The Journal of Development Studies

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/fjds20

Land Inequality and Workfare Policies

Yanan Li & Naveen Sunder

To cite this article: Yanan Li & Naveen Sunder (2022): Land Inequality and Workfare Policies, The Journal of Development Studies, DOI: 10.1080/00220388.2021.2008362

To link to this article: https://doi.org/10.1080/00220388.2021.2008362

The Journal of Development Studies, 2021
https://doi.org/10.1080/00220388.2021.2008362

Land Inequality and Workfare Policies

YANAN LI * & NAVEEN SUNDER**
*Business School, Beijing Normal University, Beijing, China, **Economics, Bentley University, Waltham, MA, USA

(Original version submitted August 2020; final version accepted November 2021)

ABSTRACT *This paper contributes to the relatively scant literature on the impacts of inequality on the efficacy of public works programmes. We study this in the context of India. In particular, we examine the effect of land inequality on the implementation of the world's largest workfare programme – the National Rural Employment Guarantee Act (NREGA). Our OLS estimates demonstrate that the concentration of land ownership reduces the efficacy of NREGA. An instrumental variable (IV) analysis, where we use the historical land tenure system as an IV for contemporaneous land inequality, further corroborates our findings. This negative relationship is consistent with the hypothesis that public work schemes raise agricultural wages in the private labour market, thereby incentivising big landlords to use their political power to oppose such programmes. We exclude the possibility that the higher provision of public jobs in more equal areas is driven by a higher demand for public jobs or by caste or religious differences. This study suggests that the concentration of land ownership, a proxy for power asymmetries, could hinder effective implementation of development policies.*

KEYWORDS: Inequality; land ownership; public works

1. Introduction

While there is a substantial literature documenting the trends in inequality and its determinants (Imbert, 2012; Piketty & Saez, 2014; Van der Weide & Milanovic, 2018), there has been relatively less literature studying its consequences, with a few exceptions (for example, Boustan, Ferreira, Winkler, & Zolt, 2013; Cinnirella & Hornung, 2016; Ramcharan, 2010). In this paper, we examine the effect of inequality, as measured by land ownership Gini coefficient, on the implementation of a workfare programme in the Indian context. This particular policy, called the National Rural Employment Guarantee Scheme (NREGA), is the world's largest public works programme. It provides 100 days of employment annually to each household. India is an excellent context to explore the effects of inequality on the implementation of developmental policy because 1) it has the world's largest workfare programme, and 2) it has experienced a historical tension arising from land inequality (Chandramohan, 1998).

Existing literature has tried to explain the allocation of public jobs and funds in NREGA implementation from a variety of perspectives, such as political competition (Gupta & Mukhopadhyay, 2016; Nath, 2015), whether bureaucrats are supervised by a single political principal or multiple politicians (Gulzar & Pasquale, 2016), the 'golden goose' effect (expectations for future rent seeking opportunities) (Niehaus & Sukhtankar, 2013), the needs of potential beneficiaries (Sheahan, Liu, Barrett, & Narayanan, 2018) and the political reservation system (Bose & Das,

Correspondence Address: Yanan Li, Business School, Beijing Normal University, Xinjiekou Wai Street 19, Haidian District, Beijing, 100875, China. Email: yananli@bnu.edu.cn
Supplementary Materials are available for this article which can be accessed via the online version of this journal available at https://doi.org/10.1080/00220388.2021.2008362.

2015; Dunning & Nilekani, 2013). This paper adds to this growing literature by directly linking landownership inequality with the heterogeneity in the provision of NREGA jobs across districts. Past understanding would suggest that 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 (Das & Maiorano, 2019; Maiorano, 2014), but broad-based quantitative testing of this notion has been rarely attempted previously.

We estimate an OLS model where we control for state-level factors in addition to a number of other temporal and district-level controls and examine how within-state cross-district variations in land ownership inequality is associated with NREGA provision. The information on district-level land distribution comes from Agricultural Census data, while the data on NREGA provisions are from a publicly available government portal. For the main specifications, we use data for the years 2006 to 2010, but show robustness checks where we use data up to 2015. Land inequality is measured using the Gini coefficient, while we have four NREGA-related outcomes: the fraction of rural households provided with employment, the per capita labour 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. Our results suggest that a one per cent increase in land Gini coefficient leads to a decline of around 0.5–0.9 per cent in NREGA provision.

One concern is that the observed effects can potentially be driven by unobserved factors. We conduct two checks to assuage this concern. First, we apply Altonji, Elder, and Taber (2005) and use the selection on observed variables to assess the probability that the estimate is being driven by unobserved heterogeneity across districts. We show that the bias of selection-on-unobservables would have to be between two and six times as large as selection on a rich set of observables to fully drive the estimated effect down to zero. Second, we use the technique outlined in Oster (2019) to calculate bounds for the estimated effect. We find that all causal bounds are relatively narrow and do not include zero. Therefore, it is unlikely that the negative effects can be fully driven by unobservables.

As a complementary approach, we conduct an instrumental variable analysis where we instrument the potentially endogenous independent variable (land inequality) using the land revenue collection system established by British colonial rulers during 1750–1861 as an instrumental variable (Banerjee & Iyer, 2005). The first stage of the two-stage least squares (2SLS) approach leverages the association between historically landlord dominated areas (in 1750–1861) and current land inequality (in 2005). Under the assumption that the instrument is exogenous, the IV estimates confirm the negative effect of land ownership inequality on public works schemes. The 2SLS estimates suggest that a one per cent increase in inequality (land Gini coefficient) leads to a 2.7 per cent reduction in the share of rural households provided with employment. As a sensitivity check, we show that the negative effect still persists if we were to relax the exclusion restriction (perfect exogeneity) by allowing for a large enough value for the direct effect of the instrumental variable on the outcome of interest (Conley, Hansen, & Rossi, 2012). We also demonstrate that both the OLS and IV results are robust to using alternative measures of land inequality, and other nationally representative datasets for NREGA outcomes.

We finally discuss potential channels that might mediate the observed effects. We hypothesise that rural elite (landlords) exercise their economic and political power to affect the implementation of pro-poor policies that might negatively impact their control over the rural economy. This is especially true in the case of public works schemes. Providing public employment to the landless and the marginal farmersincreases labour wages (Imbert & Papp, 2015; Muralidharan, Niehaus, & Sukhtankar, 2020), which would incentivise landlord elites to oppose the implementation of the public works schemes (Anderson, Francois, & Kotwal, 2015; Maiorano, 2014). Our findings are also in line with past literature that finds inequality has a detrimental effect on the emergence of human-capital accumulating and growth-promoting institutions (for example, Galor, Moav, & Vollrath, 2009; Persson & Tabellini, 1994; Sokoloff & Engerman, 2000).

Our study sheds light upon the challenges in implementing a decentralised social protection programme in the context of large economic inequality. While decentralisation can in some cases lead to improvements in public service delivery in some cases (Bardhan & Mookherjee, 2005; Nagarajan, Binswanger-Mkhize, & Meenakshisundaram, 2014), it could make it vulnerable to political capture by local elites. Our finding is consistent with other studies on the same topic. For example, Maiorano, Das, and Masiero (2018) find that decentralisation in the implementation of NREGA promoted clientelism; while Ravallion (2019) finds that rationing of work opportunities can arise under decentralised implementation in poor places due to local administrative costs and local corruption. The policy implication of these studies is that to improve effectiveness of social protection programmes, the government needs to take into account political capture by local elites.

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 & Rodrik, 1994; Persson & 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 (for example, Galor et al., 2009; Ramcharan, 2010), casts doubt on this underlying mechanism. Instead, recent evidence suggests that inequality might be a hurdle for redistribution, provided that landlords, or better-endowed agents, have sufficient political power to influence redistribution policies.

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 mandated the ‘right to work’ for all rural Indian households through the implementation of the NREGA. It was a three-phased nationwide 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 programme in the world and guarantees 100 days of working for each household per financial year (June in the current year to May in the next year). Households are issued job cards by the local government, which are then used to record work done and payment received. According to the NREGA, as long as eligible household files applications for jobs, the local government must provide employment within 15 days and within 5 km of the applicant’s home. Otherwise, states are liable to pay unemployment allowances. In practice, however, there are frictions in the implementation of the policy leading to unmet demand (Dutta, Murgai, Ravallion, & Van de Walle, 2012). On average, each household works roughly 35 person-days per financial year, far less than the claimed 100 days. The extent of the unmet demand differs by districts and by time. Moreover, the political and caste affiliation of village elected leaders may also affect the provision of NREGA jobs (Johnson, 2009; Khosla, 2011).

While more than half of the projects done under NREGA are related to water conservation, other types of works are related to irrigation provision, land development and rural connectivity. Wages are to be paid at statutory minimum wage rates, which makes this programme a means of enforcing minimum wage laws. The literature largely finds that marginalised groups have disproportionately benefitted from NREGA. Overall, NREGA has been successful in ‘ensuring social protection for the most vulnerable people living in rural India’ (Government of India, 2013) – female workers and marginalised groups belonging to scheduled castes and scheduled tribes appear to be among the main beneficiaries of the Act (Klonner & Oldiges, 2014).

2.2. Financing NREGA and the supply constraint

The central government incentivises states to provide employment through the NRERA by subsidising 100 per cent of the unskilled labour cost and 75 per cent of the material cost of the programme. The labour to material ratio could vary from 90:10 to 60:40.[1] Additionally, there is a pre-determined cap on the overall annual labour spending on NREGA at state/district/ block/ village level, which is based on the labour budget that is determined in the previous year (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 prioritised to meet job demand. Table A1 presents the various steps involved in the preparation and finalisation of annual labour budgets. Because labour budget is an estimation and NREGA is a demand driven programme, 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 labour budget, following the procedures in Table A1. However, in fact, the flexibility is limited. Once the labour budget is finalised, the maximum supply of jobs in each state/district/block will not be changed for the next financial year.

3. Data

3.1. Land inequality

We measure land inequality at the district-level, and the information comes from the Indian Agricultural Census. This provides data for 33 states (528 districts) for the years 2000 and 2005. Our main measure of inequality is constructed based on operational land holdings[2] – the 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] We 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 A3). The largest 10 per cent of landholders operate about 46 per cent of the total land in India, as shown in Figure A.1. Moreover, there is large variation in the average Gini coefficients across different states (see Figure A.2), suggesting different extents of landownership concentration across states. Land distribution did not change at the statistically significant level during 2005 and 2010 (Table A4).

3.2. NREGA-related data

NREGA implementation data come from a Government of India source that is publicly available.[4] Table A2 presents summary statistics of NREGA implementation by financial year (starting from April in the current year and ends in March the next year). On an average 30 per cent of households worked for at least one day in public works in NREGA districts. The average earnings per rural person (regardless of their work status in NREGA) was between 212 and 266 Rupees during 2006–2010.

As noted earlier, NREGA was implemented in three phases. Districts were allocated to phases based on their ranking on a Backwardness Index (Zimmermann, 2012). We extract this index and its five components from a report by the Indian Planning Commission (Commission et al., 2003). A smaller value on the index implies a more economically backward district (and an earlier receipt of NREGA).

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 A.3, the average job provision by NREGA varies a lot by states. Consistent with the empirical findings, Figure A.4 visually presents a negative (unconditional) relation between land inequality and public works provision in a binned scatter plot.

In our analysis, we use other different pieces of information on districts/states in India. We provide details on the data used, and their sources in Appendix A.

4. Empirical model and results

4.1. OLS estimation results

We examine the effect of land inequality on public works provision using a state-year fixed effects model. For this, we stack the NREGA implementation data for the years 2006 to 2010 across different districts, and regress it on district-level land gini coefficients for 2005. Using land inequality in 2005 allows us to mitigate the possibility of reverse causality (that is it is 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:

$$Y_{ist} = \alpha_0 + \beta * INE_{is,2005} + \alpha X_{ist} + \alpha_{st} D_{st} + \varepsilon_{ist} \qquad (1)$$

where β is the coefficient of interest; i indexes districts, st state-by-year cells, $t \in \{2006; 2007;\ 2008; 2009; 2010\}$.

$INE_{is,2005}$ denotes land inequality in district i of state s in 2005, measured using Gini coefficient (in logarithm). Y_{ist} denotes the implementation of the NREGA program in district i of state s in year t, measured by proportions of rural households provided with NREGA employment (in logarithm).[5] A negative value of β implies that a higher level of land inequality is negatively associated with NREGA job provision, after controlling for a variety of variables that represent the capacity of local governments, and local geographic and demographic environment (full set of controls described in Table A3 in the Appendix). We include a vector of state-by-year dummy, D_{st}, restricting the cross-sectional comparisons to within state-by-year variations. Standard errors are clustered at the district level.

The results are robust to adding extra covariates, all suggesting a significantly negative relationship between Gini coefficient and NREGA employment. Results based on Equation (1) are presented in Table 1. In column 1, we present results based on a sparse specification with no controls – there is a strong negative association between land inequality and NREGA provision. In columns 2 and 3, we present results after adding some covariates – the negative association noted above still remains. Column 4 presents results with all covariates included. Since the Backwardness Index is not available in some districts, controlling for this reduces the sample size. Given that the results are robust to dropping this variable, we exclude it from the analysis for the remainder of the paper. We consider column 2 to be our baseline model, suggesting that districts with a 10 per cent (or in absolute terms, 0.047) higher Gini coefficient would have 5 per cent (or in absolute term, 0.05 * 30 = 1.5 percentage points) fewer households provided with NREGA jobs.

4.1.1. Using selection on observables to assess the bias from unobservables.. Despite the many covariates included in Equation (1), there might still be omitted variables that are correlated with both land ownership inequality and NREGA implementation. We apply the methods by Altonji et al. (2005) and Oster (2019) to assess the potential bias from unobservables.

Altonji et al. (2005) provide a measure of how large selection of unobservables, relative to selection of observables, has to be to explain away the estimated effect. Assume β_L denotes the coefficient of land inequality from a model that contains only land inequality and state-by-year fixed effects (column 1 in Table 1) and β_F refers to the coefficient of land inequality from a model containing all control variables and state-by-year fixed effects (columns 2–4 in Table 1). Then the ratio $|\beta_F/(\beta_L - \beta_F)|$ gives a sense of the size of selections. The larger the ratio is, the greater selection effect is required to completely explain away the estimated effects. We calculate the selection ratios for three models with different sets of covariates (column 2–4). For example, in column 2, selection of unobservables would have to be at least 6 (=−0.494/(−0.412 + 0.494)) times as large as selection of observables to drive the estimated effect down to zero.[6]

Generally, we are less concerned by selection-on-unobservables if the coefficient moves further away from zero or shows only small changes towards zero when adding observables. However, Oster (2019) points out that small changes in coefficients have to be assessed along with changes in R-squared values.

Table 1. Dep. Var.: % of households provided with NREGA jobs (OLS)

	(1)	(2) Full	(3) Full	(4) Full
Panel A				
Gini coef. (log)	−0.412*	−0.494***	−0.845***	−0.856***
	(0.235)	(0.178)	(0.287)	(0.263)
log rural area(Sq. km)		0.237***	0.349***	0.314***
		(0.076)	(0.109)	(0.107)
log Rural population		−0.245***	−0.281***	−0.308***
		(0.075)	(0.097)	(0.091)
Literacy rate		−1.106***	−1.166***	−0.972***
		(0.352)	(0.403)	(0.358)
Wet season rainfall deviation		−0.067*	−0.076*	−0.075*
		(0.039)	(0.041)	(0.041)
% of land covered in fine soil		−0.078	−0.491**	−0.229
		(0.202)	(0.221)	(0.209)
% of land covered in medium soil		−0.302**	−0.491***	−0.332**
		(0.149)	(0.178)	(0.153)
% of Agricultural labourers		1.637***	2.181***	1.796***
		(0.413)	(0.496)	(0.480)
% of Main workers		0.167	0.056	−0.408
		(0.703)	(0.843)	(0.754)
% of Marginal workers		2.868***	1.519	0.913
		(0.840)	(0.943)	(0.816)
% of SC/ST population		0.494**	1.394***	0.896***
		(0.249)	(0.286)	(0.274)
% villages with Safe Drinking water		0.650	1.903	1.296
		(0.483)	(1.207)	(1.161)
% villages with Electricity		0.575**	0.681**	0.694**
		(0.261)	(0.302)	(0.290)
% villages with Primary school		0.221	0.291	0.228
		(0.295)	(0.383)	(0.356)
% villages with Medical facility		−0.695**	−0.848**	−0.710*
		(0.336)	(0.400)	(0.399)
% villages with Post and telephone facility		−0.667**	−0.605**	−0.550*
		(0.262)	(0.297)	(0.280)
Phase 2 indicator		−0.219***		−0.182**
		(0.069)		(0.074)
Phase 3 indicator		−0.538***		−0.588***
		(0.072)		(0.083)
Backwardness Index			−0.117	−0.012
			(0.096)	(0.086)
Observations	1,575	1,575	1,222	1,222
R-squared	0.527	0.720	0.710	0.734
State-by-Year Fixed Effects	Y	Y	Y	Y
Controls	N	Y	Y	Y
Panel B				
Selection Ratio $\lvert \beta_F/(\beta_L - \beta_F) \rvert$		6.0	2.0	1.9
Identified β-set		[−0.576, −0.494]	[−1.288, −0.845]	[−1.275, −0.856]

Notes: Dependent variable in Panel A is the logarithm of the share of households provided with NREGA jobs in the district. Phase 2 and 3 districts in 2006 and phase 3 districts in 2007 are excluded. Column 3 and 4 have a smaller sample size because the variable Backwardness Index is missing in some districts. Panel B assesses the severity of potential selection bias. β_L denotes the estimated coefficient for Gini coefficients from a model with a restricted set of covariates. βF denotes the estimated effect for Gini coefficients from a model containing all covariates listed in the column. The selection ratio indicates the extent of remaining selection bias due to unobservables relative to the observable variables necessary to fully explain the estimated effect. A detailed definition of the identified set is provided in the main text. The set is well identified if it does not include 0 (see Oster, 2019). Robust standard errors are in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

Only when the additional variables also explain the additional variation in the dependent variable, do small changes in coefficients help to suggest a causal interpretation. Let *Rmax* denote the maximum share of the variance of the dependent variable that could be explained, that is the R-squared value from a regression of NREGA job provision on land inequality, all observables and unobservables. Let R_F denote the R-squared value from a regression of NREGA job provision on land inequality with all covariates (columns 2–4 in Table 1) and R_L the R-squared value from a regression with a restricted set of covariates (column 1). Further, assume the ratio of selection-on-unobservables to observables is $\delta \in [0, 1]$.

We then bound the effect of land inequality on NREGA job provision using the following formula $\beta^*(R_{max}, \delta) = \beta_F - \delta[\beta_L - \beta_F]\frac{R_{max}-R_F}{R_F-R_L}$. Oster (2019) suggests that in most situations the maximum R-squared value could be set to $R_{max} = min\{1.3R_F; 1\}$ and the relative ratio could be set at $\delta = 1$. The causal effect will lie between β_F and β^*. Panel B in Table 1 shows that the bounds under different model specifications all exclude zero. For example, the set of the identified effect in column 2 is [−0.576, −0.494], far below zero. This provides strong evidence in favour of a causal and statistically significant relationship, while suggesting that selection bias is unlikely to explain away the observed effects.

4.2. Instrumental variables (IV) approach

We further address the potential endogeneity of land inequality by leveraging historical land institutions in India, more particularly the land revenue collection system established by British colonial rulers during 1750–1861. Under colonial rule, land revenue (or land tax) was the major source of government revenue in India. British administration established three different systems to collect land revenue in British India: (a) landlord-based system, where the liability for a village or a group of villages lay with a single landlord; (b) an individual cultivator-based system, where revenue settlements were 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), the village-based system, could be further classified into 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. Therefore, in effect, there were two major land revenue collection systems under the British. Banerjee and Iyer (2005) classify different districts (based on boundaries in British India) as being landlord dominant (system A) or not (system B). Table 2 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 NREGA, we use the binary indicator of land revenue system, whether a district was a landlord district (or not) under the British, to instrument for contemporaneous land inequality (in 2005). The instrumental variable strategy rests on the assumption that land revenue collection systems under British India only affect 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 any hard rule in terms of land fertility, weather or labour productivity (Banerjee & Iyer, 2005), and this is expected to determine land policy in the future in these regions.

We estimate first-stage relation using the following equation:

$$INE_{is,2005} = \alpha'_0 + \rho Z_{is} + \alpha' X_{ist} + \alpha'_{st} D_{st} + \eta_{ist} \tag{2}$$

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

Table 2. State-wise distribution of landlord and non-landlord districts

	Landlord Dominated Districts	Non-landlord Dominated Districts	Total Districts
Andhra Pradesh	3	10	13
Assam	2	14	16
Bihar	1	0	1
Chhattisgarh	11	2	13
Gujarat	1	10	11
Haryana	0	14	14
Himachal Pradesh	0	10	10
Karnataka	0	16	16
Kerala	0	6	6
Madhya Pradesh	12	0	12
Meghalaya	0	7	7
Odisha	17	6	23
Punjab	0	8	8
Rajasthan	1	0	1
Tamil Nadu	6	17	23
Uttar Pradesh	15	50	66
Uttarakhand	0	10	10
West Bengal	12	0	13
Total	81	180	261

Source: The information is extracted from the paper by Banerjee and Iyer (2005).
Notes: This table is a subsample of districts that used to be part of British India (see Banerjee & Iyer, 2005) and are nonmissing in the Agricultural Census and the NREGA dataset. The table lists the 2001 districts, incorporating state and district boundary changes over 1961–2001 (Kumar & Somanathan, 2015).

The first-stage conditional correlations suggest that landlord-dominated districts have 7 per cent lower Gini coefficient in 2005 (Panel A of Table 2). This estimated effect is equivalent to 0.033 (= −7 per cent * 0.47) in the absolute value of Gini coefficient, or 0.5 (= 0.033/0.07) standard deviation of the average Gini coefficient, given that the mean and standard deviation of Gini coefficient are respectively 0.47 and 0.07 in the sample respectively. Figure A.5 visually presents the negative relationship between landlord-dominated revenue collection system and current land inequality. The result that historically landlord-dominated districts have a lower land inequality in modern times is consistent with the findings in Banerjee and Iyer (2005). They show that states with a higher historical landlord proportion had higher Gini measures of land ownership inequality in 1885, and this inequality persisted until the end of the colonial period.[7] However, they show that major landlord-dominated states enacted more land reforms (6.5) in the period between 1957–1992 as compared to non-landlord states that implemented an average of 3.5 land reforms.[8] 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. This is also suggested by Besley and Burgess (2000) and Besley, Leight, Pande, and Rao (2016), who demonstrate that 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.

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 cultural and institutional outcomes, and such unobserved outcomes are also correlated with redistributive policies, then this IV would violate the exclusion condition. We assuage this concern in Section 4.4 using the technique of plausibly exogenous IV (Conley et al., 2012).

Table 3. IV estimates of the effects of land inequality on NREGA job provision

	(1)	(2)	(3)
	2SLS	2SLS	OLS
***Panel A**: first stage*			
Dep. var.: Gini coef. (log)			
Historical landlord district indicator	−0.071***	−0.073***	
	(0.018)	(0.018)	
Covariates	No	Yes	
State-by-Year FE	Yes	Yes	
Observations	1011	1011	
***Panel B**: second stage*			
Dep. var.: % of households provided with NREGA jobs			
Gini coef. (log)	−2.451**	−2.711**	−1.291***
	(1.212)	(1.168)	(0.306)
Covariates	No	Yes	Yes
State-by-Year FE	Yes	Yes	Yes
Observations	1011	1011	1011
Kleibergen-Paap rk Wald F statistics	15.51	16.79	

Source: Author's analysis based on data from NREGA public data portal (2006–2010) and the 2005 Indian Agricultural Census.
Notes: The sample of the analyses contains the districts that were part of the British India during 1850–1947, as the instrumental variable is only defined on this sub-sample. Column 1 reports 2SLS estimates without district-level control variables. Column 2 reports 2SLS estimates with a full set of control variables as in Table 1 . Column 3 conducts a separate OLS regression using the IV sample. Panel A reports first stage results of the 2SLS estimation, where the historical land tenure institution is an instrument variable for contemporaneous landownership inequality. 'Historical landlord district indicator' equals one if a district was a landlord district under the British, and zero otherwise. District-wise land Gini coefficient is constructed using the 2005 Indian Agricultural census. Panel B (column 1 and 2) presents second stage results. The dependent variable is the fraction of households provided with NREGA jobs in each year during 2006–2010 (logarithm). Standard errors are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.3. IV Results

Panel B of Table 3 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 by one-third of the original size because the instrumental variable, the indicator of landlord dominated districts, is only defined for a subset of districts that were under the British tenancy system during 1850–1947. To make the 2SLS results comparable to the results presented in Table 1, we present the corresponding results for this subsample in Column 3 (Panel B) of Table 3.

Column 1 of Panel B 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 sensitive to the addition of covariates. The results suggest that a 10 per cent increase in the Gini coefficient of landownership would decrease the share of employed households by 27 per cent (or equivalent to 8 percentage points, given that the average share of employed households is 30 per cent). First-stage F statistics are above 10 in both models, suggesting a low possibility of weak instruments.

Column 3 of Panel B reports estimates based on Equation (1) after restricting the sample to observations included in the IV specification. Both OLS and 2SLS estimations suggest a negative relation between land inequality and NREGA provision. In terms of magnitude, 2SLS coefficient is about twice as large as the OLS coefficient, suggesting that OLS results are biased upwards (towards finding zero effect). The difference of this magnitude between the IV and OLS estimates is in line with other studies that use geographical conditions to instrument for land inequality (Cinnirella & Hornung, 2016; Easterly, 2007; Ramcharan, 2010).

4.4. Plausibly Exogenous instrumental variable

The credibility of the 2SLS estimation rests on the identification assumption that the instrument (historical land tenure system) does not affect the outcome (provision of public works) through channels other than the main independent variable of interest (land inequality). However, it is possible that the historical land tenure system might have a direct effect on the provision of NREGA employment beyond the impact mediated by land inequality (Misra, 2019). For example, Banerjee and Iyer (2005) have found that landlord-dominated areas are associated with lower public investment in the long term. Therefore, we employ the method proposed by Conley et al. (2012) to explore what effect such a correlation will have on our findings.[9]

To illustrate how the IV estimate is sensitive to the violation of exclusion restriction, we plot the bounds for the second-stage estimate with varying prior assumptions on the exclusion condition, following the Local To Zero (LTZ) approach (Conley et al., 2012). We assume that the direct effect of the instrument (the historical landlord tenure system) on public employment provision is normally distributed with mean 0 and variance δ^2, $\gamma \sim N(0, \delta^2)$ with $\delta > 0$. By varying δ, we can find the thresholds that would make the second-stage estimate be significant at the 95 per cent level. Figure 1 confirms that the bounds for the second-stage effect exclude zero as long as the standard error of the prior distribution is smaller than 0.065 (that is $\delta < 0.065$). Note that this is providing a rather wide range for the failure of the exclusion restriction, as 0.065 is around one third of the overall reduced-form effect of the instrument on NREGA job provision.[10] Therefore, the negative impact of land inequality on NREGA job provision is robust to a large degree of potential endogeneity in the IV.

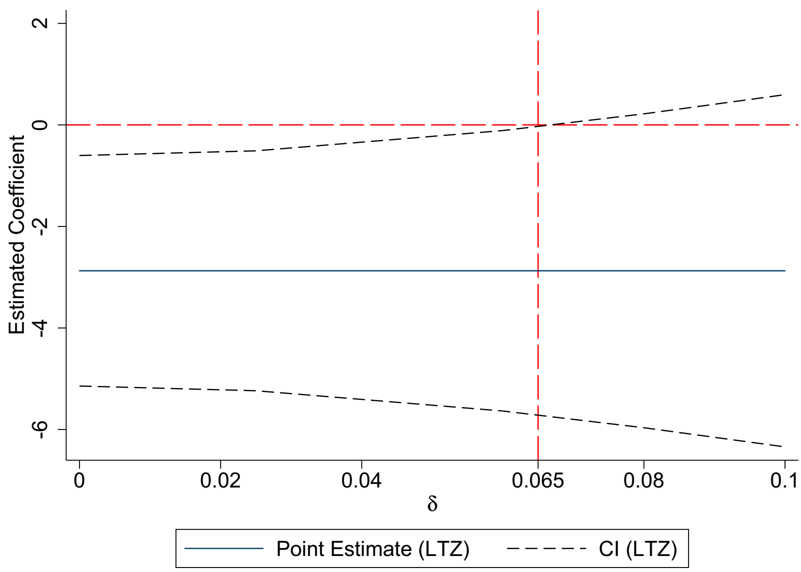

Figure 1. Plausibly Exogenous bounds for the estimated effect β.

Notes: This figure plots the plausibly exogenous bounds for the second-stage effect (that is, gini) following the Local To Zero (LTZ) approach (Conley et al., 2012). The confidence interval is at the 95 per cent level. We assume the direct effect of the instrument (that is, the historical landlord tenure system) on public employment provision is normally distributed as $\gamma \sim N(0, \delta^2)$. The vertical red line indicates the value of the upper bound for δ that would make the second-stage estimate fully to the left of zero at the 95 per cent level. The figure is plotted using the Stata command 'plausexog' (Clarke, 2019).

5. Robustness of the results

5.1. Cross-sectional results by year

The OLS results in Table 1 uses data for the years 2006–2010. In Table S.2 (in Supplementary Materials), we disaggregated results for each year separately. For each year, we first estimate a model containing only state fixed effects and land inequality and use it as the restricted model. We then estimate a model containing a full set of covariates (as in column 2, Table 1) and state fixed effects. Panel A shows that the stability of the negative effect across years and across different model specifications is remarkable. We then assess the severity of potential selection bias for the results applying Altonji et al. (2005) and Oster (2019). As shown in panel B, selection on unobservables would have to be 4–8 times as large as selection on observables to fully explain the estimated effects in each year. The bounds for the estimated effects are all tight and far below zero, suggesting a causal relationship of land inequality in NREGA job provision.

5.2. Alternative measures of NREGA implementation

All results presented thus far have used the fraction of households provided with employment to measure NREGA implementation. Table 4 presents estimates using alternative measures of NREGA efficacy – respectively, per capita labour 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. Panels 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 3.7–6.5 per cent decrease in districts with a 10 per cent higher Gini coefficient. OLS results in the IV subsample report slightly greater effects. 2SLS results give even greater effect sizes. The relative size of the estimates across models are comparable to that in the main results (Table 1).

5.3. Alternative measures of land inequality

As Gini coefficient of land ownership reflects the whole distribution of land holdings, a natural question arises – does gini coefficient capture the upper end of the land ownership distribution? 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, we create an alternative measure of land concentration – share of land owned by the top 10 per cent largest land holdings, following Besley and Burgess (2000).

The scatter plot of Gini coefficient and the shares of land owned by the top 10 per cent land holdings (Figure 2) shows that these two measurements of land inequality follow the same trends. Gini coefficient is positively correlated with the share of land owned by the top 10 per cent land holdings. This figure provides descriptive support that differences in Gini coefficients between districts are able to capture the relative differences of large landlords' land holdings.

The estimation results based on this new measure are presented in Table S.3 (in Supplementary Materials). Each column uses an alternative measurement of NREGA implementation. The results all suggest negative causal effects of land inequality on NREGA job provision.

6. Mechanism and discussions

6.1. The role of landlord elites

Providing public jobs to the landless and marginal farmers generally leads to increases in rural wages (Imbert & Papp, 2015; Merfeld, 2019; Muralidharan et al., 2020), which will potentially raise the cost of production for landlords who hire casual labourers. This wage effect may incentivise

Table 4. Robustness checks: alternative measurements of NREGA implementation

	(1) Labour Expenditure	(2) Person Days	(3) # of Projects
Panel A: OLS using the full sample			
Dep. var.: NREGA job provision			
Gini coef. (log)	−0.373*	−0.450**	−0.646*
	(0.218)	(0.208)	(0.373)
Observations	1552	1490	1530
R square	0.79	0.76	0.59
Selection Ratio $(\beta_F/(\beta_L - \beta_F))$	1.9	3.0	22.5
Identified β-set	[−0.548, 0.373]	[−0.620, −0.450]	[−0.679, −0.646]
Panel B: OLS using the IV sample			
Dep. var.: NREGA job provision			
Gini coef. (log)	−0.864**	−0.840**	−1.053**
	(0.335)	(0.338)	(0.533)
Observations	1005	965	983
R square	0.76	0.72	0.60
Selection Ratio $(\beta_F/(\beta_L - \beta_F))$	7.1	3.3	31.5
Identified β-set	[−0.945, −0.864]	[−1.048, −0.840]	[−1.095, −1.053]
Panel C: IV first stage			
Dep. var.: Gini coef. (log)			
Historical Landlord district	−0.073***	−0.075***	−0.073***
	(0.018)	(0.018)	(0.018)
Observations	1005	965	983
R square	0.67	0.67	0.67
Panel D: IV second stage			
Dep. var.: NREGA job provision			
Gini coef. (log)	−2.584**	−3.023**	−1.803
	(1.221)	(1.244)	(1.875)
Observations	1005	965	983
Kleibergen-Paap rk Wald F statistics	16.71	17.20	15.93

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 labour 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. The selection ratio indicates the extent of remaining selection bias due to unobservables relative to the observable variables in the model that would be necessary to drive the treatment effect down to 0. The identified set is calculated by applying Oster (2019). Panel C shows the first stage results and Panel D 2SLS estimates, where the instrumental variable is the binary indicator that equals one if the district used to be a landlord district 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 second column of Table 1. Standard errors are clustered at the district level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

landlords to oppose the implementation of NREGA (Anderson et al., 2015; Maiorano, 2014). Landlord elites are a big and powerful lobby, and they exert substantial authority over decision-making at the community level. They are reported to be heavily involved in nominating and mobilising consensus-based support for political candidates (Arora, 2020). A substantial literature has documented the manipulating role of landowning elites in the distribution of public goods services in a variety of contexts (Bardhan & Mookherjee, 2012; Bhattacharya, Kar, & Nandi, 2016; Faguet, Sánchez, & Villaveces, 2020). Therefore, it is not surprising that landlord elites could influence NREGA public works and employment provision to their benefit. When the NREGA was implemented, most landowners resented the increased bargaining power of the labourers and their subsequent increased social status in the village's political economy (evidence from interviews with local journalists and government officials by Das & Maiorano, 2019; Maiorano, 2014).[11] Indeed, government officials at all levels are subject to pressures from the farmers' lobby

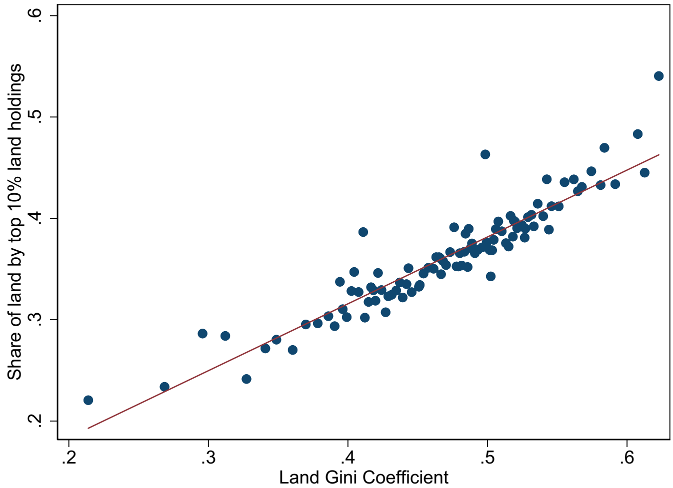

Figure 2. What does Gini coefficient capture? – Gini coefficient vs. top 10 per cent land holdings.
Source: Author's calculation based on 2005 Indian Agricultural Census.
Notes: The vertical axis is district-wise land inequality constructed as the fraction of land owned by top 10 per cent biggest landowners. The horizontal axis is district-wise gini coefficient of landownership.

against NREGA. The lobby is so strong that even the ruling party does not use the programme to win votes in swing areas (Das & Maiorano, 2019).

There are a few ways in which big landlords can intervene the process of providing NREGA jobs. At the stage of drafting the labour budget, landlords may lobby for a budget plan that does not provide enough jobs to the rural poor. Landowning elites can also exercise their power in defining the priority areas of NREGA-related projects and promote activities that are of directrelevance to themselves (Raabe et al., 2010). This is possible because of the bottom-up approach to budget planning (outlined in Table A1). The demand for NREGA jobs and the roster of projects are first identified at the Gram Panchayat level and are then 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.

Even after the labour budget is finalised, big farmers can still use their political power to block the implementation, such as delaying work assignment, payment and some complementary machinery (Maiorano, 2014). In particular, landlords, in the form of farmers' representatives, can influence NREGA implementation through the Field Assistant (FA). The FA is a government employee who is hired on a contract basis to implement the NREGA. FAs typically have 'nearly absolute power as far as the implementation of the MGNREGA is concerned'. They play a crucial role in selecting NREGA beneficiaries and are prone to influence of the political and economic elite (Maiorano, 2014). For instance, 'farmers who want MGNREGA to be stopped during the agricultural season exercise pressures on the FAs to do so and if the pressure is not sufficient, they approach the local MLA (interview with the Minister for Rural Development, Hyderabad, 5 August 2013)'.

6.2. Is NREGA demand higher in more equal areas?

The results from our analysis suggest that relatively equal (in land ownership) areas have more public jobs provided to the rural poor. It is possible that these effects are potentially driven by higher demand for NREGA in relatively more equal regions. Here, we directly examine this possibility. In particular, we explore whether the demand for NREGA jobs is higher in low Gini coefficient districts because of the prevalence of lower agricultural wages and other economic indicators (which are direct correlates of demand for NREGA jobs) in the pre-NREGA period.

We collect information on various economic indicators from two sources: the five components of the Backwardness Index (constructed in 2003) and infrastructure-related variables from the 2001 census. We use data from the pre-NREGA period to rule out reverse causality. The results are presented in Table S.4 (in Supplementary Materials) – they show the relationship between Gini coefficient and economic development in the full sample. We find that there is largely no significant relationship between the level of land inequality (gini coefficient) and various measures of economic development and labour market. In particular, agricultural wages, an important labour market variable, were slightly higher in more equal areas before the introduction of NREGA (which would imply potentially weaker NREGA demand). Therefore, there seems to be little evidence supporting a higher demand for NREGA jobs in more equal regions – thus eliminating this as a potential channel.

6.3. Caste and religious diversity

Past literature has indicated that increased social diversity significantly reduces the supply of local public goods. This is potentially due to reduced cooperation or the lack of effective social sanctions across ethnic groups (Alesina, Baqir, & Easterly, 1999; Miguel & Gugerty, 2005; Munshi & Rosenzweig, 2018). In our study, social diversity could be a potential confounding factor if district-level social diversity is correlated with both NREGA job provision (outcome) and land inequality (main independent variable of interest). Caste and to some extent religion are two main forms of social diversity in India. It is possible that the land inequality covariate picks up the effect of caste/religion-based cohesion rather than through big landlords.

Most previous estimation models discussed in this analysis have controlled for caste composition (measured as the fraction of SC/ST population out of the total population in the district). This alleviates the concern that our results are driven by the effect of caste composition. We try to assuage the concerns that religion, another important social identity in India, can be a confounding factor. We do this by examining the robustness of our results in the inclusion of an additional control variable that accounts for religious diversity (religious fractionalisation). We create a religion fractionalisation index following the definition in Alesina et al. (1999):

$$ReligionFractionalization = 1 - \sum_i religion_i^2$$

where $religion_i$ denotes the share of population that believe in religion i. This index measures the probability that two randomly drawn people from a district belong to different religious groups. Higher values of the index indicate greater religious diversity. We construct this index using data from the 2001 population census (which provides information on the population size for the three largest religious groups in each district). As shown in Table S.5 (in Supplementary Materials), our results change minimally after controlling for this religion diversity index. Thus, religion is unlikely to be a confounding factor for our estimates either.

7. Conclusion

We study how the concentration of land ownership, a proxy for landlords' political power, affects the implementation of public works schemes (the NREGA) in the Indian context. Using district-level data on land ownership distribution and NREGA implementation, we find that the concentration of land ownership causes a reduction in the efficacy of public works provision. To address the potential endogeneity issue arising from measurement errors and omitted variables, we apply two main strategies. One is using selection on observables to assess the bias of the OLS estimates (following Altonji et al., 2005; Oster, 2019), and the other is using a historical land tenure institution to instrument for contemporaneous land inequality. Our main findings are unaffected by the use of either of these empirical strategies. Our results are also robust to a variety of checks, such as relaxing the exclusion restriction, using alternative measurements of job provision and land inequality. We find that the results are most likely driven by the political power that the landed elites possess, which enables them to affect the implementation of the NREGA. We rule out demand side factors and the role of social diversity (caste and religion) in driving these effects.

Investigating the relationship between land inequality and public works provision is not only relevant to India but also has policy implications for other developing countries, in Asia and Africa, which have the dual need for job creation and investment in public services (such as road maintenance). Land distribution plays an important role in community development in an agrarian developing country. Recent evidence suggests that the concentration of land ownership could be a hurdle for redistribution, as landlords (better-endowed agents) have sufficient political power to influence redistribution policies such as public education expenditure and other anti-poverty programs (Galor et al., 2009; Ramcharan, 2010). Further, studies find that land distribution legislation are associated with poverty reduction and economic growth (Besley & Burgess, 2000). Therefore, more effective land reforms are needed to improve access to land by the poor, reduce land ownership inequality and achieve the goal of rural community development. More broadly, this paper adds to the discussion on how power asymmetries could hinder policies aimed at promoting prosperity. Future research would provide a more complete understanding of the economic consequences of land inequality and examine how power asymmetry begets economic inequality.

Notes

1. There is a widespread criticism that NREGA does not create productive durable assets that could generate sustained increases in income. According to a Bank (2011), many public works are said to be 'washed away the next monsoon'. The report attributes the inferior quality of public works to the inadequate attention given to the objective of asset creation in NREGA. However, it is difficult to draw a unified picture for the entire country as there is large heterogeneity in the type/nature and quality of assets across different parts of the country. There are no reliable nationally representative datasets that provide information on quality of the assets created under NREGA – hence our analysis cannot provide any empirical evidence on the heterogeneity between landlord and nonlandlord regions in NREGA asset creation.
2. According to the 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. Constructing the measure on owned land holdings would have been preferable. But since more than 97 per cent of holdings are owned and operated by the same individual/family (p. 29, Agriculture Census Report 2005), using operational land holdings is not a major issue.
3. We 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, p. 121). Land operated by institutions constituted less than 0.5 per cent of the total area, hence excluded from the data.
4. http://nregarep2.nic.in/netnrega/dynamic2/dynamicreportnew4.aspx.
5. By taking the logarithm, Phase 2 and 3 districts in 2006 and Phase 3 districts in 2007 are automatically dropped from the regressions. This makes sure that we do not assess the implementation of NREGA *prior to* the rollout of the programme in these districts. The advantage of taking the logarithm lies in the convenience of making comparisons across different specifications.
6. This ratio is large compared to other studies applying the same method (for example, Gehring & Schneider, 2018; Nunn & Wantchekon, 2011).
7. Banerjee and Iyer (2005) explain 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.'
8. We first plot the numbers of land reforms over time in major landlord and non-landlord states in Figure A.6. It provides consistent evidence with the literature that landlord areas enacted more frequent land reforms than non-landlord areas, especially after 1970 (based on data from Besley & Burgess, 2000). To depict since when landlord dominated districts started to have lower land ownership inequality than non-landlord dominated districts, we further plot the trends of land inequality, measured by the share of land owned by the top 10 per cent land holdings, for major landlord and non-landlord districts in Figure A.7. It shows that the shift of landlord districts from having relatively high land inequality to relatively low land inequality occurred around 1970. Therefore, interestingly, the turning point

of relative inequality in landlord versus non-landlord dominated districts coincides with the time when land reforms in major landlord states started to outnumber those in non-landlord states. Put together this further adds credibility to our first- stage estimates.

9. This method has been used in a number of other studies to examine the sensitivity of estimation results to the violations of exogeneity conditions (for example, Azar, Marinescu, & Steinbaum, 2020; Calvi & Mantovanelli, 2018; Ding, Lehrer, Rosenquist, & Audrain-mcgovern, 2009; Nunn & Wantchekon, 2011).
10. The overall reduced-form effect is obtained by estimating the following model: $Y_{ist} = \alpha_0 + \rho Z_{is} + \alpha X_{ist} + \alpha_{st} D_{st} + \eta_{ist}$, where all variables are the same as in Equation (1) and Z_{is} is the indicator for historical landlord districts.
11. For example, 'In the summer of 2011, farmers in East and West Godavari districts even declared a "crop holiday" protesting against low profitability mainly due – according to them – to the increase in rural wages witnessed in recent years' (Maiorano, 2014).

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Yanan Li http://orcid.org/0000-0002-8856-6580

References

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

Alesina, A., Baqir, R., & Easterly, W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, *114*(4), 1243–1284.

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

Altonji, J. G., Elder, T. E., & 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., & Kotwal, A. (2015). Clientelism in Indian villages. *The American Economic Review*, *105*(6), 1780–1816.

Arora, A. (2020). Election by community consensus: Effects on political selection and governance. *The Review of Economics and Statistics*, 1–45. doi:10.1162/rest_a_00970

Azar, J., Marinescu, I., & Steinbaum, M. (2020, May). Labor market concentration. *Journal of Human Resources*, *55*(4), 1218–9914R1. Publisher: University of Wisconsin Press.

Banerjee, A., & 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., & Mookherjee, D. (2005). Decentralizing antipoverty programme delivery in developing countries. *Journal of Public Economics*, *89*(4), 675–704.

Bardhan, P., & Mookherjee, D. (2012). *Political clientelism and capture: Theory and evidence from West Bengal, India* (Number 2012/97). UNU-WIDER Working paper. United Nations University.

Besley, T., & 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., & Rao, V. (2016). Long-run impacts of land regulation: Evidence from tenancy reform in India. *Journal of Development Economics*, *118*, 72–87.

Bhattacharya, A., Kar, A., & Nandi, A. (2016). *Local institutional structure and clientelistic access to employment: the case of MGNREGS in three states of India*. Discussion Papers 16/15, Department of Economics, University of York.

Bose, N., & Das, S. (2015). Political reservation for women and delivery of public works program. *Review of Development Economics, 22*, 203– 219. https://doi.org/10.1111/rode.12336

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

Calvi, R., & Mantovanelli, F. G. (2018, November). Long-term effects of access to health care: Medical missions in colonial India. *Journal of Development Economics*, *135*, 285–303.

Chandramohan, C. (1998). Political economy of agrarian conflicts in India. *Economic and Political Weekly*, *31*(41), 2647–2653.

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

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

Damian, C. (2019). *PLAUSEXOG: Stata module to implement Conley et al's plausibly exogenous bounds*. Statistical Software Components, S457832. Boston College Department of Economics.

Das, U., & Maiorano, D. (2019). Post-clientelistic initiatives in a patronage democracy: The distributive politics of India's MGNREGA. *World Development*, *117*, 239–252.

Ding, W., Lehrer, S. F., Rosenquist, J. N., & 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., & Nilekani, J. (2013). Ethnic quotas and political mobilization: Caste, parties, and distribution in Indian village councils. *American Political Science Review*, *107*(1), 35–56.

Dutta, P., Murgai, R., Ravallion, M., & Van de Walle, D. (2012). Does India's employment guarantee scheme guarantee employment? *Economic and Political Weekly, 47*(16), 55–64. http://www.jstor.org/stable/23214599

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

Faguet, J.-P., Sánchez, F., & Villaveces, M.-J. (2020). The perversion of public land distribution by landed elites: Power, inequality and development in Colombia. *World Development*, *136*, 105036.

Galor, O., Moav, O., & 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.

Gehring, K., & Schneider, S. A. (2018, February). Towards the greater good? EU commissioners' nationality and budget allocation in the European Union. *American Economic Journal: Economic Policy*, *10*(1), 214–239.

Government of India. (2013). *Mahatma Gandhi national rural employment guarantee act, 2005, operational guidelines, 4th edition* (Technical report). New Delhi, India: Department of Rural Development. Ministry of Rural Development.

Gulzar, S., & Pasquale, B. (2016). Politicians, bureaucrats, and development: Evidence from India. *American Political Science Review*, *111*(1), 1–22.

Gupta, B., & 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., & Papp, J. (2015). Labor market effects of social programmes: Evidence from India's employment guarantee. *American Economic Journal: Applied Economics*, *7*(2), 233–263.

Imbert, C. (2012). Decomposing the labor market earnings inequality: The public and private sectors in Vietnam, 1993–2006. *The World Bank Economic Review*, *27*(1), 55–79.

Johnson, D. (2009). *How do caste, gender and party affiliation of locally elected leaders affect implementation of nrega?* Emerging Markets: Regional Perspective eJournal. Emerging Markets: Regional Perspective eJournal. .

Khosla, R. (2011). Caste, politics and public good distribution in India: Evidence from nregs in Andhra Pradesh. *Economic and Political Weekly*, *46*(12), 63–69.

Klonner, S., & Oldiges, C. (2014). *Safety net for India's poor or waste of public funds? Poverty and welfare in the wake of the world's largest job guarantee programme*. Working Papers 0564, University of Heidelberg, Department of Economics.

Kumar, H., & Somanathan, R. (2015). *State and district boundary changes in India: 1961-2001*. Available at SSRN 2687484. http://dx.doi.org/10.2139/ssrn.2687484

Maiorano, D., Das, U., & Masiero, S. (2018). Decentralisation, clientelism and social protection programmes: A study of India's mgnrega. *Oxford Development Studies*, *46*(4), 536–549.

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 programme. *Journal of Development Economics*, *136*, 151–167.

Miguel, E., & Gugerty, M. K. (2005). Ethnic diversity, social sanctions, and public goods in Kenya. *Journal of Public Economics*, *89*(11–12), 2325–2368.

Misra, K. (2019, December). Does historical land inequality attenuate the positive impact of India's employment guarantee programme? *World Development*, *124*, 104662.

Munshi, K., & Rosenzweig, M. (2018). *Insiders and outsiders: local ethnic politics and public goods provision* (No. w21720). National Bureau of Economic Research.

Muralidharan, K., Niehaus, P., & Sukhtankar, S. (2020). *General equilibrium effects of (improving) public employment programmes: Experimental evidence from India*. National Bureau of Economic Research.

Nagarajan, H. K., Binswanger-Mkhize, H. P., & Meenakshisundaram, S. (2014). *Decentralization and empowerment for rural development*. Cambridge, United Kingdom: Cambridge University Press.

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

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

NREGA Operational Guidelines. (2013). *Mahatma Gandhi national rural employment guarantee act 2005 – Operational guidelines, 4th version* (Technical report). New Delhi, India: Ministry of Rural Development, Department of Rural Development.

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

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204.

Persson, T., & Tabellini, G. (1994). Is inequality harmful for growth? *The American Economic Review*, *84*(3), 600–621.

Piketty, T., & Saez, E. (2014). Inequality in the long run. *Science*, *344*(6186), 838–843.

Planning Commission . (2003). *Report of the task force:Identification of districts for wage and self employment programmes*. New Delhi: Planning Commission, Government of India.

Raabe, K., Birner, R., Sekher, M., Gayathridevi, K., Shilpi, A., Schiffer, E. (2010). *How to overcome the governance challenges of implementing nrega* (IFPRI discussion papers series (00963)).

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

Ravallion, M. (2019). *Is a decentralized right-to-work policy feasible?* NBER Working Papers 25687. National Bureau of Economic Research.

Sheahan, M., Liu, Y., Barrett, C. B., & Narayanan, S. (2018). Preferential resource spending under an employment guarantee: The political economy of mgnregs in Andhra Pradesh. *The World Bank Economic Review*, *32*(3), 44.

Sokoloff, K. L., & 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.

Van der Weide, R., & Milanovic, B. (2018). Inequality is bad for growth of the poor (but not for that of the rich). *The World Bank Economic Review*, *32*(3), 507–530.

World, Bank (2011). *Social protection for a changing India: Main Report*. World Bank. https://openknowledge.worldbank.org/handle/10986/2745

Zimmermann, L. (2012). *Labor market impacts of a large-scale public works programme: Evidence from the Indian employment guarantee scheme* (Working Paper). IZA Discussion Papers, No. 6858, Institute for the Study of Labor (IZA), Bonn.

Appendix A. Data and Descriptive Statistics

Other datasets and merging districts across time

In our analysis we use other different pieces of information on districts/states in India. 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. Populations between 2001 and 2010 are filled using these two years' census data, assuming a growth rate equal to that during 1991–2001. Table A3 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, we compute wet season precipitation by aggregating the amount of precipitation between June and November in the year. Soil information is obtained from the Food and Agriculture Organization (FAO) Digital Soil Map of the World and Derived Soil Properties (CDROM). Table A3 shows that 90 per cent of the land contains medium or fine level soil, and the other 10 per cent of land is covered by course soil.

Finally, compiling these data sets from different sources into a district-wise panel is challenging – this is because of changes in district boundaries/definition during the study period. The 2005 Agricultural census contains information for 528 districts in 33 states. The NREGA data portal only contains information for 501 of these districts (in year 2005 districts), which are spread across 30 states. We lose another 30 districts when merging this dataset with the 2001 census (largely due to changes in district boundaries between 2001 and 2005). Further, we merge in information on other covariates (such as soil and rainfall) which results in the loss of 16 districts (as covariates have missing values). Finally, we drop about 41 districts that had boundary changes during 2001–2010 (results are robust to adding these districts back). This gives 414 districts in 28 states (Andaman and Nicobar, Andhra Pradesh, Assam, Bihar, Chhattisgarh, Dadra & Nagar Haveli, Goa, Gujarat, Haryana, Himachal Pradesh, Jammu and Kashmir, Karnataka, Kerala, Madhya Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Odisha, Puducherry, Punjab, Rajasthan, Sikkim, Tamil Nadu, Tripura, Uttar Pradesh, Uttarakhand, West Bengal).

Table A1. Timelines for various steps involved in preparation and finalisation of annual labour 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 scrutinises the Labour Budget and requests for compliance for deficiencies, if any
February	Meetings of Empowered Committee are held and LB finalised
February, March	'Agreed to' LB communicated to States. States feed data of Month-wise and District-wise break-up 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.

Table A2. Summary statistics of NREGA implementation

	(1)	(2)	(3)	(4)	(5)
	2006	2007	2008	2009	2010
% of households provided employment	41.51	34.78	27.33	29.76	31.16
	(23.27)	(25.33)	(26.06)	(24.22)	(23.85)
Avg days of employment provided per rural SC/ST person	6.47	5.80	5.26	6.63	6.49
	(5.03)	(5.82)	(6.61)	(6.94)	(7.13)
Labour expense per rural person (2006 Rs.)	265.95	235.55	212.11	234.22	229.86
	(243.25)	(248.35)	(312.28)	(291.48)	(279.94)
Number of completed works per 1000 rural persons	5.65	9.98	9.53	13.61	7.81
	(7.23)	(23.73)	(26.17)	(20.90)	(13.66)
Number of districts with employment provided	123	210	414	414	414

Source: Author's calculation based on MGNREGA Public Data Portal.
Notes: Labour expense is deflated by state-wise Consumer Price Index, using 2006 as the base year. Only districts in the regression analysis are included. Phase 2 and 3 districts in 2006 and phase 3 districts in 2007 are excluded. Numbers in parentheses are standard deviations.

Table A3. Summary statistics of land inequality and demographic information

	Mean	Standard Deviation
Gini coef.	0.47	(0.07)
Rural area (1000 Sq km)	5.17	(4.69)
Rural population (1000,000)	1.44	(1.03)
Wet season rainfall (100 mm)	1.15	(0.78)
Literacy rate (%)	64.24	(11.76)
% of land covered in fine soil	20.00	(23.96)
% of land covered in medium soil	69.50	(28.13)
% of Agricultural labourers	23.78	(13.16)
% of Main workers	30.65	(5.93)
% of Marginal workers	10.63	(4.28)
% of SC/ST population	33.19	(21.85)
% villages with Safe Drinking water	96.23	(10.43)
% villages with Electricity (Power Supply)	83.28	(19.72)
% villages with Primary school	83.90	(14.23)
% villages with Medical facility	40.59	(25.30)
% villages with Post and telephone facility	50.40	(26.74)
Observations	1575	

Source: Author's calculation based on 2005 Indian Agricultural Census and 2001 Population Census.
Notes: 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 regarded as an agricultural labourer if she/he worked in another person's land for wages in cash, kind or share.

Table A4. Comparison of land distribution in 2005 and 2010

Variables	2005 (mean)	2010 (mean)	Mean Difference
Top 10%	0.359	0.357	−0.002
Top 20%	0.528	0.525	−0.003
Top 30%	0.648	0.645	−0.003
Top 40%	0.739	0.735	−0.004
Bottom 40%	0.136	0.140	0.004
Bottom 30%	0.092	0.095	0.003
Bottom 20%	0.056	0.059	0.002
Bottom 10%	0.026	0.028	0.001
Middle 40–80%	0.204	0.206	0.002
Middle 50–80%	0.134	0.137	0.002
Gini Coefficient	0.466	0.460	−0.006
Number of districts	528	561	

Source: Authors' calculation based on Indian Agricultural Census (2005 and 2010).
Notes: This table tests whether land distribution changed in a statistically significant way between 2005 and 2010 by examining land ownership concentration within certain intervals of the distribution. 'Top 10 per cent' means the share of land by top 10 per cent biggest land holdings. 'Middle 40–80' denotes the share of land owned by the middle 40–80 percentile of land holdings. Last column shows none of the changes is statistically significant at a 10 per cent significance level.

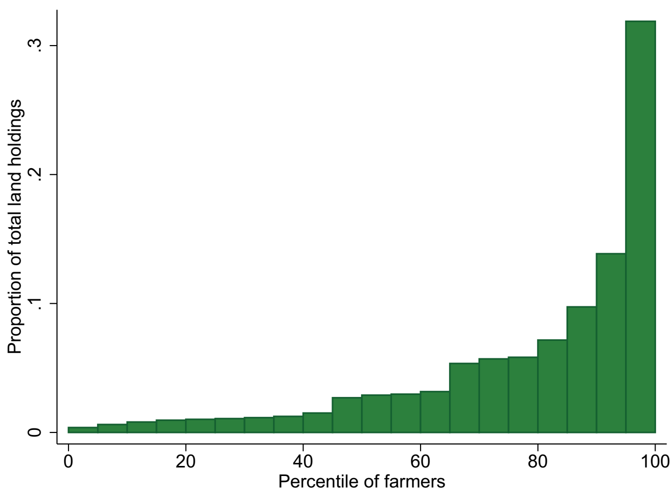

Figure A.1. Shares of land area by percentiles of holdings, 2005.
Source: Author's calculation based on the 2005 Indian Agricultural Census.
Notes: This graph reports the share of land holdings by percentiles of farm ownership. The graph is derived by first ranking all land holdings from the biggest farmers to the smallest farmers in India, then calculating the share of land operated at each five percentile.

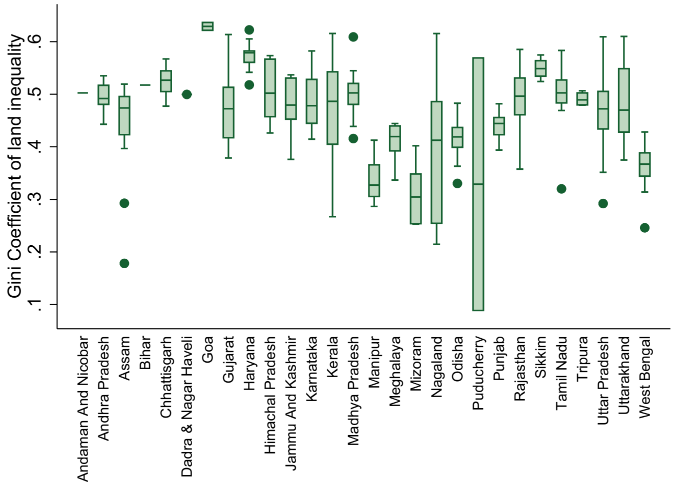

Figure A.2. Land inequality (Gini coefficient) by state, 2005.

Source: Author's calculation based on the 2005 Indian Agricultural Census.

Notes: The author calculated Gini coefficients based on district-wise land distribution. Only states in the OLS regression sample are included.

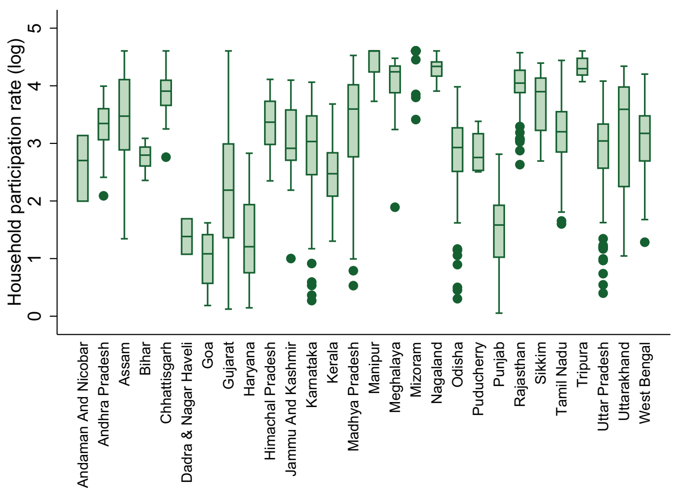

Figure A.3. Shares of households provided with NREGA employment by state, 2006–2010.

Source: Author's calculation based on NREGA Public Data Portal.

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

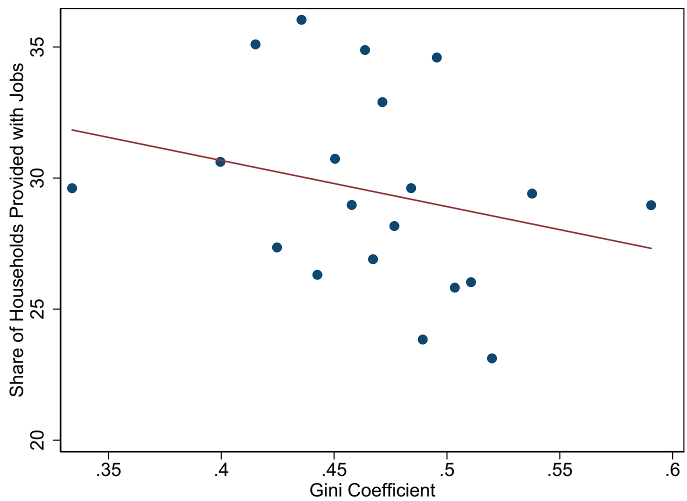

Figure A.4. NREGA job provision and land ownership inequality.
Source: Author's calculation based on NREGA Public Data Portal and 2005 India Agricultural Census.
Notes: This figure is a binned scatter plot of NREGA job provision versus Gini landownership inequality. NREGA implementation is measured by the share of households provided with public jobs in each district. Gini coefficient measures landownership inequality at the district level.

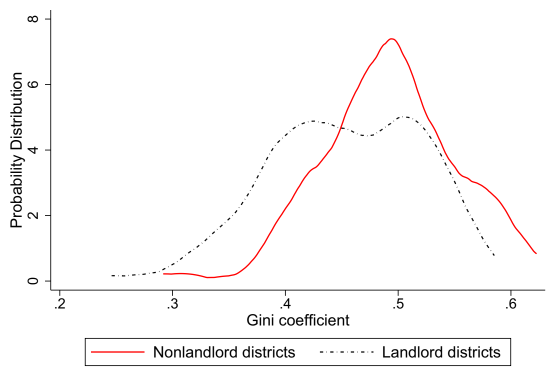

Figure A.5. Visualise first stage – land inequality (Gini coefficient) in landlord/non-landlord districts.
Source: Author's calculation based on the 2005 Indian Agricultural Census.
Notes: This figure plots the probability distribution of Gini coefficients (in 2005) in historically landlord dominated areas v.s. non-landlord dominated areas.

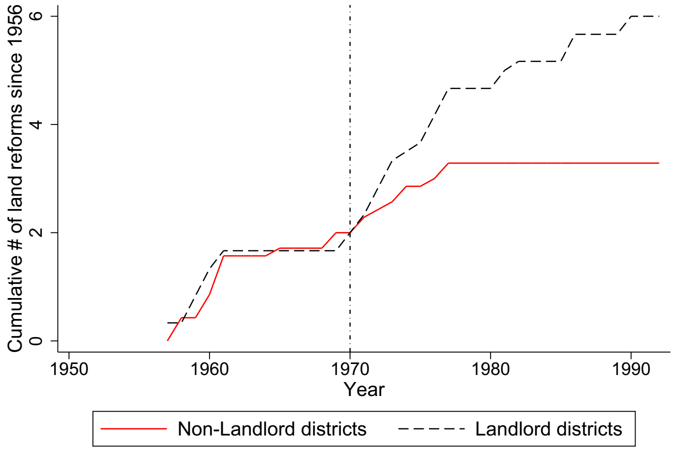

Figure A.6. Frequencies of land reforms in landlord/non-landlord districts.

Source: Information on the cumulative number of land reforms at the state level is extracted from Besley and Burgess (2000).

Notes: Major landlord-dominated areas are states with an above-median share of districts belonging to landlord dominated districts in British India.

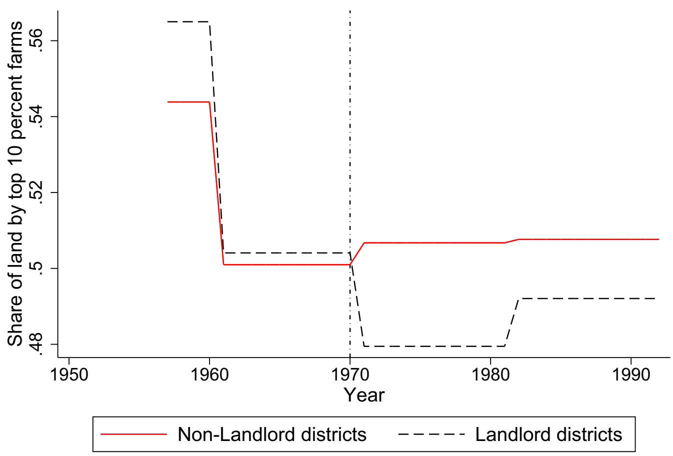

Figure A.7. Trends of land inequality in landlord/non-landlord districts.

Source: Data on land inequaliy is extracted from Besley and Burgess (2000).

Notes: 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 per cent of the biggest farmers.