66)	(3.97)	(5.50)	(5.90)	(5.45)
avg days of employment provided per NREGA-woman	19.00	16.35	17.00	21.61	20.84
	(16.27)	(17.03)	(16.80)	(15.88)	(13.09)
labor expense per rural person (2006 Rs.)	68.64	92.41	163.12	175.46	167.70
	(192.87)	(179.99)	(258.53)	(224.40)	(202.98)
number of completed works per 1000 rural persons	1.73	5.06	9.48	13.39	7.68
	(5.06)	(17.57)	(26.09)	(20.69)	(13.50)
# of districts with employemnt provided	122	202	409	410	415
Observations	416	416	416	416	416

Notes: Original data come from MGNREGA public portal. Only districts in the regression analysis are included. Labor expense is deflated by state-wise Consumer Price Index, using 2006 as the base year.

Information on the three-phased roll-out comes from the document by NREGA Report (2007).[5] Phase 1 includes 200 districts; Phase 2 includes 130 districts and Phase 3 includes the rest of districts. Phases are determined based on the ranking of Backwardness Index (Zimmermann, 2012). I extract this index and its five components from Indian Planning Commission 2003 Report, including agricultural wages in 1996, agricultural productivity per person in 1990-93, agricultural productivity per hectare in 1990-93, the population ratio of Schedule Caste to Schedule Tribe from the 1991 Population Census and the poverty ratio in 1994 (Commission et al., 2003).

Despite an increasing provision of NREGA jobs over time during the study period, there is substantial heterogeneity of NREGA implementations across districts. As shown in Figure 3, the average job provision by NREGA varies a lot by states. Consistent with the empirical findings, Figure 4 visually presents a negative relation between land inequality and public works provision by a kernel regression of the shares of households provided with public jobs on the share of land occupied by the top 10 percent biggest farmers.

4.3 Demographic and Geographic Information

District profiles are downloaded from the 2001 population census, including caste composition, employment and industry structure, literacy rate, amenities and infrastructural facilities, district area size and so on. Population between 2001 and 2010 are filled using these two years' census data, assuming

[5]This online document nicely presents the phase-in progress *http://nrega.nic.in/MNREGA_Dist.pdf*

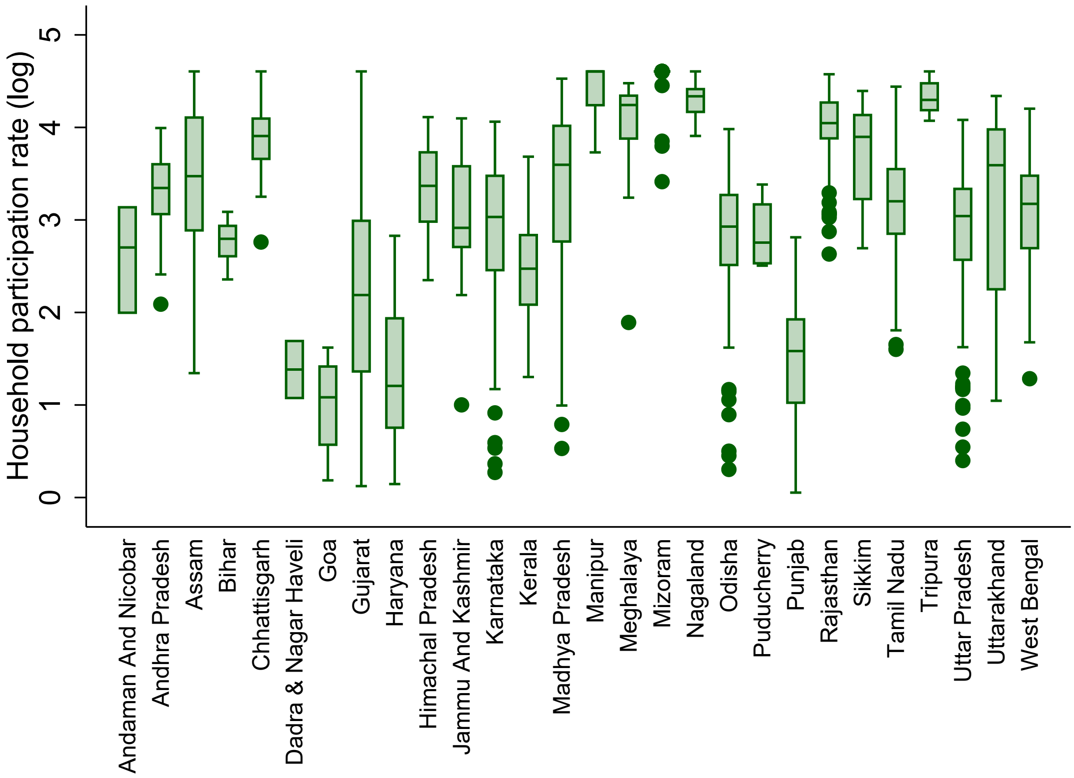

Figure 3: Shares of Households provided with NREGA employment by state, 2006-2010

Note: Shares are calculated as total number of households provided with NREGA employment divided by total rural households in the district. All districts in the OLS regression sample are included.

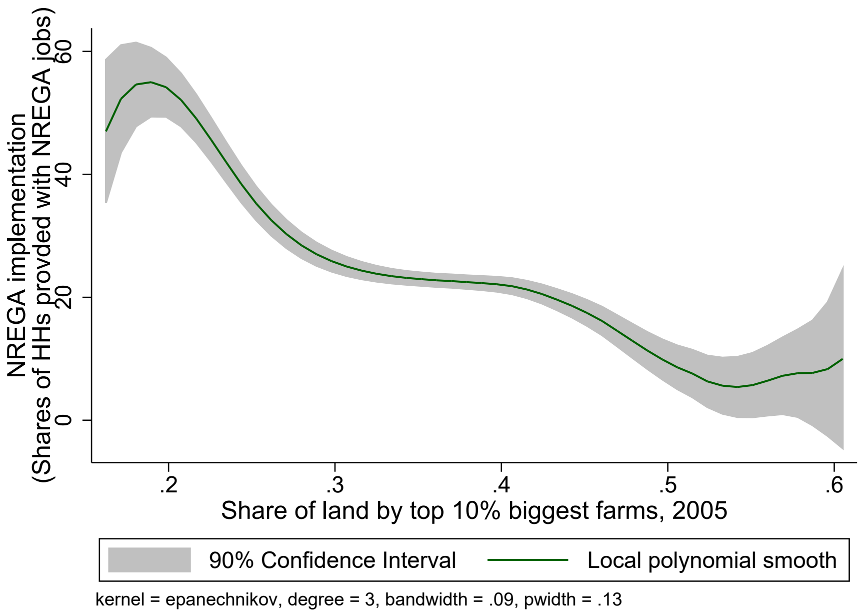

Figure 4: Local polynomial smoothing

Note: Author calculates the share of land by top 10% biggest farms based on 2005 India Agricultural Census. NREGA implementation is measured by the share of households provided with public jobs. Kernel = epanechnikov, degree=3, bandwidth=.09.

a growth rate equal to that during 1991-2001. Table 3 presents summary statistics of district-wise demographic information in 2005.

The monthly rainfall data are obtained from Center for Climatic Research, University of Delaware. Indian agricultural year is split into two distinct seasons– wet season (from June to November) and dry season (from December to May). Existing studies document that NREGA participation is strongly associated with rainfall shocks in wet season. Therefore, I compute wet season precipitation by aggregating the amount of precipitation between June and November in the study year. Soil information is obtained from Food and Agriculture Organization (FAO) Digital Soil Map of the World and Derived Soil Properties (CDROM). Table 3 shows that 91% of the land contains medium or fine level soil; 9% of land is covered by course soil.

Compiling these data sets into a district-wise panel is complicated by district jurisdictional changes during 2001-2011. There were 640 districts in 2011, as opposed to 593 districts in 2001 (Census, 2011). In the analysis, districts with boundary changes are excluded, although results are robust to adding these districts back. The final sample includes 416 districts at the 2001 district level.

Table 3: Summary statistics of land inequality and demographic information

	mean	sd	min	max
Gini coef.	0.47	(0.08)	0	1
Rural area (Sq. km)	5044.34	(4862.42)	119	45382
% of Rural pop	77.94	(15.42)	12	100
Literacy rate	65.29	(11.70)	31	97
wet season rainfall (100 mm)	1.03	(0.69)	0	6
% of land covered in fine soil	20.08	(24.19)	0	97
% of land covered in medium soil	70.70	(27.68)	0	100
% of SC population	15.84	(8.84)	0	50
% of ST population	15.89	(26.18)	0	98
Work-population ratio	40.91	(6.98)	24	63
% of Main workers	30.68	(5.95)	17	52
% of Marginal workers	10.23	(4.20)	2	24
% of Agricultural labourers	22.63	(12.98)	1	63
% of Cultivators	37.77	(18.04)	1	82
% of Household industry workers	4.05	(3.89)	1	31
% of Other industries	35.60	(17.71)	8	91
% villages with Safe Drinking water	96.19	(10.49)	24	100
% villages with Electricity (Power Supply)	84.76	(18.89)	10	100
% villages with Paved approach road	60.87	(25.51)	12	100
% villages with Primary school	84.35	(14.15)	31	100
% villages with Medical facility	41.82	(25.66)	3	100
% villages with Post and telephone facility	52.43	(26.81)	4	100
Observations	416			

Notes: Gini coefficient is calculated based on Agricultural census 2005 in our sample districts. Demographic characteristics are from 2001 population census. Work-population ratio is calculated as the number of total workers divided by total population. Total workers = main workers + marginal workers = Ag laborers + cultivators + household industry workers + Other workers.

Main workers were those who were engaged in any economically productive activity for 183 days (or six months) or more during the year. Marginal workers were those who worked for less than 183 days (or six months).

A person was considered as cultivator if he or she was engaged either as employer, single worker or family worker in cultivation of land owned or held from government (or private persons, institutions). In contrast, A person was regarded as an agricultural labourer if she/he worked in another person's land for wages in cash, kind or share.

5 Empirical Model and Results

5.1 Cross-sectional Results

I examine the effect of land inequality on public works provision by pooling the NREGA implementation data during 2006-2010 and using across district (within-state) variations of land concentration in 2005.[6] Using land inequality in 2005 (much ahead of the initiation of NREGA) allows for some control of potential reverse causality (i.e. it's reasonable that land inequality in 2005 will affect public work provision in post-2006, but unlikely that public works in post-2006 will affect land inequality in 2005). The model specification is:

$$Y_{i,s,t} = \alpha_0 + \beta * INE_{i,2005} + \alpha X_{it} + \alpha_s D_s + \alpha_t D_t + \varepsilon_{it}, \forall t \in \{2006, 2007, 2008, 2009, 2010\} \qquad (3)$$

where β is the coefficient of interest; i indexes districts, s states and t years.

$INE_{i,2005}$ denotes land inequality in district i in 2005, measured by Gini coefficient (in logarithm). $Y_{i,s,t}$ denotes the implementation of the NREGA program in district i, state s, in year t, measured by proportions of rural households provided with NREGA employment (in logarithm). A negative sign of β means NREGA job provision is negatively associated with land inequality. Standard errors are clustered at the district level.

To identify the effect of the concentration of landownership on public works provision, β, I need to control for variables that are correlated with land inequality and at the same time affect NREGA implementations. The first set of confounding factors contains the capacity of local governments to accommodate job supply to job demand. Such variables include the fraction of villages that have access to drinking water in the district, the fraction of villages with electricity, the fraction of villages with paved road and the fraction of villages with schools and other rural infrastructure variables. Second, I also include a more general economic development variable, the "Backwardness Index", a score constructed by Indian planning commission in 2003, with smaller numbers meaning being more economically backward. The literature has shown that NREGA program rolls out from backward districts to more affluent districts, in the order of their rankings on this index(Zimmermann, 2012; Dasgupta et al., 2017). Despite this, this roll-out rule might not be absolutely enforced, hence I also include phase dummies to better capture the heterogeneous implementation by phases. Third, I control for soil texture and the current wet season's rainfall deviations from historical means, because these geographic variables could affect both the demand for and the supply of NREGA jobs, and are also

[6] Table A.1 shows that land distribution didn't change at the statistically significant level during 2005 and 2010.

documented to be associated with land distribution.[7] Last, I include a vector of state dummy, D_s, and year dummy, D_t, restricting the cross-sectional comparisons to within-state variations.

The results of OLS estimates are presented in Table 4. The results are robust to adding extra covariates, all suggesting a significantly negative relationship between Gini coefficient and the proportion of rural households provided with NREGA employment. Column 1 has no covariates. Column 2 includes all covariates other than the Backwardness Index. Column 3 replaces the Phase indicators with the Backwardness Index. Column 4 includes all covariates. As the Backwardness Index is not available in some districts, controlling for this variable will decrease the sample size. Given that the results are robust to dropping this additional variable, for the remainder of the paper I exclude it from the covariates. Thus Column 2 becomes the baseline model, suggesting that districts with a 1% (or in absolute term, 0.0047) higher Gini coefficient would have 0.5% (or in absolute term, 0.005 * 30=1.5 percentage points) fewer households provided with NREGA jobs.

5.1.1 Using Selection on Observables to Assess the Bias from Unobservables

Despite the many observables included in the OLS model, there might still be omitted variables that are correlated with both land ownership and the demand/supply side of NREGA implementation. For instance, adverse geographical and climatic characteristics, on the one hand, may concentrate landownership by reducing the demand for land by marginal farmers,[7] and, on the other hand, may increase the demand for NREGA jobs. Therefore, if there are such geographical and climatic variables omitted, OLS estimates will be biased (upward and toward zero in the given example). I follow the method by Altonji et al. (2005) to assess the potential bias from unobservables.

This method provides a measure how much selection on unobservables, relative to selection on observables, has to be to explain away the estimated effect.[8] Assume β^R denotes the estimated coefficient for the variable of interest from the regression with a restricted set of covariates. Denote β^F the estimated effect for the variable from the regression with a full set of covariates. Then the ratio $\beta^F/(\beta^R - \beta^F)$ gives a sense of the size of selections. The larger the ratio is, the greater effect is required to fully explain away the estimated effects. In the current setting, the ratio could be calculated from the estimates in Column 1 and 2 of Table 4. To attribute the entire OLS estimate to selection effects,

[7] This relation between geographic and climate information and land ownership distribution is established by existing studies that use various geographical conditions to instrument for land inequality, including climatic information, soil quality and the share of cash crop (inequality-rising) and wheat/rice crop etc. (e.g. Easterly, 2007; Sokoloff and Engerman, 2000; Galor et al., 2009; Ramcharan, 2010; Cinnirella and Hornung, 2016; Baten and Juif, 2014). The spirits of these IVs are, small farmers are usually less able to hedge against negative weather shocks, and have a smaller demand for land in areas with poor soil quality (or in areas with violent rainfall variability). Thus, regions with poorer soil quality (or more rain variability) have higher land concentration.

[8] Nunn and Wantchekon (2011) provide a nice example of the application of this method.

Table 4: Dep. var.: % of households provided with NREGA jobs (OLS)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Gini coef. (log)	-0.550**	-0.529***	-0.751***	-0.777***
	(0.218)	(0.176)	(0.257)	(0.241)
log rural area(Sq. km)		0.214***	0.324***	0.291***
		(0.069)	(0.099)	(0.098)
log Rural population		-0.195***	-0.216**	-0.246***
		(0.066)	(0.088)	(0.083)
Literacy rate		-0.948***	-1.113***	-1.016***
		(0.330)	(0.386)	(0.345)
Wet season rainfall deviation		-0.095***	-0.091**	-0.091**
		(0.031)	(0.036)	(0.036)
% of land covered in fine soil		-0.168	-0.505**	-0.297
		(0.196)	(0.219)	(0.206)
% of land covered in medium soil		-0.344**	-0.506***	-0.389**
		(0.142)	(0.175)	(0.155)
% of Agricultural labourers		1.607***	2.039***	1.730***
		(0.392)	(0.471)	(0.461)
% of Main workers		-0.029	0.245	-0.104
		(0.697)	(0.828)	(0.737)
% of Marginal workers		2.678***	1.712*	1.175
		(0.718)	(0.912)	(0.792)
% of SCST population		0.467**	1.322***	0.851***
		(0.216)	(0.268)	(0.265)
% villages with Safe Drinking water		0.840**	1.758	1.438
		(0.331)	(1.223)	(1.199)
% villages with Electricity (Power Supply)		0.501**	0.573**	0.579**
		(0.215)	(0.283)	(0.276)
% villages with Paved approach road		-0.879***	-0.778***	-0.600**
		(0.235)	(0.296)	(0.281)
% villages with Primary school		0.198	0.188	0.096
		(0.262)	(0.341)	(0.328)
% villages with Medical facility		-0.485	-0.591	-0.508
		(0.325)	(0.396)	(0.396)
% villages with Post and telephone facility		-0.521**	-0.373	-0.358
		(0.230)	(0.258)	(0.246)
Phase 2 indicator		-0.164***		-0.180**
		(0.062)		(0.071)
Phase 3 indicator		-0.470***		-0.517***
		(0.069)		(0.083)
Backwardness Index			-0.107	-0.018
			(0.093)	(0.088)
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	1563	1563	1227	1227
R square	0.49	0.71	0.66	0.68

Notes: Dependent variable is the logarithm of the share of households provided with NREGA jobs in the district. Column 3 and 4 have a smaller sample size because the variable "Backwardness Index" is missing in some districts. Standard errors are in parentheses, clustered at district level. * $p < 0.10$, ** $p<0.05$, *** $p<0.01$.

Table 5: State-wise distribution of landlord and non-landlord districts

	Landlord	Non-landlord	Total districts
Andhra Pradesh	2	8	10
Bihar	1	0	1
Chhattisgarh	4	1	5
Gujarat	0	6	6
Haryana	0	4	4
Karnataka	0	11	11
Madhya Pradesh	10	1	11
Odisha	6	2	8
Punjab	0	5	5
Rajasthan	1	0	1
Tamil Nadu	2	9	11
Uttar Pradesh	12	34	46
Uttarakhand	0	3	3
West Bengal	10	0	10
Total	48	84	132

Source: This table is a subsample of districts that used to be part of British India (see Banerjee and Iyer, 2005) and are available in the Agricultural Census and the NREGA dataset. The table lists the 2001 districts, incorporating state and district boundary changes over 1961-2001 (Kumar and Somanathan, 2009).

selection on unobservables would have to be at least 25 (=-0.529/(-0.550+0.529)) times greater than selection on observables. This makes it less likely that the estimated effect is fully driven by selection on unobservables.

5.2 Addressing Endogeneity

I further address the potential endogeneity issues by taking advantage of historical institutions in India — land revenue collection system, established by British colonial rulers during 1750-1861. This variable is constructed based on the study by Banerjee and Iyer (2005). Land revenue, or land tax, was the major source of government revenue in India and during British times as well. British administration established three systems to collect land revenue in all cultivable land in British India: (a) landlord-based system, where the liability for a village or a group of villages lay with with a single landlord; (b) an individual cultivator-based system, where revenue settlements was made directly with individual cultivators; (c) village-based system, where village bodies which jointly owned the village were responsible for the land revenue. System (c), village-based system, could be further grouped as either system (a) or (b), depending on whether the village body was a single landlord or a large number of members with each person being responsible for a fixed share of the revenue. Table 5 presents state-wise distributions of landlord and non-landlord districts.

To identify a causal relation between land distribution and the provision of public works under

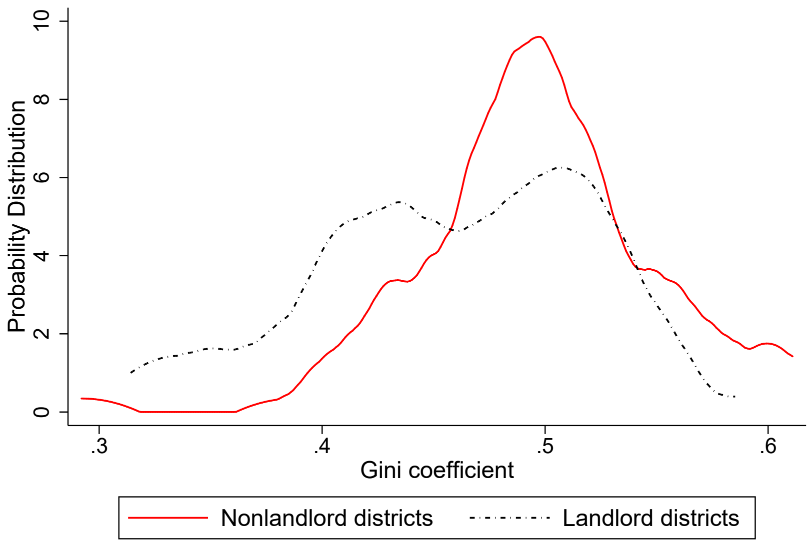

Figure 5: Visualize first stage —Land Inequality (Gini coefficient) in landlord/non-landlord districts
Source: The author calculated landownership Gini coefficients based on the 2005 India Agricultural Census.

NREGA, I use the binary indicator of land revenue system — whether this district was a landlord district in British India — to instrument for land inequality in 2005. The instrumental variable strategy rests on the assumption that land revenue collection systems under British India only affects redistributive policies through contemporary and current land inequality, after controlling for all observables. This is plausible because the way that British colonial rulers decided land revenue system in different areas was not based on a hard rule in terms of land fertility, weather or labor productivity (Banerjee and Iyer, 2005). Figure 5 visually presents the negative relationship between landlord-dominated revenue collection system and current land inequality.

I estimate first stage relation using the following equation:

$$INE_{i,2005} = \alpha_0' + \rho Z_i + \alpha' X_{it} + \alpha_s' D_s + \alpha_t' D_t + \eta_{it} \tag{4}$$

where Z_i is the binary indicator that equals to 1 if district i used to be a landlord-dominated district in British India, and zero otherwise; $INE_{i,2005}$ denotes land inequality in district i in 2005, measured by Gini coefficient; X_i denotes the same vector of district-wise covariates as in Equation(3).[9]

The first-stage conditional correlations suggest that landlord-dominated districts have 8% lower Gini landownership inequality in 2005 (Table 6). This estimated effect is equivalent to -0.038 (= -8%

[9]By restricting the variations to be within-state, in this IV estimation, states where land tenure systems do not vary across districts within the state will be absorbed in the state fixed effects, such as Bihar, Gujarat, Haryana, Karnataka, Punjab, Rajasthan, Uttarakhand and West Bengal. As a result, 91 districts in 6 states are left and contribute to the variations in the IV estimation.

Table 6: Dep. var.: Land inequality (gini coefficient) in 2005, (First stage)

	(1) OLS	(2) OLS
Landlord district indicator	-0.076***	-0.080***
	(0.021)	(0.020)
State FE	Yes	Yes
Year FE	Yes	Yes
Control variables	No	Yes
Observations	515	515
R square	0.57	0.74
F test: landlord indicator coef=0	13.38	16.17

Notes: "Landlord district indicator" equals 1 if the district in question was a landlord district (i.e. landlords were responsible for collecting land revenue) in British Raj. Land Gini coefficient is constructed using 2005 Indian Agricultural census. Standard errors are clustered at district level. * p < 0.10, ** p<0.05, *** p<0.01.

* 0.47) in the absolute term of Gini coefficient, or 0.5 (=0.038/0.08) standard deviation of the average Gini coefficient, given that the mean and standard deviation of Gini coefficient are respectively 0.47 and 0.08 in the sample.

The first-stage result that previously landlord-dominated districts in British India has a lower land inequality today is consistent with the study by Banerjee and Iyer (2005). They show that states with a higher landlord proportion had higher Gini measures of land ownership inequality in 1885, and this inequality persisted until the end of the colonial period.[10] However, as they argue, major landlord-dominated states enacted 6.5 land reforms in the period between 1957-1992, while non-landlord states had an average of 3.5. It was the greater number of land reforms in landlord districts in the post-Independence period that drove down the relative landownership concentration in those areas compared to non-landlord areas. According to Besley and Burgess (2000), states that enacted a larger number of land reforms had a greater decline of Gini coefficient of land inequality. Therefore, the previously landlord-dominated areas turn out to have lower landownership concentration than non-landlord-dominated areas due to more land reforms. Furthermore, the negative sign of first-stage results is consistent with the study by Besley et al. (2016) that shows in the long-run land inequality is lower in areas that saw greater intensity of tenancy reform.

I first plot the numbers of land reforms over time in major landlord and non-landlord states in Figure 6. It provides consistent evidence with the literature that landlord areas enacted more frequent

[10] Banerjee and Iyer (2005) explains why the choice of landlord revenue system had a strong effect on the distribution of land and wealth in the British India period. "Under landlord-based systems, the landlords were given a more or less free hand to set the terms for the tenants and, as a result, they were in a position to appropriate most of the gains in productivity."

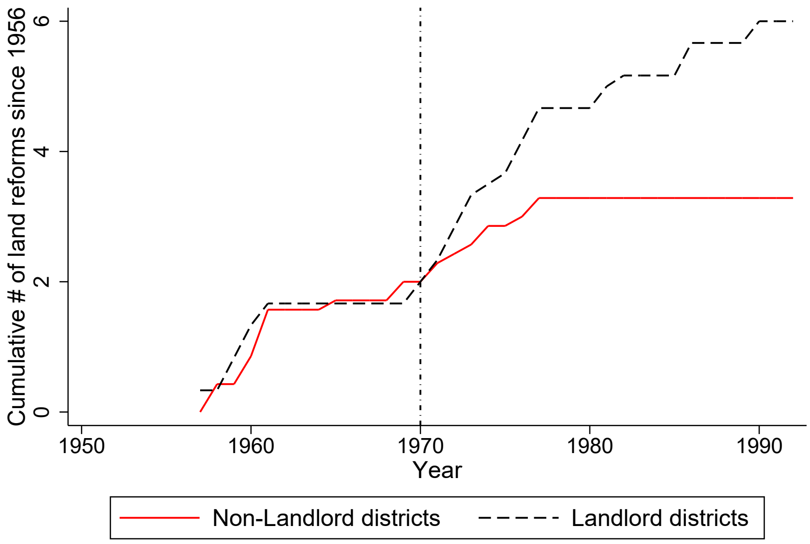

Figure 6: Frequencies of land reforms

Note: Information on the cumulative number of land reforms at the state level come from Besley and Burgess (2000). Major landlord-dominated areas are states with an above-median share of districts belonging to landlord dominated districts in British India.

land reforms than non-landlord areas, especially after 1970.[11] To depict since *when* landlord dominated districts started to have lower landownership inequality than non-landlord dominated districts, I further plot the trends of land inequality, measured by the share of land owned by the top 10% land holdings, for major landlord and non-landlord districts in Figure 7.[12] It shows that the shift of landlord districts from having relatively high land inequality to relatively low land inequality also occurred in 1970. Therefore, interestingly, the turning point of relative inequality in landlord versus non-landlord dominated districts coincides with the time when major landlord states started to outnumber non-landlord states in land reforms. All such information together explains the negative sign of first-stage estimate— landlord districts, although starting with higher land inequality in British Indian period, enacted more land reforms after Indian Independence, and hence ended up having lower land inequality in 2005.

The instrumental variable strategy relies on the assumption that land revenue collection system under British India only affects redistributive policies through contemporary and current land inequality, after controlling for observables. However, if different historical property rights institutions lead to persistent unobserved culture and institutional outcomes, and such unobserved outcomes are also correlated with redistributive policies, then this IV would violate the exclusion condition.

[11] State-wise land reform data come from Besley and Burgess (2000). Major landlord-dominated areas are states with an above-median share of districts belonging to landlord dominated districts.

[12] State-wise land distribution data come from Besley and Burgess (2000).

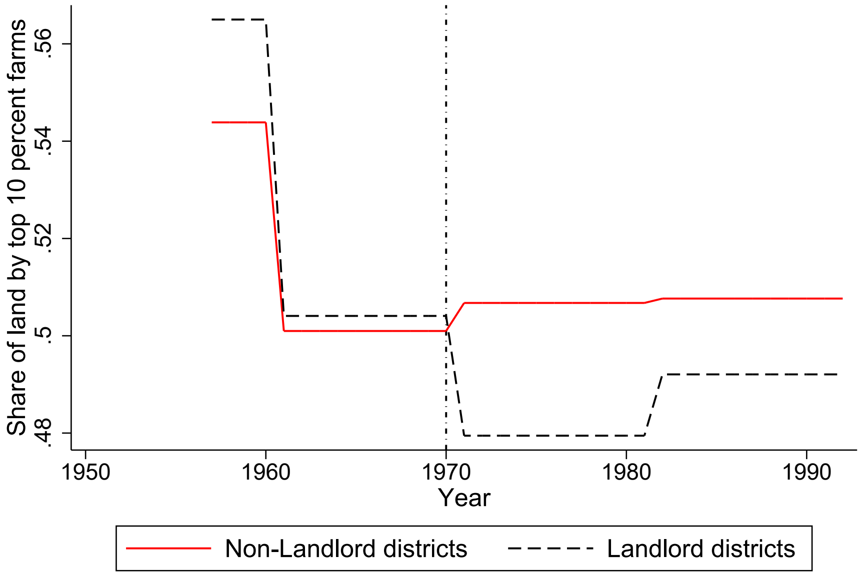

Figure 7: Trends of Land Inequality in landlord/nonlandlord districts

Note: Major landlord-dominated areas are states with an above-median share of districts belonging to landlord dominated districts in British India. Land inequality is measured by the share of land occupied by the top 10 percent biggest farmers, for which data come from Besley and Burgess (2000).

5.3 IV Results

Table 7 presents two-stage least square (2SLS) estimates of the effect of land inequality on NREGA implementation (measured by proportions of household provided with NREGA jobs). The sample size drops to one third of the original size, because the instrumental variable, the indicator of landlord dominated districts, is only defined in districts that were under British India during 1850-1947. To make the 2SLS results comparable to the foregoing OLS results, I present the subsample OLS results in Column 3 after presenting the 2SLS estimates in Column 1 and 2.

Column 1 only includes the inequality variable. Column 2 includes a full set of covariates in the estimation. The size of the estimated effect of the Gini land inequality is not sensible to adding covariates. The results show that a 1 percent increase in the Gini coefficient of landownership would have decreased the share of employed households by 5 percent (or equivalent to 1.5 percentage points, given that the average share of employed households is 30 percent). First-stage F statistics are above 10 in both models, suggesting a rejection of weak instrument null hypothesis.

Among all the covariates, the estimated coefficients of Phase 2 and Phase 3 indicator are both negative, with the former having a smaller magnitude than the latter. Phase 1 indicator serves as the reference group and omitted from the model. The results suggest that Phase 1 districts have the highest level of public employment provision, followed by Phase 2 districts and last Phase 3 districts. This

Table 7: Dep. var.: % of households provided with NREGA jobs (second stage)

	(1) 2SLS	(2) 2SLS	(3) OLS
Gini coef. (log)	-5.109***	-4.506***	-1.568***
	(1.983)	(1.496)	(0.431)
log rural area(Sq. km)		0.867***	0.557***
		(0.241)	(0.189)
log Rural population		-0.426***	-0.281*
		(0.162)	(0.155)
Literacy rate		0.605	-0.063
		(0.746)	(0.569)
Wet season rainfall deviation		-0.115*	-0.100*
		(0.062)	(0.060)
% of land covered in fine soil		-1.810***	-0.588
		(0.689)	(0.468)
% of land covered in medium soil		-1.126***	-0.505*
		(0.394)	(0.296)
% of Agricultural labourers		3.411***	2.613***
		(0.870)	(0.702)
% of Main workers		1.332	-0.044
		(1.835)	(1.663)
% of Marginal workers		-0.764	-1.145
		(1.839)	(1.523)
% of SCST population		0.254	0.450
		(0.512)	(0.486)
% villages with Safe Drinking water		-0.219	0.678
		(1.041)	(0.868)
% villages with Electricity (Power Supply)		0.548	0.546
		(0.500)	(0.399)
% villages with Paved approach road		-1.002**	-1.221***
		(0.464)	(0.392)
% villages with Primary school		1.582***	1.176***
		(0.455)	(0.385)
% villages with Medical facility		-0.381	-0.587
		(0.455)	(0.446)
% villages with Post and telephone facility		-1.477***	-1.328***
		(0.367)	(0.409)
Phase 2 indicator		-0.218*	-0.268**
		(0.127)	(0.111)
Phase 3 indicator		-0.384**	-0.486***
		(0.150)	(0.130)
State FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	515	515	515
First-stage F statistics	13.38	16.17	

Notes: Column 3 shows OLS results; Column 1-2 present 2SLS estimates. Dependent variable is the fraction of households provided with NREGA jobs in each year during 2006-2010. District-wise land Gini coefficient is constructed using 2005 Indian Agricultural census. Instrumental variable is a binary indicator that equals 1 if the district in question was a landlord district (i.e. landlords were responsible for collecting land revenue) in British Raj. Standard errors are clustered at the district level. * $p < 0.10$, ** $p<0.05$, *** $p<0.01$.

relative position resonates with the fact that NREGA rolls out from the most economically backward districts to richer districts.

To put the results into perspective, consider the difference of land Gini coefficients between two districts in Uttar Pradesh, Ballia and Allahabad. In Ballia, Gini coefficient of landownership was 0.486 (which was at the 50th percentile of the land distribution) in 2005; and in Allahabad, this number was 0.519 (which was at the 80th percentile). Using the estimate from Column 2, the difference of 0.033 points, or 6.4 percent (=0.033/0.518), in Gini coefficient implies that 28.8 percent (=6.4 * 4.5%) more households would have been provided with NREGA jobs in Allahabad if it had a land gini coefficient as small as Bellia's. Given the shares of households provided with NREGA jobs are, respectively, 22 percent in Bellia and 12 percent in Allahabad, this 28.8 percent increase would have eliminated one third (=28.8% * 12 / (22-12)) of the actual gap in job allocation rates between these two districts.

Column 3 reports OLS estimates by restricting the sample to the IV subsample, which serves as a reference to 2SLS estimates.[13] Both OLS and 2SLS estimations suggest a negative relation between land inequality and NREGA provision. In terms of the magnitude, 2SLS coefficient is about 3 times the OLS coefficient, suggesting that OLS results are biased upward (toward finding zero effect). A simple and possible source of endogeneity that leads to the upward bias of OLS results is measurement error in landownership distributions. As I approximate the size of land by the average size of land holdings in the size bin it belongs to, it unavoidably creates noise. Another source of bias might be some omitted variables that lead to less job provision in more equal areas. The three-times difference between IV estimates and OLS estimates is also in line with other studies that use geographical conditions to instrument for land inequality (Easterly, 2007; Cinnirella and Hornung, 2016; Ramcharan, 2010) [14].

[13]The OLS estimate in the subsample has a smaller magnitude than in the full sample. I explain this in Section 6.4.

[14] Note that I take the logarithm of the dependent variable and the Gini coefficient so that the estimated effect can be easily interpreted as percent changes. This is especially convenient when comparing estimates across different measurements of NREGA implementation (in the robustness check section). The disadvantage of taking logarithm lies in losing more than 100 observations that had zero NREGA jobs provided, most of which were in Phase 2 and Phase 3 areas before 2007. Therefore, to address the concern that the estimated results may be driven by sample selections, I use level regressions which will include districts that have zero NREGA jobs provided as well as districts with a positive number of jobs provided. The level of dependent variable and gini coefficient here are standardized by the standard deviation of the sample observations. All models in Table A.2 in the Appendix give negative signs, suggesting that the estimated effect does not rest on the choice of the logarithm. Column 4 shows that a 1 standard deviation increase of gini coefficient is associated with 0.6 standard deviation decline of the share of households provided with NREGA jobs.

6 Robustness of the Results

6.1 Plausible Exogenous Instrumental Variable

The credibility of 2SLS estimations rests on the identification assumption that the historical institution of landlord versus non-landlord dominated areas does not directly relate to the provision of public works other than through land distribution. However, this instrumental variable may only be plausibly exogenous rather than perfectly exogenous. In particular, Banerjee and Iyer (2005) have found that landlord-dominated areas are associated with lower public investment in the long term. If we take public works provision as one kind of public investments, then landlord-dominated areas might have lower provision of public employment today. Now I employ the method proposed by Conley et al. (2012) to show that this potential negative correlation will not invalidate the 2SLS estimate.[15]

Conley et al. (2012) relax the IV exclusion restriction by allowing the instrumental variable to also enter linearly in the second-stage regression with a coefficient, γ. The following equation is a generalization of this method:

$$Y = \beta X + \gamma Z + \varepsilon,$$

where β is the effect of interest, Z is the instrument variable, γ reflects how close the exclusion restriction is satisfied. The IV exclusion restriction is equivalent to the prior belief that $\gamma = 0$. The definition of plausible exogeneity is having prior information that implies γ is near 0 but perhaps not exactly 0. Without prior information or assumptions about γ, the parameters β and γ can not be jointly identified. Conley et al. (2012) show how to obtain the bounds for the IV estimate of the effect of interest (in the current paper, the effect of land inequality on public works provision, β) with prior information or assumptions about γ.

The concern that landlord-dominated areas might have lower public investment for reasons not captured by any observables is equal to say $\gamma < 0$. Applying the "Union of intervals" approach to the current paper, I find that if $\gamma < 0$, the bounds of β are further away from zero relative to the 2SLS estimate of β that assumes perfect exogeneity (that is, $\gamma = 0$). In other words, if landlord-dominated areas are associated with less NREGA employment even after controlling for all covariates, then 2SLS estimates provide an underestimation (in terms of absolute values) of the true effect of land inequality on NREGA job provision. Therefore, this alleviates the concern that long-term effects of colonial

[15] This method has been used in other studies to examine the sensitivity of estimation results to the violations of exogeneity conditions (e.g. Nunn and Wantchekon, 2011; Ding et al., 2009).

history (Banerjee and Iyer, 2005) may put a threat on the identification of β.

6.2 Alternative Measurements of NREGA Implementations

Both the OLS and 2SLS results presented in the previous section use the fraction of households provided with public employment to measure NREGA implementations. Table 8 presents estimates using alternative measurements of public works implementation — respectively, per capita labor expenditure, average days of employment provided to a person in either Schedule Caste or Schedule Tribe, and the total number of completed works per rural person.

Panel A shows OLS results using the full sample. Panel B, C and D present OLS estimates and 2SLS estimates using the IV subsample. The estimates are overall comparable across columns when using different measurements of land inequality. OLS results using the full sample show a 0.4-1 percent decrease in districts with a 1 percent higher Gini coefficient. OLS results in the IV subsample show slightly greater effects, between 1.2 and 1.8 percent. 2SLS results give even greater effects, around 3 percent. The relativeness of the estimates across models is comparable to that in the main results.

6.3 Does Gini Coefficient Capture the Top Distribution?—Alternative Measurement of Land Inequality

As Gini coefficients of landownership reflect the whole distribution of land holdings, a natural question arises—does gini coefficient capture the top distribution of land ownership? It is after all the top, rather than the middle or bottom, distribution of land holdings that reflects the concentration of big farms and big landlords' political power. Therefore, I construct shares of land owned by the top 10% largest land holdings to measure land inequality, following Besley and Burgess (2000).

The scatter plot of Gini coefficient and the shares of land owned by the top 10% land holdings, in Figure 8, shows that these two measurements of land inequality follow the same trends. The shares of land owned by the top 10% land holdings are higher wherever Gini coefficients are greater. This figure provides descriptive support that differences in Gini coefficients between districts are able to capture the relative differences of large landlords' land holdings. Then I use this alternative inequality measurement of landownership distribution to re-estimate the effect of land inequality on public works provision.

The estimation results are presented in Table 9. Each column uses an alternative measurement of NREGA implementation—share of households participating in NREGA employment, per capita labor expenditure, average days that each person in either Schedule Caste or Schedule Tribe worked

Table 8: Robustness checks: Alternative measurements of NREGA implementation

	(1) Labor Expenditure	(2) Persondays	(3) # of projects
Panel A: OLS using the full sample			
Dep. var.: NREGA job provision			
Gini coef. (log)	-0.452**	-0.587***	-1.227**
	(0.223)	(0.225)	(0.475)
Observations	1563	1500	1540
R square	0.70	0.67	0.47
Panel B: OLS using the IV sample			
Dep. var.: NREGA job provision			
Gini coef. (log)	-1.212**	-1.471**	-1.807*
	(0.553)	(0.600)	(0.994)
Observations	514	505	510
R square	0.68	0.68	0.47
Panel C: IV first stage			
Dep. var.: Gini coef. (log)			
Landlord district indicator	-0.080***	-0.080***	-0.078***
	(0.020)	(0.020)	(0.020)
Observations	514	505	510
R square	0.74	0.74	0.74
Panel D: IV second stage			
Dep. var.: NREGA job provision			
Gini coef. (log)	-2.941*	-3.404*	-3.872
	(1.705)	(1.916)	(2.963)
Observations	514	505	510
Kleibergen-Paap rk Wald F statistics	16.10	15.70	15.53

Notes: The table shows district-level OLS and IV estimates using three alternative measurements of NREGA implementation. The dependent variables in each model are, respectively, (log) per capita labor expenditure, (log) average days that each person in Schedule Caste or Schedule Tribe worked in NREGA, (log) total number of completed works per rural person. Panel A shows OLS results using the full sample. Panel B shows OLS results using the IV sample. Panel C shows the first stage results and Panel D 2SLS estimates, where the instrumental variable is the binary indicator that equals 1 if the district used to be a landlord district (i.e. landlords were responsible for collecting land revenue collection) in British Raj. All specifications include a full set of covariates, including year and state fixed effects, phase indicators and other covariates listed in the last column of Table 7. Standard errors are clustered at the district level. * $p < 0.10$, ** $p<0.05$, *** $p<0.01$.

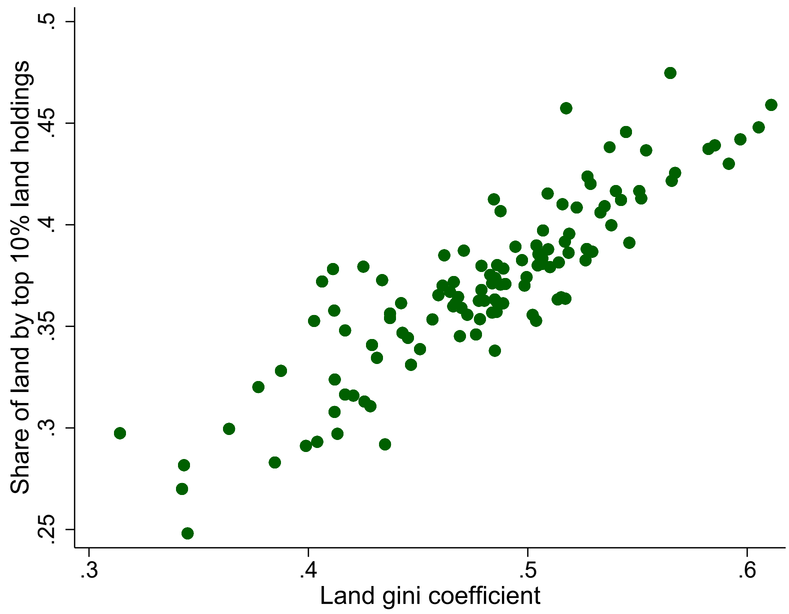

Figure 8: What does Gini coefficient capture? — Gini coefficient VS top 10% land holdings

Note: 2005 India Agricultural census.

in NREGA, the total number of completed works per rural person. Panel A shows OLS results using the full sample. Panel B-D show OLS and 2SLS results using the IV sample, where the instrumental variable is the historical land tenure indicator. The estimations all give negative signs on land inequality. For instance, if the share of land owned by the top 10 percent biggest land holdings increases by 1 percent (or in the absolute term, 0.46 percentage points, given the fact that the average shares of land owned by the top 10% is 46 percent), the share of households provided with NREGA jobs will decrease by 6 percent (or equivalent to 0.06 * 30%=1.8 percentage points, given that the average share of households participation is 30%). In addition, similar to results using Gini coefficient as the measurement of land inequality, 2SLS estimates have greater magnitudes than OLS estimates.

6.4 Is NREGA Demand Higher in More Equal Areas?

This paper mainly argues that relatively equal (in landownership) areas have more public jobs provided to the rural poor because of less interferences by landlords, rather than through other mechanisms such as a higher demand for jobs. The previous sections conclude this argument by controlling for a series of demand side factors in the empirical model. In this section, I directly examine the possibility of a higher demand for public jobs in areas with more equal land distributions, which will shed light upon what is driving the smaller job provision in these areas. Again, as the sample size in the 2SLS estimation drops by two thirds compared to the original OLS sample, I examine the relation between

Table 9: Robustness check: land inequality measured as the share of land by top 10% land holdings

	(1) % HHs employed	(2) Labor Expenditure	(3) Persondays	(4) # of projects
Panel A: OLS using the full sample				
Dep. var.: NREGA job provision				
Share of land by top 10% (log)	-0.553***	-0.486*	-0.484*	-0.897**
	(0.205)	(0.251)	(0.248)	(0.448)
Observations	1563	1563	1500	1540
R square	0.71	0.70	0.67	0.47
Panel B: OLS using the IV sample				
Dep. var.: NREGA job provision				
Share of land by top 10% (log)	-1.655***	-1.592***	-1.729**	-1.790**
	(0.442)	(0.556)	(0.664)	(0.844)
Observations	515	514	505	510
R square	0.65	0.68	0.68	0.47
Panel C: IV first stage				
Dep. var.: Share of land by top 10%				
Landlord district indicator	-0.059***	-0.059***	-0.058***	-0.057***
	(0.016)	(0.016)	(0.016)	(0.015)
Observations	515	514	505	510
R square	0.73	0.73	0.73	0.73
Panel D: IV second stage				
Dep. var.: NREGA job provision				
Share of land by top 10% (log)	-6.156***	-4.012*	-4.676*	-5.352
	(1.972)	(2.197)	(2.580)	(3.984)
Observations	515	514	505	510
Kleibergen-Paap rk Wald F statistics	14.24	14.20	13.88	14.07

The table shows OLS and IV estimates using the alternative measurement of land inequality, and using four alternative measurements of NREGA implementation. District-wise land inequality is constructed as the fraction of land owned by top 10% biggest land holdings from 2005 Indian Agricultural census. The dependent variables in each model are, respectively, (log) share of households participating in NREGA employment, (log) per capita labor expenditure, (log) average days that each SC/ST person worked in NREGA, (log) total number of completed works per rural person.
Panel A shows OLS results using the full sample. Panel B shows OLS results using the IV sample. Panel C shows first-stage results and Panel D 2SLS estimates, where the instrumental variable is the binary indicator that equals 1 if the district used to be a landlord district (i.e. landlords were responsible for collecting land revenue collection) in British Raj. All specifications include a full set of covariates, including year and state fixed effects, phase indicators and other covariates listed in the last column of Table 7. Standard errors are clustered at the district level. * $p < 0.10$, ** $p<0.05$, *** $p<0.01$.

the Gini coefficient and economic development indicators separately for these two samples.[16] Then I explore how agricultural wages could bias the estimated effects.

The top panel in Table 10 shows the relationship between Gini coefficient and economic development in the full sample. By some measures of economic indicators, such as the fractions of villages with access to paved roads and schooling, areas with more equal land distributions were less developed before 2006. The Backwardness Index and agricultural productivity (Rupees per hectare), however, were about the same in areas with different levels of Gini coefficients. In particular, agricultural wages, an important labor market variable, were slightly higher in these more equal areas before the introduction of NREGA. With higher agricultural wages in private sectors, the demand for NREGA jobs is presumably lower in such areas, hence eliminating the concern that the higher participation rate in these areas are caused by job demand rather than job supply. Column 1 and 2 in Table 11 reaffirm this argument by showing that adding the additional covariate, agricultural wages, into the original OLS model does not change the estimated effect much, and in fact, it slightly raises the estimate.

The second panel in Table 10 shows that in the IV sample, other than literacy rates and access to schools, economic characteristics do not vary by Gini coefficient at the 10% significant level. The fact that labor market characteristics are not worse in areas with more equal land distributions teases out the possibility that the higher NREGA participation in these equal areas is driven by a higher demand for public jobs. The comparisons in Column 3 and 4 (and Column 5 and 6) in Table 11 further reaffirm this argument by showing that adding the additional covariate, agricultural wages, into the old OLS model (IV model) only slightly affects the estimated effects.

The examination of the relation between Gini coefficient and economic development also provides insight on the advantage and disadvantage of OLS and IV estimation methods. As mentioned earlier, OLS estimates have the potential endogeneity issue arising from unobserved geographic and climate variables. By switching to the IV method, I can address the endogeneity issue in this regard, but meanwhile introduce another issue due to the change of sample representativeness. When sample size drops to 1/3, some properties of the original full sample disappear. For instance, as Column 1 in Table 10 shows, the more equal areas have higher agricultural wages in the full sample but not in the IV sample. As higher agricultural wages in local markets indicate a lower demand for NREGA jobs, the difference of this property in these two samples probably explains why OLS estimates have smaller magnitudes than 2SLS estimates.

[16] The economic variables are from two sources: components of the Backwardness Index (constructed in 2003) and infrastructure variables from the 2001 census. The fact that these variables were all collected before 2006 rules out the concern that agricultural wages and other economic variables were caused by NREGA implementations.

Table 10: Do districts with more equal land distributions have worse labor markets?

	(1) Ag wage	(2) Ag productivity	(3) Backward Index	(4) Literacy Rate	(5) % Marginal Worker	(6) % Ag Labor	(7) % Electr	(8) % Paved Road	(9) % School
Panel A: Full sample									
Gini coef.	-0.230**	0.187	-0.099	0.081*	0.014	0.056	0.078	0.104**	0.156**
	(0.108)	(0.223)	(0.140)	(0.042)	(0.014)	(0.037)	(0.059)	(0.052)	(0.073)
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1227	1227	1227	1227	1227	1227	1227	1227	1227
R square	0.62	0.27	0.43	0.44	0.30	0.55	0.74	0.81	0.44
Panel B: IV sample									
Gini coef.	0.206	0.368	0.264	0.129*	0.027	0.040	0.149	0.182	0.249**
	(0.175)	(0.528)	(0.310)	(0.070)	(0.032)	(0.073)	(0.135)	(0.134)	(0.121)
State dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	457	457	457	457	457	457	457	457	457
R square	0.60	0.19	0.39	0.46	0.34	0.57	0.73	0.75	0.47

Note: The table shows comparative economic characteristics in districts with differential land inequality. Each estimate is derived by regressing the economic indicator (column title) on Gini coefficient and state fixed effects. Information on agricultural wages (Rupees per day in 1995), agricultural productivity (Rupees/hectare in 1995) and Backwardness Index all comes from the 2003 Indian Planning Commission (Commission et al., 2003). "% of Electr" means the fractions of villages that have access to electricity in the district. The top panel provides the results for the full sample; the bottom panel uses the IV subsample. Standard errors are clustered at the district level.
* $p < 0.10$, ** $p<0.05$, *** $p<0.01$.

Table 11: Robustness of main effects to the control of agricultural wages

	Full sample		IV Subsampe			
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	2SLS	2SLS
Gini coef.	-0.78***	-0.80***	-1.59***	-1.44***	-4.16***	-4.13***
	(0.24)	(0.25)	(0.45)	(0.45)	(1.24)	(1.41)
Ag Wages (Rs/day, 1996)		-0.24*		-0.35		-0.03
		(0.13)		(0.24)		(0.32)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1227	1227	457	457	457	457
Kleibergen-Paap rk Wald F statistics					25.76	20.89

Note: Column 1 restates the OLS result from Table 4. Column 3 and 5 restate the main results from Table 7.
* p < 0.10, ** p<0.05, *** p<0.01.

6.5 Panel Data Model

Finally, I address the identification issue by undertaking an alternative strategy: panel data model. I collect a new wave of data, NREGA implementations during 2011-2014 and the Agricultural Census in 2010, into the previous dataset, so the new dataset covers 2006-2014. This makes a panel data structure and allows me to control for district fixed effects and take advantage of the variations of landownership distributions over time. All time-invariant unobservables will be removed by the district fixed effects. The estimation equation is as follows,

$$Y_{it} = \alpha_0 + \beta * INE_{it} + \alpha_i + \lambda_t + \varepsilon_{it}, \forall t \in \{2006, 2007, ..., 2014\} \tag{5}$$

where α_i denotes district fixed effects and λ_t year fixed effects.

The estimates in Table 12 reaffirm a negative relationship between land inequality and the provision of NREGA jobs. Column 1 includes state fixed effects, as in previous models. It reports a similar magnitude as the OLS estimations when using a short period of data (Table 4). Column 2 adds district fixed effects, but still reports a similar estimate. This suggests that including district fixed effects does not add more explanation power once I control for state fixed effects.

Table 12: Panel model results (2006-2014)

	(1) OLS	(2) FE
Gini coef. (log)	-0.496**	-0.598**
	(0.236)	(0.278)
Year FE	Yes	Yes
State FE	Yes	No
District FE	No	Yes
Observations	3200	3200
R square	0.47	0.07

Notes: This table reports results from estimating a fixed effect model. NREGA implementation data cover 2006-2014 and land inequality includes 2005 and 2010 two rounds. The dependent variable is the fraction of households provided with NREGA jobs (logarithm). No time varying control variables are included. Standard errors are clustered at the district level. * $p < 0.10$, ** $p<0.05$, *** $p<0.01$.

7 Conclusion

This paper studies how the concentration of landownership, a proxy for landlords' political power, affects the effective implementation of public works schemes in the context of The National Rural Employment Guarantee Scheme in India. Using district-level data on landownership distribution in 2005 and NREGA implementation during 2006-2010, I find that the concentration of landownership causes the reduction of public works provision. OLS estimates suggest that a 1 percent difference of land Gini coefficient leads to a 0.5 percent gap in NREGA provision in terms of household participation rates. To address the potential endogeneity issue arising from measurement errors and unobserved omitted variables, I apply three methods.

First, I apply the method of selections on observables (Altonji et al., 2005) to show that the results are less likely to be completely driven by unobservables. Second, I use a historical institution, the land revenue collection system established by British colonial rulers during 1750-1861, to instrument for land inequality. First-stage results show that previously landlord-dominated areas in British India has lower land inequality today, as a result of a higher number of land reforms after the Indian Independence. Under the assumption that the instrument is exogenous, the 2SLS estimates confirm the negative effect of landownership inequality on public works schemes. Both OLS and 2SLS results are robust to the use of alternative measurements of public works provision and land inequality. Third, by collecting a panel data set of land inequality and NREGA implementations, I estimate a panel data model and derive similar results. Finally, I exclude the possibility that the higher provision of public jobs in the more equal areas (in terms of landownership) is driven by a higher demand for public jobs, by showing that their local labor markets (especially agricultural wages in the private sector) is not worse than that in the more unequal areas.

Investigating the relation between land inequality and public works provision is not only relevant to India, but also has policy implementations for other developing countries, such as South Africa and Kenya, that have the dual need for job creation and investment in public services (such as road maintenance). More broadly, this paper adds to the discussion on how power asymmetries could hinder policies aimed at promoting equity. To improve policy effectiveness, the government needs to take into account asymmetries in bargaining power, which point is highlighted in 2017 World Bank Report (López-Calva et al., 2017). Future research would provide a more complete understanding of the economic consequences of land inequality and examine how power asymmetry begets economic inequality.

References

Agriculture Census Report (2005). All india report on agriculture census 2005-06. Technical report, Ministry of Agriculture, Department of Agriculture & Cooperation.

Alesina, A. and Rodrik, D. (1994). Distributive politics and economic growth. *The Quarterly Journal of Economics*, 109(2):465–490.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of political economy*, 113(1):151–184.

Anderson, S., Francois, P., and Kotwal, A. (2015). Clientelism in indian villages. *The American Economic Review*, 105(6):1780–1816.

Banerjee, A. and Iyer, L. (2005). History, institutions, and economic performance: the legacy of colonial land tenure systems in india. *The American Economic Review*, 95(4):1190–1213.

Bardhan, P. and Mookherjee, D. (2005). Decentralizing antipoverty program delivery in developing countries. *Journal of Public Economics*, 89(4):675–704.

Baten, J. and Juif, D. (2014). A story of large landowners and math skills: Inequality and human capital formation in long-run development, 1820–2000. *Journal of Comparative Economics*, 42(2):375–401.

Beg, S. (2016). Traditional elites: Agricultural productivity and the persistence of political power. *Working paper*.

Benabou, R. (2000). Unequal societies: Income distribution and the social contract. *American Economic Review*, pages 96–129.

Besley, T. and Burgess, R. (2000). Land reform, poverty reduction, and growth: Evidence from india. *The Quarterly Journal of Economics*, 115(2):389–430.

Besley, T., Leight, J., Pande, R., and Rao, V. (2016). Long-run impacts of land regulation: Evidence from tenancy reform in india. *Journal of Development Economics*, 118:72–87.

Bose, N. and Das, S. (2015). Political reservation for women and delivery of public works program.

Boustan, L., Ferreira, F., Winkler, H., and Zolt, E. M. (2013). The effect of rising income inequality on taxation and public expenditures: Evidence from us municipalities and school districts, 1970–2000. *Review of Economics and Statistics*, 95(4):1291–1302.

Census (2011). Administrative atlas of india - census of india 2011. Technical report, http://censusindia.gov.in/2011census/maps/maps2011.html.

Cinnirella, F. and Hornung, E. (2016). Landownership concentration and the expansion of education. *Journal of Development Economics*, 121:135–152.

Commission, P. et al. (2003). Identification of districts for wage and self employment programmes. *Report of the Task Force*.

Conley, T. G., Hansen, C. B., and Rossi, P. E. (2012). Plausibly exogenous. *Review of Economics and Statistics*, 94(1):260–272.

Dasgupta, A., Gawande, K., and Kapur, D. (2017). (when) do antipoverty programs reduce violence? india's rural employment guarantee and maoist conflict. *International Organization*, pages 1–28.

Ding, W., Lehrer, S. F., Rosenquist, J. N., and Audrain-McGovern, J. (2009). The impact of poor health on academic performance: New evidence using genetic markers. *Journal of Health Economics*, 28(3):578–597.

Dunning, T. and Nilekani, J. (2013). Ethnic quotas and political mobilization: caste, parties, and distribution in indian village councils. *American Political Science Review*, 107(01):35–56.

Dutta, P., Murgai, R., Ravallion, M., and Van de Walle, D. (2014). Does india's employment guarantee scheme guarantee employment?

Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics*, 84(2):755–776.

Galor, O., Moav, O., and Vollrath, D. (2009). Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence. *The Review of Economic Studies*, 76(1):143–179.

Gulzar, S. and Pasquale, B. (2016). Politicians, bureaucrats, and development: Evidence from india. *American Political Science Review*, pages 1–22.

Gupta, B. and Mukhopadhyay, A. (2016). Local funds and political competition: evidence from the national rural employment guarantee scheme in india. *European Journal of Political Economy*, 41:14–30.

Imbert, C. and Papp, J. (2015). Labor market effects of social programs: Evidence from india's employment guarantee. *American Economic Journal: Applied Economics*, 7(2):233–263.

Kumar, H. and Somanathan, R. (2009). Mapping indian districts across census years, 1971-2001. *Economic and Political Weekly*, pages 69–73.

López-Calva, L. F. et al. (2017). World development report 2017: Governance and the law.

Maiorano, D. (2014). The politics of the mahatma gandhi national rural employment guarantee act in andhra pradesh. *World Development*, 58:95–105.

Merfeld, J. D. (2019). Spatially heterogeneous effects of a public works program. *Journal of Development Economics*, 136:151–167.

Muralidharan, K., Niehaus, P., and Sukhtankar, S. (2017). General equilibrium effects of (improving) public employment programs: Experimental evidence from india. Technical report, National Bureau of Economic Research.

Nath, A. (2015). Bureaucrats and politicians: How does electoral competition affect bureaucratic performance? *Institute for Economic Development (IED) Working Paper*, 269.

Niehaus, P. and Sukhtankar, S. (2013). The marginal rate of corruption in public programs: Evidence from india. *Journal of Public Economics*, 104:52–64.

NREGA Operational Guidelines (2013). Mahatma gandhi national rural employment gurantee act 2005 – operational guidelines, 4th version. Technical report, Ministry of Rural Development, Department of Rural Development.

NREGA Report (2007). Mahatma gandhi national rural employment gurantee act 2005 – report of the second year april 2006 – march 2007. Technical report, Ministry of Rural Development, Department of Rural Development.

Nunn, N. and Wantchekon, L. (2011). The slave trade and the origins of mistrust in africa. *The American Economic Review*, 101(7):3221–3252.

Persson, T. and Tabellini, G. (1994). Is inequality harmful for growth? *The American Economic Review*, pages 600–621.

Ramcharan, R. (2010). Inequality and redistribution: Evidence from us counties and states, 1890–1930. *The Review of Economics and Statistics*, 92(4):729–744.

Santangelo, G. (2016). Firms and farms: The impact of agricultural productivity on the local indian economy. *Working Paper*.

Sheahan, M., Liu, Y., Barrett, C. B., and Narayanan, S. (2016). Preferential resource spending under an employment guarantee: the political economy of mgnregs in andhra pradesh. *The World Bank Economic Review*, page 044.

Sokoloff, K. L. and Engerman, S. L. (2000). History lessons: Institutions, factors endowments, and paths of development in the new world. *The Journal of Economic Perspectives*, 14(3):217–232.

Subbarao, K. (1997). Public works as an anti-poverty program: An overview of cross-country experience. *American Journal of Agricultural Economics*, 79(2):678–683.

Zimmermann, L. (2012). Labor market impacts of a large-scale public works program: evidence from the indian employment guarantee scheme. *Working Paper*.

Appendix A

Table A.1: Comparison of Land distribution in 2005 and 2010

Variables	2005 (mean)	2010 (mean)	Mean Diff
top10	0.359	0.357	-0.002
top20	0.528	0.525	-0.003
top30	0.648	0.645	-0.003
top40	0.739	0.735	-0.004
bot40	0.136	0.140	0.004
bot30	0.092	0.095	0.003
bot20	0.056	0.059	0.002
bot10	0.026	0.028	0.001
mid40_80	0.380	0.380	0
mid50_80	0.336	0.336	0
Gini	0.466	0.460	-0.006
Number of districts	528	561	

Notes: Top10 means the share of land by top 10% land holdings. mid40-80 denotes the share of land by middle 40-80 percent of land holdings.

Table A.2: Dependent variable: Share of households provided with NREGA jobs

	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS
Gini coef.	-0.04*	-0.18***	-0.61**	-0.61***
	(0.02)	(0.04)	(0.27)	(0.20)
Covariates	Yes	Yes	No	Yes
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	2085	660	660	660
R square	0.62	0.67	0.02	0.59
Kleibergen-Paap rk Wald F statistics			12.46	15.71

Notes: Column 1 shows OLS results in the full sample; Column 2-4 present estimates using the IV subsample. Dependent variable is the share of households provided with NREGA jobs in each year during 2006-2010. It is standardized by the respective standard deviations in each year. District-wise land Gini coefficient is constructed using 2005 Indian Agricultural census. Instrumental variable is a binary indicator that equals 1 if the district in question was a landlord district (i.e. landlords were responsible for collecting land revenue collection) in British Raj. * $p < 0.10$, ** $p<0.05$, *** $p<0.01$.