History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India
by Abhijit Banerjee and Lakshmi Iyer

Shekhar Mittal

Global Economics & Management
UCLA-Anderson

December 3rd, 2014

Summary

Summary

- Land taxes were major source of revenue for the British in India, different institutions for tax collection across districts

Summary

- Land taxes were major source of revenue for the British in India, different institutions for tax collection across districts
- Landlord districts systematically underperform the districts which had a direct mechanism
    - States with a higher proportion of landlord districts have much lower levels of public development expenditures

Summary

- Land taxes were major source of revenue for the British in India, different institutions for tax collection across districts
- Landlord districts systematically underperform the districts which had a direct mechanism
  - States with a higher proportion of landlord districts have much lower levels of public development expenditures
- There is no evidence of landlord districts being at a disadvantage historically

Literature

Literature

- Acemoglu et al. (2001, 2002): Mortality rates among early Europeans strong predictor of ”good institutions”

Literature

- Acemoglu et al. (2001, 2002): Mortality rates among early Europeans strong predictor of ”good institutions”
- Engerman & Sokoloff (1997, 2000, 2002): Difference in Brazil and USA due to suitability of sugar cultivation and resulting slave population

Literature

- Acemoglu et al. (2001, 2002): Mortality rates among early Europeans strong predictor of ”good institutions”
- Engerman & Sokoloff (1997, 2000, 2002): Difference in Brazil and USA due to suitability of sugar cultivation and resulting slave population
- This paper:
  - There are no direct taxes on agricultural incomes in independent India
  - Therefore, a clean example of institutional overhang, underscoring how hard it is to reform the institutional environment

Background

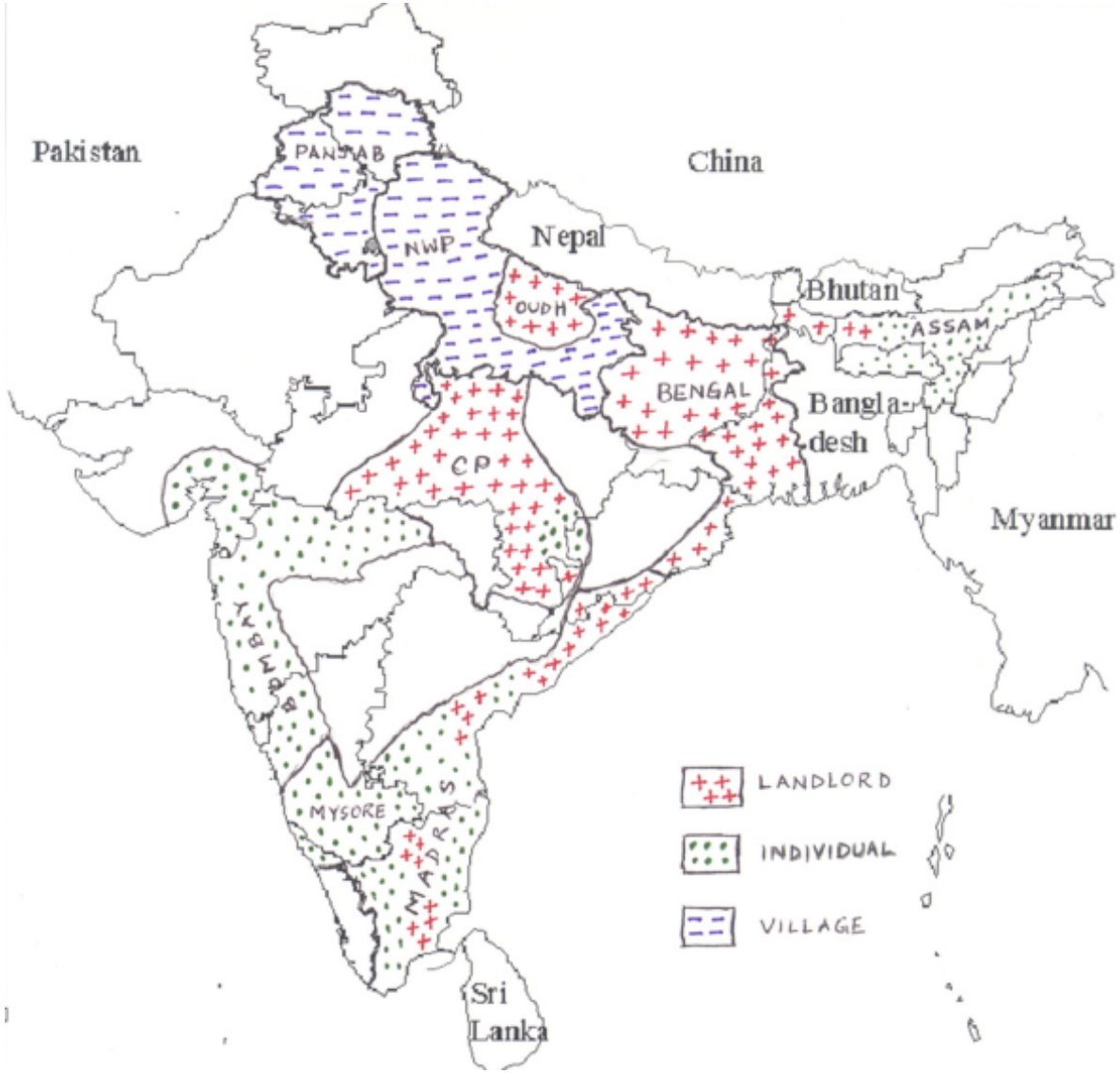


Figure: Map of India

Data

- **District level analysis**
  - Pros: larger sample size, more persistence in boundaries, harder to compute state level measures
  - Cons: Don't have measures of GDP or per-capita income

Data

- **District level analysis**
  - Pros: larger sample size, more persistence in boundaries, harder to compute state level measures
  - Cons: Don't have measures of GDP or per-capita income
- **Outcome variables**
  - Agricultural investment outcomes - irrigated gross cropped area, fertilizer used per hectare, and the proportion of area sown with high-yielding varieties (HYV) of rice, wheat, and other cereals
  - Agricultural productivity - crop yields

Data..

- **India agriculture and climate data set**
  - period 1956-1987, 271 districts in 13 major states at 1961 district level.
  - Source: World Bank

Data..

- **India agriculture and climate data set**
  - period 1956-1987, 271 districts in 13 major states at 1961 district level.
  - Source: World Bank
- **1981 census**
  - Stock of health and education infrastructure

Data..

- **India agriculture and climate data set**
  - period 1956-1987, 271 districts in 13 major states at 1961 district level.
  - Source: World Bank
- **1981 census**
  - Stock of health and education infrastructure
- For each district of British India, construct a measure of non-landlord control - proportion of land area not under landlords.

Question

$$y_{it} = constant + \alpha_t + \beta NL_i + X_{it}\gamma + \epsilon_{it}$$

Question

$$y_{it} = constant + \alpha_t + \beta NL_i + X_{it}\gamma + \epsilon_{it}$$

- Outcome variables are investment, productivity etc. in district i and time t (t is recent)

Question

$$y_{it} = constant + \alpha_t + \beta NL_i + X_{it}\gamma + \epsilon_{it}$$

- Outcome variables are investment, productivity etc. in district i and time t (t is recent)
- Time fixed effects, controls such as latitude, altitude, soil type, rainfall, coastal, length of British rule

Question

$$y_{it} = constant + \alpha_t + \beta NL_i + X_{it}\gamma + \epsilon_{it}$$

- Outcome variables are investment, productivity etc. in district i and time t (t is recent)
- Time fixed effects, controls such as latitude, altitude, soil type, rainfall, coastal, length of British rule
- NL_i is the measure of non landlord control. β is the co-efficient of interest

Question

$$y_{it} = constant + \alpha_t + \beta NL_i + X_{it}\gamma + \epsilon_{it}$$

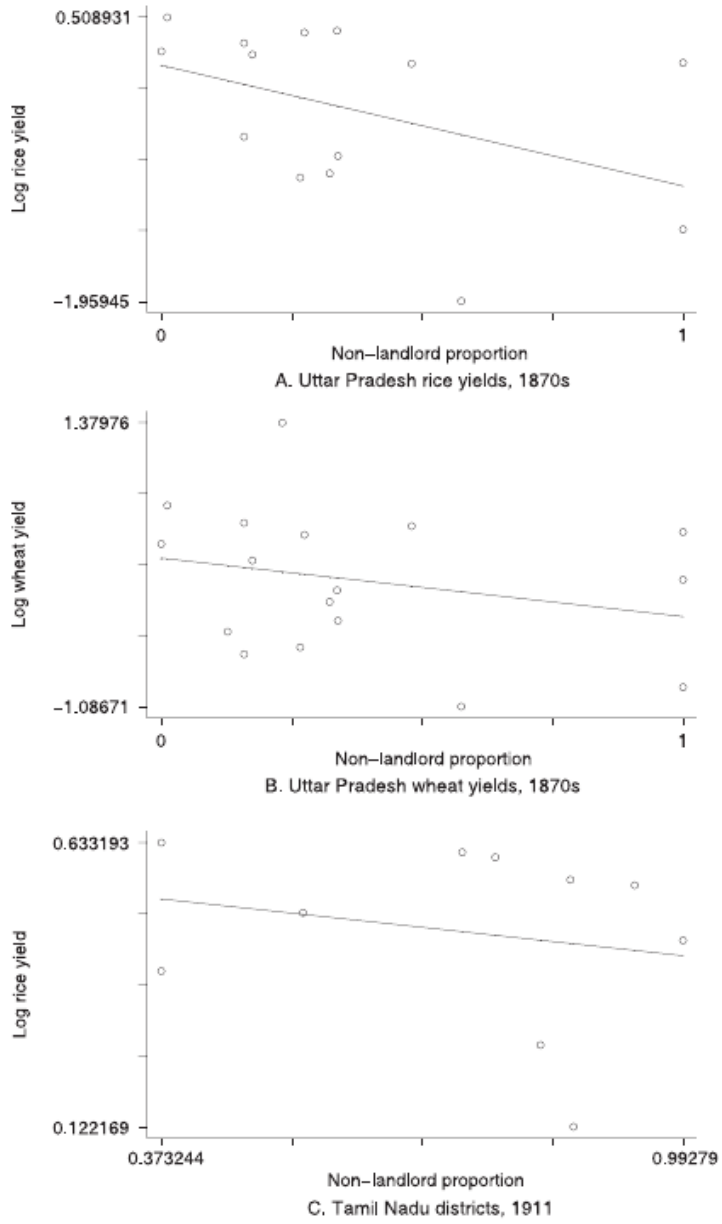
- Outcome variables are investment, productivity etc. in district i and time t (t is recent)
- Time fixed effects, controls such as latitude, altitude, soil type, rainfall, coastal, length of British rule
- NL_i is the measure of non landlord control. β is the co-efficient of interest
- Tackle exogeneity concerns
  - Looking only at the difference between neighboring districts
  - Using IV: areas where the land revenue collection was taken over by the British between 1820 and 1856 (but not before or after) are much more likely to have a non-landlord system

Question

$$y_{it} = constant + \alpha_t + \beta NL_i + X_{it}\gamma + \epsilon_{it}$$

- Outcome variables are investment, productivity etc. in district i and time t (t is recent)
- Time fixed effects, controls such as latitude, altitude, soil type, rainfall, coastal, length of British rule
- NL_i is the measure of non landlord control. β is the co-efficient of interest
- Tackle exogeneity concerns
  - Looking only at the difference between neighboring districts
  - Using IV: areas where the land revenue collection was taken over by the British between 1820 and 1856 (but not before or after) are much more likely to have a non-landlord system
- No state or district fixed effects

Yields in colonial period



Differences in geography and demographics

- (L areas) lower altitudes, higher rainfall

		Mean	Standard deviation	Difference[a]	Standard error of difference
Geography					
Latitude		22.19	5.60	−4.35***	(0.961)
Altitude		366.41	148.14	93.64***	(25.98)
				—	
Mean annual rainfall (mm)		1263.09	471.64	373.99***	(80.83)
Coastal dummy		0.1497	0.3579	0.084	(0.065)
Top 2 soil types	Black soil	0.2096	0.4082	0.244***	(0.072)
	Alluvial soil	0.1677	0.3747	−0.135**	(0.067)
	Red soil	0.5689	0.4967	0.075	(0.090)
Top-soil depth	<25 cm	0.0181	0.1336	0.016	(0.024)
	25–50 cm	0.1145	0.3193	−0.076	(0.058)
	50–100 cm	0.2289	0.4214	0.193	(0.075)
	100–300 cm	0.0904	0.2876	0.135***	(0.051)
	>300 cm	0.5482	0.4991	−0.268***	(0.088)
Area share of various crops: 1956–1987					
Area share of rice		0.366	0.298	−0.194***	(0.054)
Area share of wheat		0.149	0.157	−0.058**	(0.026)
Area share of other cereals		0.205	0.172	0.128***	(0.031)
Area share of oilseeds		0.067	0.088	0.065***	(0.013)
Area share of cotton		0.041	0.096	0.066***	(0.018)
Area share of tobacco		0.003	0.015	0.005**	(0.002)
Area share of sugarcane		0.031	0.053	0.005	(0.008)
Cash crops-to-cereals ratio		0.149	0.257	0.152***	(0.048)
Demographics: 1961, 1971, 1981, 1991					
Log (Population)		14.26	0.634	−0.088	(0.109)
Population density		36.44	85.92	−11.22**	(4.02)
Proportion of scheduled castes		0.1598	0.0733	−0.034**	(0.014)
Proportion of scheduled tribes		0.0980	0.1630	−0.010	(0.031)
Proportion rural		0.8102	0.1237	−0.066***	(0.023)
Proportion of working population in farming		0.7119	0.1352	−0.050*	(0.027)

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. For the area under different crops and demographics, the difference is calculated after controlling for year fixed effects.

[a] Difference represents the average difference between non-landlord and landlord districts, computed as the regression coefficient on the non-landlord proportion.

Differences in geography and demographics

- (L areas) lower altitudes, higher rainfall
- greater proportion of minorities

		Mean	Standard deviation	Difference[a]	Standard error of difference
Geography					
Latitude		22.19	5.60	−4.35***	(0.961)
Altitude		366.41	148.14	93.64***	(25.98)
				—	
Mean annual rainfall (mm)		1263.09	471.64	373.99***	(80.83)
Coastal dummy		0.1497	0.3579	0.084	(0.065)
Top 2 soil types	Black soil	0.2096	0.4082	0.244***	(0.072)
	Alluvial soil	0.1677	0.3747	−0.135**	(0.067)
	Red soil	0.5689	0.4967	0.075	(0.090)
Top-soil depth	<25 cm	0.0181	0.1336	0.016	(0.024)
	25–50 cm	0.1145	0.3193	−0.076	(0.058)
	50–100 cm	0.2289	0.4214	0.193	(0.075)
	100–300 cm	0.0904	0.2876	0.135***	(0.051)
	>300 cm	0.5482	0.4991	−0.268***	(0.088)
Area share of various crops: 1956–1987					
Area share of rice		0.366	0.298	−0.194***	(0.054)
Area share of wheat		0.149	0.157	−0.058**	(0.026)
Area share of other cereals		0.205	0.172	0.128***	(0.031)
Area share of oilseeds		0.067	0.088	0.065***	(0.013)
Area share of cotton		0.041	0.096	0.066***	(0.018)
Area share of tobacco		0.003	0.015	0.005**	(0.002)
Area share of sugarcane		0.031	0.053	0.005	(0.008)
Cash crops-to-cereals ratio		0.149	0.257	0.152***	(0.048)
Demographics: 1961, 1971, 1981, 1991					
Log (Population)		14.26	0.634	−0.088	(0.109)
Population density		36.44	85.92	−11.22**	(4.02)
Proportion of scheduled castes		0.1598	0.0733	−0.034**	(0.014)
Proportion of scheduled tribes		0.0980	0.1630	−0.010	(0.031)
Proportion rural		0.8102	0.1237	−0.066***	(0.023)
Proportion of working population in farming		0.7119	0.1352	−0.050*	(0.027)

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. For the area under different crops and demographics, the difference is calculated after controlling for year fixed effects.

[a] Difference represents the average difference between non-landlord and landlord districts, computed as the regression coefficient on the non-landlord proportion.

Differences in geography and demographics

- (L areas) lower altitudes, higher rainfall
- greater proportion of minorities
- greater share of food crops

		Mean	Standard deviation	Difference[a]	Standard error of difference
Geography					
Latitude		22.19	5.60	−4.35***	(0.961)
Altitude		366.41	148.14	93.64***	(25.98)
				—	
Mean annual rainfall (mm)		1263.09	471.64	373.99***	(80.83)
Coastal dummy		0.1497	0.3579	0.084	(0.065)
Top 2 soil types	Black soil	0.2096	0.4082	0.244***	(0.072)
	Alluvial soil	0.1677	0.3747	−0.135**	(0.067)
	Red soil	0.5689	0.4967	0.075	(0.090)
Top-soil depth	<25 cm	0.0181	0.1336	0.016	(0.024)
	25–50 cm	0.1145	0.3193	−0.076	(0.058)
	50–100 cm	0.2289	0.4214	0.193	(0.075)
	100–300 cm	0.0904	0.2876	0.135***	(0.051)
	>300 cm	0.5482	0.4991	−0.268***	(0.088)
Area share of various crops: 1956–1987					
Area share of rice		0.366	0.298	−0.194***	(0.054)
Area share of wheat		0.149	0.157	−0.058**	(0.026)
Area share of other cereals		0.205	0.172	0.128***	(0.031)
Area share of oilseeds		0.067	0.088	0.065***	(0.013)
Area share of cotton		0.041	0.096	0.066***	(0.018)
Area share of tobacco		0.003	0.015	0.005**	(0.002)
Area share of sugarcane		0.031	0.053	0.005	(0.008)
Cash crops-to-cereals ratio		0.149	0.257	0.152***	(0.048)
Demographics: 1961, 1971, 1981, 1991					
Log (Population)		14.26	0.634	−0.088	(0.109)
Population density		36.44	85.92	−11.22**	(4.02)
Proportion of scheduled castes		0.1598	0.0733	−0.034**	(0.014)
Proportion of scheduled tribes		0.0980	0.1630	−0.010	(0.031)
Proportion rural		0.8102	0.1237	−0.066***	(0.023)
Proportion of working population in farming		0.7119	0.1352	−0.050*	(0.027)

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. For the area under different crops and demographics, the difference is calculated after controlling for year fixed effects.

[a] Difference represents the average difference between non-landlord and landlord districts, computed as the regression coefficient on the non-landlord proportion.

Differences in geography and demographics

- (L areas) lower altitudes, higher rainfall
- greater proportion of minorities
- greater share of food crops
- Significantly higher population density

		Mean	Standard deviation	Difference[a]	Standard error of difference
Geography					
Latitude		22.19	5.60	−4.35***	(0.961)
Altitude		366.41	148.14	93.64***	(25.98)
				—	
Mean annual rainfall (mm)		1263.09	471.64	373.99***	(80.83)
Coastal dummy		0.1497	0.3579	0.084	(0.065)
Top 2 soil types	Black soil	0.2096	0.4082	0.244***	(0.072)
	Alluvial soil	0.1677	0.3747	−0.135**	(0.067)
	Red soil	0.5689	0.4967	0.075	(0.090)
Top-soil depth	<25 cm	0.0181	0.1336	0.016	(0.024)
	25–50 cm	0.1145	0.3193	−0.076	(0.058)
	50–100 cm	0.2289	0.4214	0.193	(0.075)
	100–300 cm	0.0904	0.2876	0.135***	(0.051)
	>300 cm	0.5482	0.4991	−0.268***	(0.088)
Area share of various crops: 1956–1987					
Area share of rice		0.366	0.298	−0.194***	(0.054)
Area share of wheat		0.149	0.157	−0.058**	(0.026)
Area share of other cereals		0.205	0.172	0.128***	(0.031)
Area share of oilseeds		0.067	0.088	0.065***	(0.013)
Area share of cotton		0.041	0.096	0.066***	(0.018)
Area share of tobacco		0.003	0.015	0.005**	(0.002)
Area share of sugarcane		0.031	0.053	0.005	(0.008)
Cash crops-to-cereals ratio		0.149	0.257	0.152***	(0.048)
Demographics: 1961, 1971, 1981, 1991					
Log (Population)		14.26	0.634	−0.088	(0.109)
Population density		36.44	85.92	−11.22**	(4.02)
Proportion of scheduled castes		0.1598	0.0733	−0.034**	(0.014)
Proportion of scheduled tribes		0.0980	0.1630	−0.010	(0.031)
Proportion rural		0.8102	0.1237	−0.066***	(0.023)
Proportion of working population in farming		0.7119	0.1352	−0.050*	(0.027)

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. For the area under different crops and demographics, the difference is calculated after controlling for year fixed effects.

[a] Difference represents the average difference between non-landlord and landlord districts, computed as the regression coefficient on the non-landlord proportion.

Agricultural investments and yields

- Investigate differences in 1956-85 period

(*Mean non-landlord proportion = 0.5051 (s.d. = 0.4274)*)

Dependent variable	Mean of dependent variable	Coefficient on non-landlord proportion: OLS Full sample (1)	Coefficient on non-landlord proportion: OLS Excluding Bengal and Bihar (2)	Coefficient on non-landlord dummy: OLS Full sample (3)	Coefficient on non-landlord dummy: OLS Excluding village-based districts (4)
Agricultural investments					
Proportion of gross cropped area irrigated	0.276	0.065*	0.066*	0.077***	0.005
		(0.034)	(0.035)	(0.027)	(0.032)
Fertilizer use (kg/ha)	24.64	10.708***	10.992***	9.988***	10.695***
		(3.345)	(3.406)	(2.301)	(3.040)
Proportion of rice area under HYV	0.298	0.079*	0.094**	0.016	0.074*
		(0.044)	(0.043)	(0.032)	(0.038)
Proportion of wheat area under HYV	0.518	0.092**	0.119***	0.031	0.107**
		(0.046)	(0.045)	(0.036)	(0.052)
Proportion of other cereals area under HYV	0.196	0.057*	0.084***	−0.035	0.109***
		(0.031)	(0.024)	(0.025)	(0.041)
Agricultural productivity					
log (yield of 15 major crops)		0.157**	0.152**	0.173***	0.089
		(0.071)	(0.074)	(0.053)	(0.085)
log (rice yield)		0.171**	0.195**	0.099	0.173**
		(0.081)	(0.081)	(0.062)	(0.079)
log (wheat yield)		0.229***	0.228***	0.188***	0.143
		(0.067)	(0.070)	(0.054)	(0.098)
No. of districts		166	143	166	109
Year fixed effects		YES	YES	YES	YES
Geographic controls		YES	YES	YES	YES
Date of British land revenue control		YES	YES	YES	YES

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Each cell represents the coefficient from a regression of the dependent variable on the measure of non-landlord control. Data are from 1956 to 1987. Data for area under high-yielding varieties (HYV) is after 1965. Geographic controls are altitude, latitude, mean annual rainfall, and dummies for soil type and coastal regions. The non-landlord dummy is assigned as follows: the dummy equals one for all individual-based districts and all village-based districts except those in Oudh. For landlord-based districts and the village-based districts of Oudh, the dummy is zero.

Agricultural investments and yields

- Investigate differences in 1956-85 period
- Non-landlord districts have a 24% higher proportion of irrigated area & 43% higher levels of fertilizer use

(*Mean non-landlord proportion = 0.5051 (s.d. = 0.4274)*)

Dependent variable	Mean of dependent variable	Coefficient on non-landlord proportion: OLS Full sample (1)	Coefficient on non-landlord proportion: OLS Excluding Bengal and Bihar (2)	Coefficient on non-landlord dummy: OLS Full sample (3)	Coefficient on non-landlord dummy: OLS Excluding village-based districts (4)
Agricultural investments					
Proportion of gross cropped area irrigated	0.276	0.065*	0.066*	0.077***	0.005
		(0.034)	(0.035)	(0.027)	(0.032)
Fertilizer use (kg/ha)	24.64	10.708***	10.992***	9.988***	10.695***
		(3.345)	(3.406)	(2.301)	(3.040)
Proportion of rice area under HYV	0.298	0.079*	0.094**	0.016	0.074*
		(0.044)	(0.043)	(0.032)	(0.038)
Proportion of wheat area under HYV	0.518	0.092**	0.119***	0.031	0.107**
		(0.046)	(0.045)	(0.036)	(0.052)
Proportion of other cereals area under HYV	0.196	0.057*	0.084***	−0.035	0.109***
		(0.031)	(0.024)	(0.025)	(0.041)
Agricultural productivity					
log (yield of 15 major crops)		0.157**	0.152**	0.173***	0.089
		(0.071)	(0.074)	(0.053)	(0.085)
log (rice yield)		0.171**	0.195**	0.099	0.173**
		(0.081)	(0.081)	(0.062)	(0.079)
log (wheat yield)		0.229***	0.228***	0.188***	0.143
		(0.067)	(0.070)	(0.054)	(0.098)
No. of districts		166	143	166	109
Year fixed effects		YES	YES	YES	YES
Geographic controls		YES	YES	YES	YES
Date of British land revenue control		YES	YES	YES	YES

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Each cell represents the coefficient from a regression of the dependent variable on the measure of non-landlord control. Data are from 1956 to 1987. Data for area under high-yielding varieties (HYV) is after 1965. Geographic controls are altitude, latitude, mean annual rainfall, and dummies for soil type and coastal regions. The non-landlord dummy is assigned as follows: the dummy equals one for all individual-based districts and all village-based districts except those in Oudh. For landlord-based districts and the village-based districts of Oudh, the dummy is zero.

Agricultural investments and yields

- Investigate differences in 1956-85 period
- Non-landlord districts have a 24% higher proportion of irrigated area & 43% higher levels of fertilizer use
- Robust to dropping the 2 most feudal states

(*Mean non-landlord proportion = 0.5051 (s.d. = 0.4274)*)

		Coefficient on non-landlord proportion		Coefficient on non-landlord dummy	
Dependent variable	Mean of dependent variable	OLS Full sample (1)	OLS Excluding Bengal and Bihar (2)	OLS Full sample (3)	OLS Excluding village-based districts (4)
Agricultural investments					
Proportion of gross cropped area irrigated	0.276	0.065*	0.066*	0.077***	0.005
		(0.034)	(0.035)	(0.027)	(0.032)
Fertilizer use (kg/ha)	24.64	10.708***	10.992***	9.988***	10.695***
		(3.345)	(3.406)	(2.301)	(3.040)
Proportion of rice area under HYV	0.298	0.079*	0.094**	0.016	0.074*
		(0.044)	(0.043)	(0.032)	(0.038)
Proportion of wheat area under HYV	0.518	0.092**	0.119***	0.031	0.107**
		(0.046)	(0.045)	(0.036)	(0.052)
Proportion of other cereals area under HYV	0.196	0.057*	0.084***	−0.035	0.109***
		(0.031)	(0.024)	(0.025)	(0.041)
Agricultural productivity					
log (yield of 15 major crops)		0.157**	0.152**	0.173***	0.089
		(0.071)	(0.074)	(0.053)	(0.085)
log (rice yield)		0.171**	0.195**	0.099	0.173**
		(0.081)	(0.081)	(0.062)	(0.079)
log (wheat yield)		0.229***	0.228***	0.188***	0.143
		(0.067)	(0.070)	(0.054)	(0.098)
No. of districts		166	143	166	109
Year fixed effects		YES	YES	YES	YES
Geographic controls		YES	YES	YES	YES
Date of British land revenue control		YES	YES	YES	YES

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Each cell represents the coefficient from a regression of the dependent variable on the measure of non-landlord control. Data are from 1956 to 1987. Data for area under high-yielding varieties (HYV) is after 1965. Geographic controls are altitude, latitude, mean annual rainfall, and dummies for soil type and coastal regions. The non-landlord dummy is assigned as follows: the dummy equals one for all individual-based districts and all village-based districts except those in Oudh. For landlord-based districts and the village-based districts of Oudh, the dummy is zero.

Robustness of OLS

- Total yields are 15 percent higher and wheat yields 25 percent higher

Panel A: Robustness checks

	Coefficient on non-landlord proportion	
Dependent variable	OLS Neighbors only (1)	IV Full sample (2)
Agricultural investments		
Proportion of gross cropped area irrigated	0.101**	0.216
	(0.041)	(0.137)
Fertilizer use (kg/ha)	10.589**	26.198**
	(4.979)	(13.244)
Proportion of rice area under HYV	−0.015	0.411**
	(0.083)	(0.163)
Proportion of wheat area under HYV	0.078**	0.584***
	(0.034)	(0.163)
Proportion of other cereals area under HYV	−0.025	0.526***
	(0.024)	(0.129)
Agricultural productivity		
log (yield of 15 major crops)	0.145**	0.409
	(0.061)	(0.261)
log (rice yield)	0.126	0.554*
	(0.098)	(0.285)
log (wheat yield)	0.253***	0.706***
	(0.084)	(0.214)
No. of districts	35	166
Year fixed effects	YES	YES
Geographic controls	YES	YES
Date of British land revenue control	YES	YES

Panel B: First-stage regressions for IV
Dependent variable: Non-landlord proportion

Coefficient on	(1)	(2)	(3)
Instrument (=1 if date of British revenue control is between 1820 and 1856)	0.331***	0.430***	0.419***
	(0.086)	(0.092)	(0.087)
R-squared	0.40	0.43	0.63
No. of observations	166	166	166
Geographic controls	YES	YES	YES
Date of British land revenue control	YES	YES	YES
Date of British land revenue control squared	NO	YES	NO
State fixed effects	NO	NO	YES

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Each cell in Panel A represents the coefficient from a regression of the dependent variable on the non-landlord proportion. Data are from 1956–1987. Data for area under high-yielding varieties (HYV) is after 1965. Geographic controls are altitude, latitude, mean annual rainfall, and dummies for soil type and coastal regions. Instrument is a dummy that equals one if the date of British revenue control is after 1820 and before 1856.

Robustness of OLS

- Total yields are 15 percent higher and wheat yields 25 percent higher
- IV results indicate that OLS results are not biased upward due to ommitted characteristics

Panel A: Robustness checks

Dependent variable	Coefficient on non-landlord proportion: OLS Neighbors only (1)	IV Full sample (2)
Agricultural investments		
Proportion of gross cropped area irrigated	0.101**	0.216
	(0.041)	(0.137)
Fertilizer use (kg/ha)	10.589**	26.198**
	(4.979)	(13.244)
Proportion of rice area under HYV	−0.015	0.411**
	(0.083)	(0.163)
Proportion of wheat area under HYV	0.078**	0.584***
	(0.034)	(0.163)
Proportion of other cereals area under HYV	−0.025	0.526***
	(0.024)	(0.129)
Agricultural productivity		
log (yield of 15 major crops)	0.145**	0.409
	(0.061)	(0.261)
log (rice yield)	0.126	0.554*
	(0.098)	(0.285)
log (wheat yield)	0.253***	0.706***
	(0.084)	(0.214)
No. of districts	35	166
Year fixed effects	YES	YES
Geographic controls	YES	YES
Date of British land revenue control	YES	YES

Panel B: First-stage regressions for IV
Dependent variable: Non-landlord proportion

Coefficient on	(1)	(2)	(3)
Instrument (=1 if date of British revenue control is between 1820 and 1856)	0.331***	0.430***	0.419***
	(0.086)	(0.092)	(0.087)
R-squared	0.40	0.43	0.63
No. of observations	166	166	166
Geographic controls	YES	YES	YES
Date of British land revenue control	YES	YES	YES
Date of British land revenue control squared	NO	YES	NO
State fixed effects	NO	NO	YES

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Each cell in Panel A represents the coefficient from a regression of the dependent variable on the non-landlord proportion. Data are from 1956–1987. Data for area under high-yielding varieties (HYV) is after 1965. Geographic controls are altitude, latitude, mean annual rainfall, and dummies for soil type and coastal regions. Instrument is a dummy that equals one if the date of British revenue control is after 1820 and before 1856.

Are yields explained by investments?

- All measures are positive and stongly significant

	Dependent variables		
	Log total yield OLS (1)	Log rice yield OLS (2)	Log wheat yield OLS (3)
Proportion non-landlord	0.035	0.070	0.109
	(0.053)	(0.063)	(0.063)
Proportion of gross cropped area irrigated	0.693**	0.439**	0.435**
	(0.112)	(0.096)	(0.117)
Fertilizer use (kg/ha)	0.007**	0.004**	0.001
	(0.001)	(0.001)	(0.001)
Percent area under HYV	4.274**	0.580**	0.618**
	(1.122)	(0.063)	(0.070)
Adjusted *R*-squared	0.60	0.52	0.56
No. of districts	166	166	166
Year fixed effects	YES	YES	YES
Geographic controls	YES	YES	YES
Date of British land revenue control	YES	YES	YES

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Data are from 1956–1987. Data for area under high-yielding varieties (HYV) is after 1965. Geographic controls are altitude, latitude, mean annual rainfall, and dummies for soil type and coastal regions.

Are yields explained by investments?

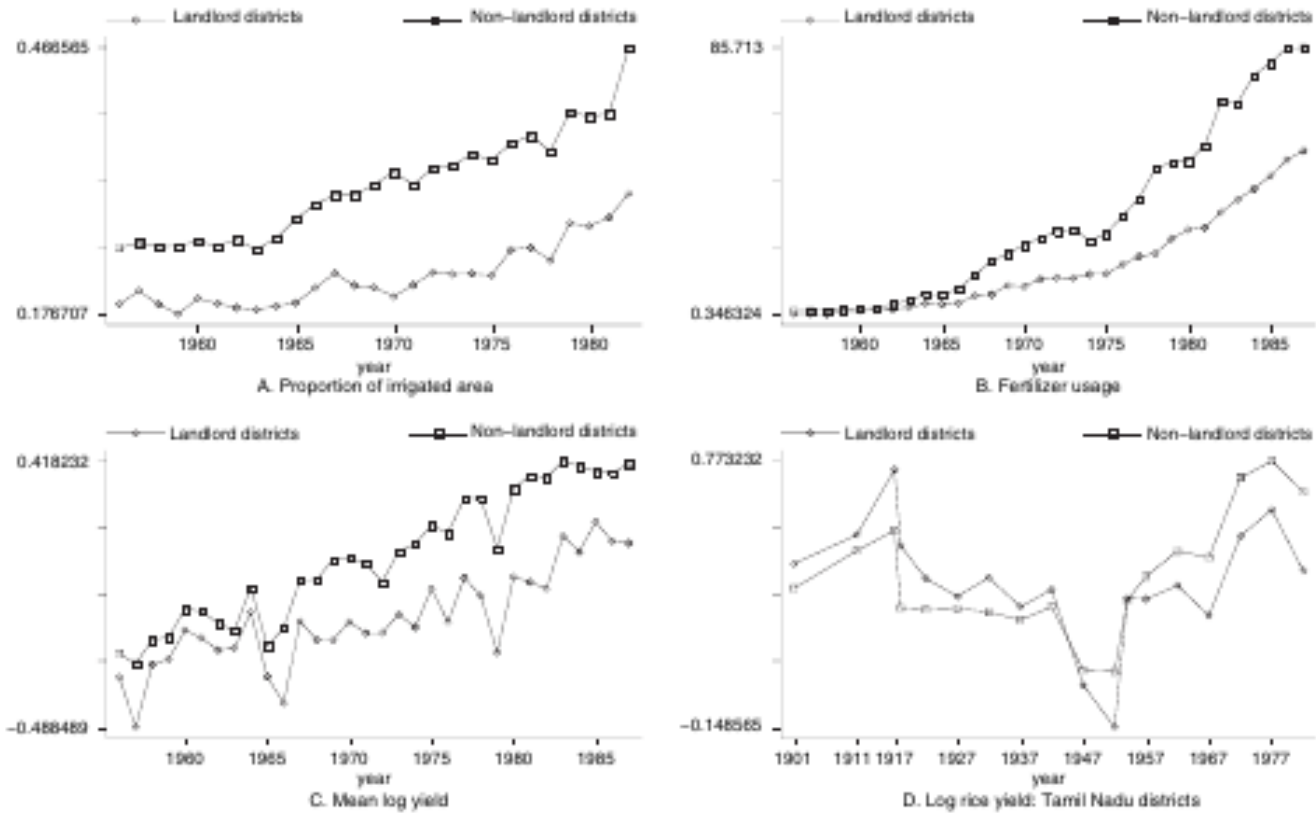
- All measures are positive and stongly significant
- Non-landlord variable is no longer significant

	Dependent variables		
	Log total yield OLS (1)	Log rice yield OLS (2)	Log wheat yield OLS (3)
Proportion non-landlord	0.035	0.070	0.109
	(0.053)	(0.063)	(0.063)
Proportion of gross cropped area irrigated	0.693**	0.439**	0.435**
	(0.112)	(0.096)	(0.117)
Fertilizer use (kg/ha)	0.007**	0.004**	0.001
	(0.001)	(0.001)	(0.001)
Percent area under HYV	4.274**	0.580**	0.618**
	(1.122)	(0.063)	(0.070)
Adjusted *R*-squared	0.60	0.52	0.56
No. of districts	166	166	166
Year fixed effects	YES	YES	YES
Geographic controls	YES	YES	YES
Date of British land revenue control	YES	YES	YES

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Data are from 1956–1987. Data for area under high-yielding varieties (HYV) is after 1965. Geographic controls are altitude, latitude, mean annual rainfall, and dummies for soil type and coastal regions.

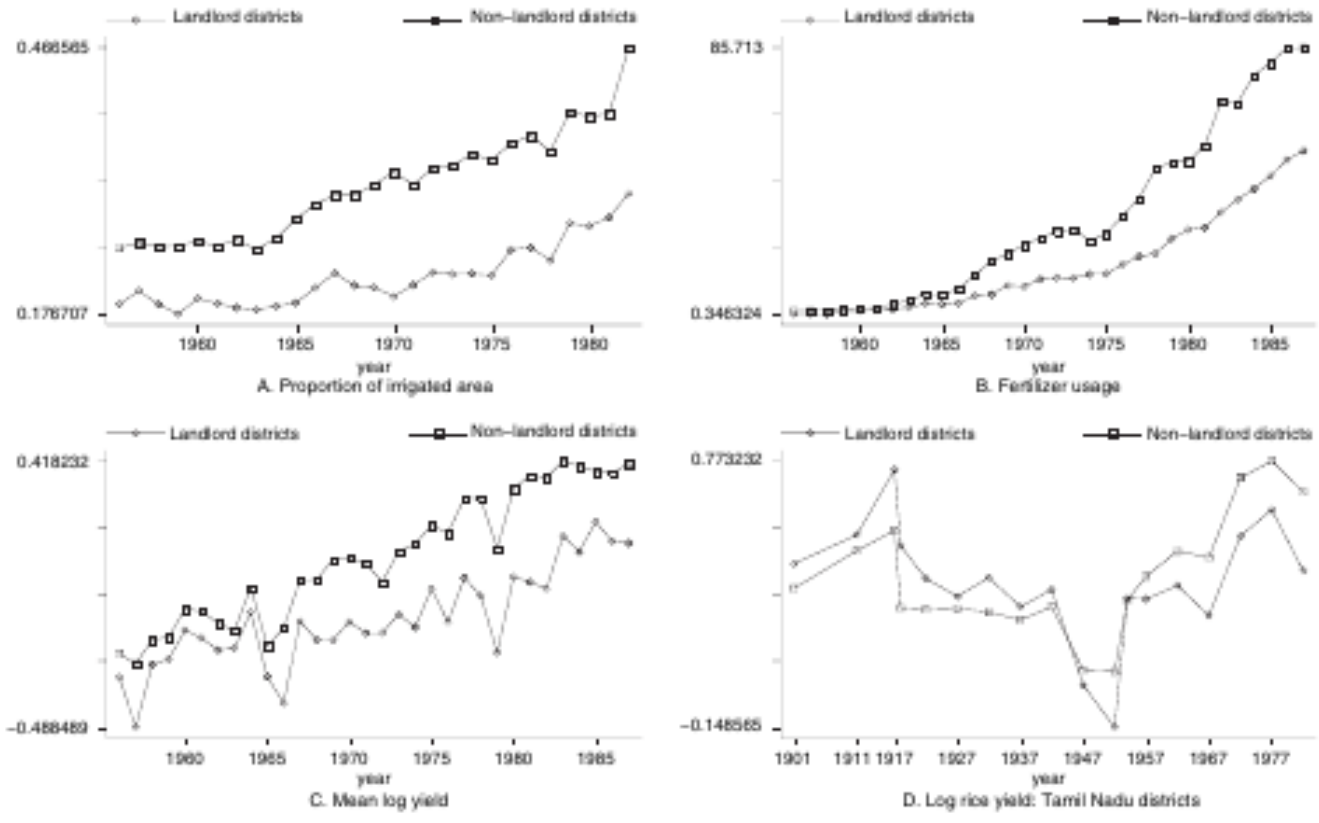
Investment and productivity

- NL districts not more productive than landlord districts in the colonial period



Investment and productivity

- NL districts not more productive than landlord districts in the colonial period
- NL areas overtake the landlord areas during the mid-1960s



When do the differences appear?

Panel A: Full sample

Dependent variable	Coefficient on non-landlord proportion: 1956–1965 (1)	After 1965 (2)	Difference (3)
Agricultural investments			
Proportion of gross cropped area irrigated	0.046	0.079**	0.033**
	(0.033)	(0.036)	(0.016)
Fertilizer use (kg/ha)	1.026**	15.581***	14.55***
	(0.425)	(4.763)	(4.44)
Agricultural productivity			
log (yield of 15 major crops)	0.066	0.201***	0.135***
	(0.065)	(0.076)	(0.033)
log (rice yield)	0.108	0.196**	0.088**
	(0.069)	(0.089)	(0.044)
log (wheat yield)	0.146**	0.268***	0.122*
	(0.058)	(0.079)	(0.063)
No. of districts	166	166	166
Year fixed effects	YES	YES	YES
Geographic controls	YES	YES	YES
Date of British land revenue control	YES	YES	YES

Panel B: Rice yields for Tamil Nadu districts
Sample: 10 districts of Tamil Nadu. Data are for 1870, 1901, 1911, 1917, 1919, and five-yearly intervals from 1922 to 1982.

Dependent variable	Coefficient on non-landlord proportion: Before 1965	After 1965	Difference
Log rice yield	−0.099	0.415	0.514**
	(0.172)	(0.366)	(0.217)
No. of districts	10	10	10
Year fixed effects	YES	YES	YES

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Data are from 1956–1987. Geographic controls are altitude, latitude, mean annual rainfall, and dummies for soil type and coastal regions. Estimates in column (3) are computed from a regression of the dependent variable on the interaction of the non-landlord proportion and a dummy for year >1965, after controlling for the main effects of these variables, as well as geographic controls.

Why do the landlord districts fall behind?

- Landlord areas had fewer villages with primary, middle and high schools

Dependent variables	Mean of dependent variable	Coefficient on non-landlord proportion: OLS Base specification (1)	OLS Control for state dev exp per capita (2)	OLS State FE (3)
Panel A: Agricultural investments				
Proportion of gross cropped area irrigated	0.276	0.065*	0.074**	0.028
		(0.034)	(0.035)	(0.036)
Fertilizer use (kg/ha)	24.64	10.708***	10.805***	4.297
		(3.345)	(3.717)	(3.308)
Proportion of rice area under HYV	0.298	0.079*	0.007	0.000
		(0.044)	(0.040)	(0.042)
Proportion of wheat area under HYV	0.518	0.092**	0.061	0.028
		(0.046)	(0.047)	(0.039)
Proportion of other cereals area under HYV	0.196	0.057*	0.025	0.043*
		(0.031)	(0.030)	(0.026)
Panel B: Agricultural productivity				
log (yield of 15 major crops)		0.157**	0.174**	0.059
		(0.071)	(0.076)	(0.072)
log (rice yield)		0.171**	0.083	0.016
		(0.081)	(0.082)	(0.078)
log (wheat yield)		0.229***	0.243***	0.150***
		(0.067)	(0.072)	(0.045)
Panel C: Education and health investments, 1981				
Proportion of villages having:				
Primary school	0.745	0.154***	0.062*	0.102***
		(0.036)	(0.037)	(0.039)
Middle school	0.204	0.125***	0.093***	0.064***
		(0.023)	(0.021)	(0.018)
High school	0.082	0.052***	0.019	0.030**
		(0.018)	(0.014)	(0.013)
Primary health center	0.023	0.011***	0.002	0.012***
		(0.004)	(0.004)	(0.004)
Primary health subcenter	0.031	0.033***	0.011	0.006
		(0.011)	(0.009)	(0.006)
Panel D: Education and health outcomes				
Literacy rate (1961, 1971, 1981, 1991)	0.2945	0.0524**	0.0290*	0.0241
		(0.0190)	(0.0171)	(0.0176)
Infant mortality rate (1991)	82.17	−32.71***	−25.43***	−15.81***
		(5.38)	(5.28)	(5.40)
State fixed effects		NO	NO	YES
Year fixed effects		YES	YES	YES
Geographic controls		YES	YES	YES
Date of British land revenue control		YES	YES	YES

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Geographic controls are altitude, latitude, mean annual rainfall, and dummies for soil type and coastal regions.

Why do the landlord districts fall behind?

- Landlord areas had fewer villages with primary, middle and high schools
- L areas have lower literacy and higher IMR

Dependent variables	Mean of dependent variable	Coefficient on non-landlord proportion: OLS Base specification (1)	OLS Control for state dev exp per capita (2)	OLS State FE (3)
Panel A: Agricultural investments				
Proportion of gross cropped area irrigated	0.276	0.065*	0.074**	0.028
		(0.034)	(0.035)	(0.036)
Fertilizer use (kg/ha)	24.64	10.708***	10.805***	4.297
		(3.345)	(3.717)	(3.308)
Proportion of rice area under HYV	0.298	0.079*	0.007	0.000
		(0.044)	(0.040)	(0.042)
Proportion of wheat area under HYV	0.518	0.092**	0.061	0.028
		(0.046)	(0.047)	(0.039)
Proportion of other cereals area under HYV	0.196	0.057*	0.025	0.043*
		(0.031)	(0.030)	(0.026)
Panel B: Agricultural productivity				
log (yield of 15 major crops)		0.157**	0.174**	0.059
		(0.071)	(0.076)	(0.072)
log (rice yield)		0.171**	0.083	0.016
		(0.081)	(0.082)	(0.078)
log (wheat yield)		0.229***	0.243***	0.150***
		(0.067)	(0.072)	(0.045)
Panel C: Education and health investments, 1981				
Proportion of villages having:				
Primary school	0.745	0.154***	0.062*	0.102***
		(0.036)	(0.037)	(0.039)
Middle school	0.204	0.125***	0.093***	0.064***
		(0.023)	(0.021)	(0.018)
High school	0.082	0.052***	0.019	0.030**
		(0.018)	(0.014)	(0.013)
Primary health center	0.023	0.011***	0.002	0.012***
		(0.004)	(0.004)	(0.004)
Primary health subcenter	0.031	0.033***	0.011	0.006
		(0.011)	(0.009)	(0.006)
Panel D: Education and health outcomes				
Literacy rate (1961, 1971, 1981, 1991)	0.2945	0.0524**	0.0290*	0.0241
		(0.0190)	(0.0171)	(0.0176)
Infant mortality rate (1991)	82.17	−32.71***	−25.43***	−15.81***
		(5.38)	(5.28)	(5.40)
State fixed effects		NO	NO	YES
Year fixed effects		YES	YES	YES
Geographic controls		YES	YES	YES
Date of British land revenue control		YES	YES	YES

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Geographic controls are altitude, latitude, mean annual rainfall, and dummies for soil type and coastal regions.

Why do the landlord districts fall behind?

- Landlord areas had fewer villages with primary, middle and high schools
- L areas have lower literacy and higher IMR
- Differences due to state development expenditure

Dependent variables	Mean of dependent variable	Coefficient on non-landlord proportion: OLS Base specification (1)	OLS Control for state dev exp per capita (2)	OLS State FE (3)
Panel A: Agricultural investments				
Proportion of gross cropped area irrigated	0.276	0.065*	0.074**	0.028
		(0.034)	(0.035)	(0.036)
Fertilizer use (kg/ha)	24.64	10.708***	10.805***	4.297
		(3.345)	(3.717)	(3.308)
Proportion of rice area under HYV	0.298	0.079*	0.007	0.000
		(0.044)	(0.040)	(0.042)
Proportion of wheat area under HYV	0.518	0.092**	0.061	0.028
		(0.046)	(0.047)	(0.039)
Proportion of other cereals area under HYV	0.196	0.057*	0.025	0.043*
		(0.031)	(0.030)	(0.026)
Panel B: Agricultural productivity				
log (yield of 15 major crops)		0.157**	0.174**	0.059
		(0.071)	(0.076)	(0.072)
log (rice yield)		0.171**	0.083	0.016
		(0.081)	(0.082)	(0.078)
log (wheat yield)		0.229***	0.243***	0.150***
		(0.067)	(0.072)	(0.045)
Panel C: Education and health investments, 1981				
Proportion of villages having:				
Primary school	0.745	0.154***	0.062*	0.102***
		(0.036)	(0.037)	(0.039)
Middle school	0.204	0.125***	0.093***	0.064***
		(0.023)	(0.021)	(0.018)
High school	0.082	0.052***	0.019	0.030**
		(0.018)	(0.014)	(0.013)
Primary health center	0.023	0.011***	0.002	0.012***
		(0.004)	(0.004)	(0.004)
Primary health subcenter	0.031	0.033***	0.011	0.006
		(0.011)	(0.009)	(0.006)
Panel D: Education and health outcomes				
Literacy rate (1961, 1971, 1981, 1991)	0.2945	0.0524**	0.0290*	0.0241
		(0.0190)	(0.0171)	(0.0176)
Infant mortality rate (1991)	82.17	−32.71***	−25.43***	−15.81***
		(5.38)	(5.28)	(5.40)
State fixed effects		NO	NO	YES
Year fixed effects		YES	YES	YES
Geographic controls		YES	YES	YES
Date of British land revenue control		YES	YES	YES

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Geographic controls are altitude, latitude, mean annual rainfall, and dummies for soil type and coastal regions.

Why do the landlord districts fall behind?

- Landlord areas had fewer villages with primary, middle and high schools
- L areas have lower literacy and higher IMR
- Differences due to state development expenditure
- Higher crime rates in L districts (Not in table)

Dependent variables	Mean of dependent variable	Coefficient on non-landlord proportion: OLS Base specification (1)	OLS Control for state dev exp per capita (2)	OLS State FE (3)
Panel A: Agricultural investments				
Proportion of gross cropped area irrigated	0.276	0.065*	0.074**	0.028
		(0.034)	(0.035)	(0.036)
Fertilizer use (kg/ha)	24.64	10.708***	10.805***	4.297
		(3.345)	(3.717)	(3.308)
Proportion of rice area under HYV	0.298	0.079*	0.007	0.000
		(0.044)	(0.040)	(0.042)
Proportion of wheat area under HYV	0.518	0.092**	0.061	0.028
		(0.046)	(0.047)	(0.039)
Proportion of other cereals area under HYV	0.196	0.057*	0.025	0.043*
		(0.031)	(0.030)	(0.026)
Panel B: Agricultural productivity				
log (yield of 15 major crops)		0.157**	0.174**	0.059
		(0.071)	(0.076)	(0.072)
log (rice yield)		0.171**	0.083	0.016
		(0.081)	(0.082)	(0.078)
log (wheat yield)		0.229***	0.243***	0.150***
		(0.067)	(0.072)	(0.045)
Panel C: Education and health investments, 1981				
Proportion of villages having:				
Primary school	0.745	0.154***	0.062*	0.102***
		(0.036)	(0.037)	(0.039)
Middle school	0.204	0.125***	0.093***	0.064***
		(0.023)	(0.021)	(0.018)
High school	0.082	0.052***	0.019	0.030**
		(0.018)	(0.014)	(0.013)
Primary health center	0.023	0.011***	0.002	0.012***
		(0.004)	(0.004)	(0.004)
Primary health subcenter	0.031	0.033***	0.011	0.006
		(0.011)	(0.009)	(0.006)
Panel D: Education and health outcomes				
Literacy rate (1961, 1971, 1981, 1991)	0.2945	0.0524**	0.0290*	0.0241
		(0.0190)	(0.0171)	(0.0176)
Infant mortality rate (1991)	82.17	−32.71***	−25.43***	−15.81***
		(5.38)	(5.28)	(5.40)
State fixed effects		NO	NO	YES
Year fixed effects		YES	YES	YES
Geographic controls		YES	YES	YES
Date of British land revenue control		YES	YES	YES

Notes: Standard errors in parentheses, corrected for district-level clustering. * Significant at 10-percent level; ** significant at 5-percent level; *** significant at 1-percent level. Geographic controls are altitude, latitude, mean annual rainfall, and dummies for soil type and coastal regions.

Conclusions/Concerns

- The concentration of economic and political power continues to be a heavy burden on the economic life of landlord areas

Conclusions/Concerns

- The concentration of economic and political power continues to be a heavy burden on the economic life of landlord areas
- Are the results peculiar to India? Results could be more interesting if they could identify the channels (instead of just hinting them)

Conclusions/Concerns

- The concentration of economic and political power continues to be a heavy burden on the economic life of landlord areas
- Are the results peculiar to India? Results could be more interesting if they could identify the channels (instead of just hinting them)
- Measurement error in NL variable? (IV should take care of it)

Conclusions/Concerns

- The concentration of economic and political power continues to be a heavy burden on the economic life of landlord areas
- Are the results peculiar to India? Results could be more interesting if they could identify the channels (instead of just hinting them)
- Measurement error in NL variable? (IV should take care of it)
- What about socio-economic indicators before mid 1960s?

Conclusions/Concerns

- The concentration of economic and political power continues to be a heavy burden on the economic life of landlord areas
- Are the results peculiar to India? Results could be more interesting if they could identify the channels (instead of just hinting them)
- Measurement error in NL variable? (IV should take care of it)
- What about socio-economic indicators before mid 1960s?
- Interesting fact - that the areas most associated with Maoist peasant uprisings are all landlord areas

Thanks!

Historical background

- Empire building began in 1757 and 1764, and revenue collection rights granted in 1765.
- Different part of the country came under British rule in different periods.
- All cultivable land in British India fell under one of three alternative systems: (a) a landlord-based system (also known as zamindari or malguzari), (b) an individual cultivator-based system (raiyatwari), and (c) a village-based system (mahalwari)

Historical background..

- In some of the areas with the landlord based system, the British declared the landlords' revenue commitments to the government to be fixed in perpetuity (the "Permanent Settlement" of 1793). In other areas, a "temporary" settlement was implemented whereby the revenue was fixed for a certain number of years, after which it was subject to revision.
- Under Raiyatwari system, the revenue commitment was not fixed; it was usually calculated as the money value of a share of the estimated average annual output. This share typically varied from place to place, was different for different soil types, and was adjusted periodically in response to changes in the productivity of the land.
- Choice of land revenue system - influence of individual administrators, political events, date of conquest, maintaining status quo.
- In early 1950s, several states passed legislation towards abolishing landlords, ceiling on land holdings, tenancy reforms etc.

Why should historical land system matter?

- Differences in the distribution of wealth (refer appendix 2).
  - Landlords in a position to appropriate most of the gains in productivity.
  - Due to permanent settlement act 1793, dues fixed permanently in nominal terms.
  - As the nineteenth century was a period of significant productivity growth and inflation, the landlord class grew rich over this period and inequality went up.
- Differences in the political environment
  - Risk of expropriation discouraged investment, created a persistent class system
  - The areas most associated with Maoist peasant uprisings (known as "Naxalite" movements)— clearly the most extreme form of the politics of class conflict in India are West Bengal, Bihar, and the Srikakulam district of Andhra Pradesh, all landlord areas
  - Plausible that, in the post-independence period, the political energies of the masses were directed more toward expropriating from the rich (via land reforms, for example) than toward trying to get more public goods (schools, tap water, electricity) from the state, while the political energies of the rich were aimed at trying to ensure that the poor did not get their way
- Differences in the relationship with the colonial state
  - Since the state could capture some of the productivity gains from non-landlord areas, and hence had more reason to invest in irrigation, railways, schools, and other infrastructure in these areas during the colonial period.
    - Almost all canals constructed by the British were in non-landlord areas.

Appendix: More results

- Post independence, there is substantial convergence in inequality between the landlord and non-landlord areas, probably because states with landlord-dominated areas tend to enact a greater number of land reforms.
- Despite that, the class-based antagonism that it created within the communities in these areas has persisted. Consistent with the fact that the gap between the non-landlord and landlord districts grows particularly fast in the period 1965–1980 when there is extensive public investment in rural areas.

Appendix 1

State	Mean non-landlord proportion	Classification of revenue systems				Total districts
		Landlord based	Individual based	Village bodies		
				Landlord	Non-landlord	
Andhra Pradesh	0.66	2	8	0	0	10
Bihar	0.00	12	0	0	0	12
Gujarat	1.00	0	7	0	0	7
Haryana	0.85	0	0	0	5	5
Karnataka	1.00	0	15	0	0	15
Madhya Pradesh	0.10	14	1	0	0	15
Maharashtra	0.78	4	14	0	0	18
Orissa	0.32	6	2	0	0	8
Punjab	0.87	0	0	0	6	6
Rajasthan	0.00	1	0	0	0	1
Tamil Nadu	0.75	2	9	0	0	11
Uttar Pradesh	0.42	0	0	12	35	47
West Bengal	0.00	11	0	0	0	11
Total	0.51	52	56	12	46	166

Notes: This table lists only districts that used to be part of British India. Areas where the British did not set up the land revenue system are excluded. Districts of British India currently in Pakistan, Bangladesh, or Burma are excluded. The table also excludes the states of Assam and Kerala, for which agricultural data are not available in the World Bank dataset. The table lists 1960 districts, some of which were split into two or more districts over time. We use unsplit districts in all our analyses.

Figure: Distribution of districts

Appendix 2

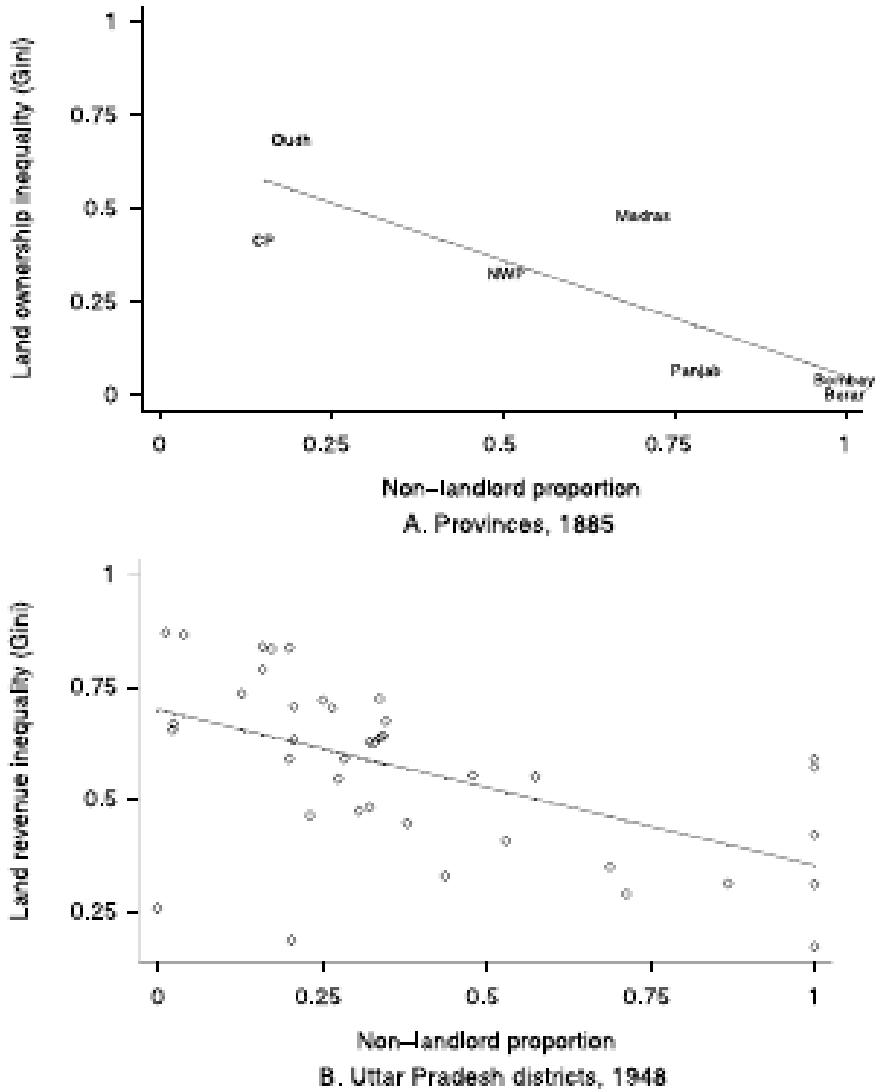


Figure: Land tenure and inequality

Appendix 4

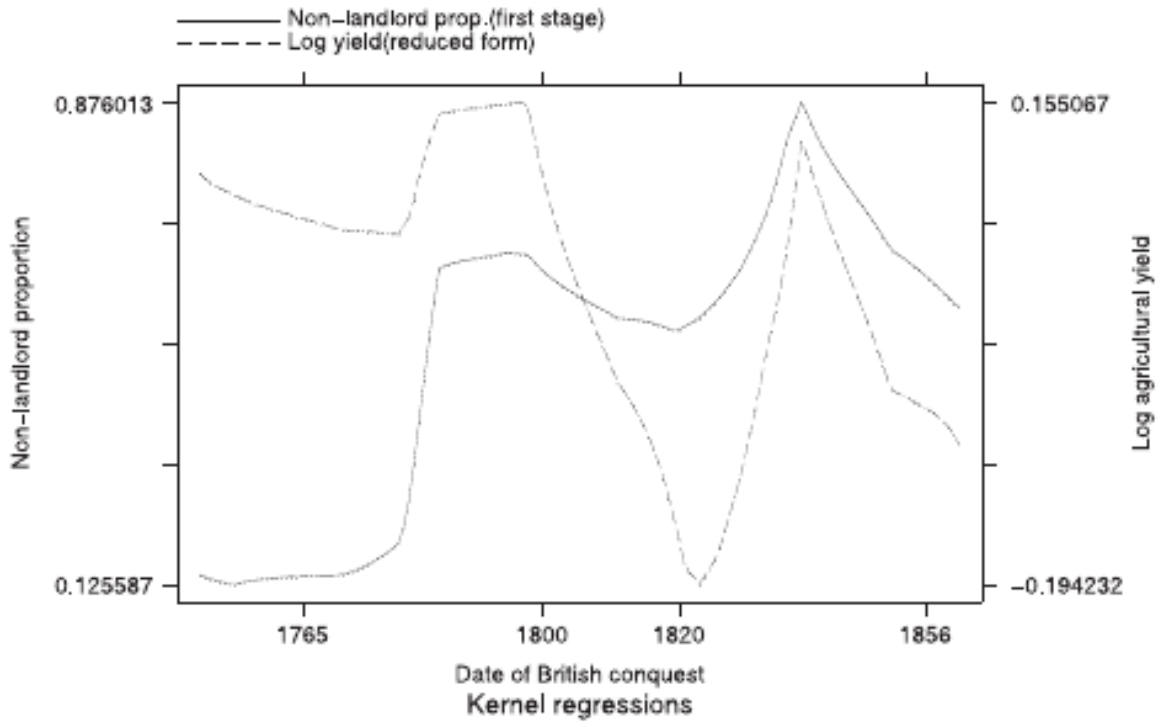


Figure: Intrumental variables strategy

Appendix 5

Figure: Land tenure and date of conquest