Rational Peasants: Revisiting Elite Capture

Cheng CHENG, Flora LI *

May 9, 2019

Abstract

Is failed local accountability the best explanation for the perplexing phenomena of land expropriation and the promoting role of lineage leaders? By replicating the analysis in Mattingly 2016 and conducting extension based on a brief review of land expropriation process in PRC, we argue that lineage leaders who takes village offices promotes the local economy by facilitating the expansion of local enterprises, which results in more land expropriation. We evaluate our argument by presenting the OLS regression results that supports: 1) local enterprises are driving forces for land expropriation; 2) lineage leaders taking office has an incremental effect on the enterprise-expropriation relation; 3) lineage leaders are more devoted to promoting local economy as compared to non-embedded cadres.

Why do Chinese lineage norms served to "encroach" on property rights of its members? According to Mattingly's research on land expropriation and political decentralization, a village cadre who is also a lineage leader of the village can exert coercive power on his villagers when there's need for land expropriation.(Mattingly, 2016) This phenomenon is argued by Mattingly to be a strategy of CPC to achieve social control via informal institutions. (Mattingly, 2016) Are Chinese peasants more vulnerable in face of lineage norms such that they behave myopically and irrationally? Could there be economic reasons that rationalize the compliant behavior of peasants? By taking into consideration of developing local enterprises, we argue that the village cadres who are simultaneously lineage leaders are better able to mobilize land resource in promoting local economy. It is likely because that the embedded cadres have more at stake improving wealth condition of the village than non-embedded. (Oi, 1985)

This paper will proceed by first presenting the replication results and model specification checks of the original models. Then to set up our own argument, we provide a brief review of land expropriation history and unexplained traces in Mattingly 2016 that provoked our extension work. Then we proceed to evaluating our argument by running OLS regression targeted at testing three successive hypothesis: 1) more local enterprises is associated with higher land expropriation

*Cheng CHENG is a first year grad student in Regional Studies: East Asia focusing on PRC domestic politics. (cheng_cheng@g.harvard.edu) Flora LI is a senior at Harvard College studying Applied Mathematics with an application in Government. We are very grateful to the author of the original paper: Professor Daniel Mattingly, whose swift response and generous help made this project possible. We would also like to thank the TF's of Gov 2001 , Professor Gary King, whose persistent support made this project a group effort. We are also grateful to our anonymous peer graders, whose comments pointed us in a new direction.

probability; 2) the embedded local cadre has an incremental effect on the enterprise-expropriation correlation; 3) the embedded cadres are more devoted to managing the local enterprises. Among the three hypothesis, hypothesis 2 speaks directly to our argument, and therefore we conducted model specification tests on the corresponding model as well as performing matching tests. Due to limited space of the paper, we present part of the robustness checks in the appendix.

1 Replication of Mattingly's Work

To verify the results that were presented by Mattingly, we first replicated the figures that were presented in his paper. We were able to accurately replicate every figure in the paper and appendix, meaning that his analysis and results were accurate. The figures are located in Figure 1, and Tables 1-4.

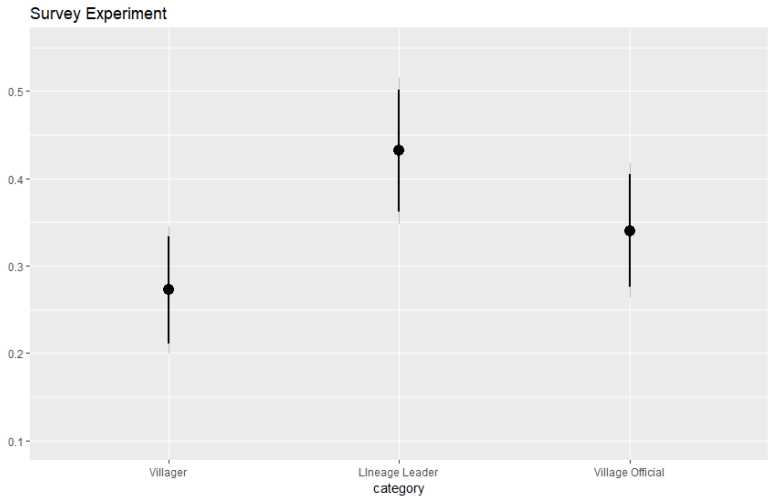

Figure 1: Figure 1 in Mattingly

From the tables above, we can see that the indicator variable for whether or not a lineage leader is a cadre member is significant, which is the basis of Mattingly's argument. Therefore, we can conclude that his conclusions have basis in statistical analysis. However, we will proceed to look deeper into his model and conclusions using our own analysis.

2 Robustness and Model Fit

One aspect that we wanted to extend for Mattingly's work was to evaluate the fit of the model that was used in his analysis. In order to do this, we calculated the robust standard errors of the

Table 1: Table 3 in Mattingly (Ordinary Least Squares Regression: Dependent Variable is Indicator for Land Expropriation)

	Dependent variable:			
	Land expropriations			
	(1)	(2)	(3)	(4)
Lineage leader is cadre	0.164**	0.164**	0.165**	0.173**
	(0.073)	(0.073)	(0.073)	(0.073)
Active lineage	−0.082	−0.068	−0.068	−0.067
	(0.058)	(0.058)	(0.058)	(0.058)
Distance to county seat		−0.0003	0.0005	0.001
		(0.001)	(0.001)	(0.001)
Terrain roughness		−0.00001	−0.00002	−0.00002
		(0.0001)	(0.0001)	(0.0001)
Agricultural suitability index		0.032	0.031	0.034
		(0.024)	(0.024)	(0.024)
Wealth (nighttime lights proxy)		0.003	0.003	0.002
		(0.003)	(0.003)	(0.003)
Township control over elections			0.111	0.137
			(0.089)	(0.090)
Distance to township			−0.009**	−0.009**
			(0.004)	(0.004)
Surname fragmentation index				0.159*
				(0.085)
Ethnic fragmentation index				−0.249
				(0.184)
Number of households (log)				0.019
				(0.028)
Constant	0.149***	0.208***	0.212***	−0.013
	(0.022)	(0.063)	(0.065)	(0.177)
Province Fixed Effects	no	yes	yes	yes
Observations	359	359	359	359
R^2	0.014	0.045	0.064	0.082
Adjusted R^2	0.009	0.026	0.040	0.050
Residual Std. Error	0.356 (df = 356)	0.353 (df = 351)	0.351 (df = 349)	0.349 (df = 346)
F Statistic	2.566* (df = 2; 356)	2.388** (df = 7; 351)	2.642*** (df = 9; 349)	2.565*** (df = 12; 346)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 2: Table 4 in Mattingly (Strength of Lineage Institutions)

	Dependent variable:			
	Land expropriations occurring in the village			
	(1)	(2)	(3)	(4)
Lineage leader is cadre	0.216***	0.217**	0.209**	0.193**
	(0.083)	(0.085)	(0.085)	(0.086)
Active lineage with ancestral hall or zupu	−0.103*	−0.058	−0.052	−0.049
	(0.059)	(0.065)	(0.064)	(0.064)
Economic controls	no	yes	yes	yes
Political controls	no	no	yes	yes
Social controls	no	no	no	yes
Province fixed effects	no	yes	yes	yes
Observations	392	390	390	376
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3: Table 5 in Mattingly (Placebo Tests)

	Dependent variable:			
	Land expropriations before the current leadership			
	(1)	(2)	(3)	(4)
Lineage leader is cadre	0.009	−0.003	−0.006	−0.005
	(0.042)	(0.041)	(0.042)	(0.043)
Active lineage	0.018	0.041	0.044	0.045
	(0.033)	(0.034)	(0.034)	(0.035)
Economic controls	no	yes	yes	yes
Political controls	no	no	yes	yes
Social controls	no	no	no	yes
Province fixed effects	no	yes	yes	yes
Observations	392	390	390	376
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 4: Table 6 in Mattingly (Petitioning and Land Expropriations)

	Dependent variable:			
	Village participation in collective petitioning			
	(1)	(2)	(3)	(4)
Lineage leader is cadre	0.000	−0.011	−0.017	−0.016
	(0.051)	(0.055)	(0.055)	(0.055)
Land seizure	−0.046	−0.082	−0.088*	−0.088
	(0.050)	(0.053)	(0.053)	(0.055)
Lineage leader × Land seizure	0.319***	0.362***	0.378***	0.352***
	(0.111)	(0.114)	(0.114)	(0.118)
Constant	0.111***	−0.021	−0.025	−0.489**
	(0.019)	(0.084)	(0.086)	(0.204)
Observations	392	390	390	376
Economic controls	no	yes	yes	yes
Political controls	no	no	yes	yes
Social controls	no	no	no	yes
Province fixed effects	no	yes	yes	yes
Note:				*p<0.1; **p<0.05; ***p<0.01

model as well as cluster-robust standard errors that were clustered on province. We could not cluster based on village as the data was collected at a village level. The main test, which can be found in Table 5, was conducted on the OLS model specified in Table 1 of this paper. All other calculations on the other models can be found in the appendix under Tables 18-20.

Table 5: Standard Errors, Robust Standard Errors, and Cluster-Robust Standard Errors (at the province level) for the original OLS model by Mattingly

	Dependent variable:		
	Land expropriations		
	(SE)	(RSE)	(CRSE)
Lineage leader is cadre	0.141^{**}	0.141^{*}	0.141^{*}
	(0.072)	(0.073)	(0.083)
Active lineage	-0.034	-0.034	-0.034
	(0.058)	(0.051)	(0.040)
Distance to county seat	0.0004	0.0004	0.0004
	(0.001)	(0.001)	(0.001)
Terrain roughness	0.0001	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)
Agricultural suitability index	0.035	0.035	0.035
	(0.030)	(0.047)	(0.064)
Wealth (nighttime lights proxy)	0.003	0.003	0.003
	(0.003)	(0.004)	(0.004)
Township control over elections	0.108	0.108	0.108^{*}
	(0.097)	(0.104)	(0.060)
Distance to township	-0.010^{**}	-0.010^{***}	-0.010^{**}
	(0.004)	(0.003)	(0.004)
Surname fragmentation index	0.159^{*}	0.159	0.159
	(0.090)	(0.111)	(0.130)
Ethnic fragmentation index	-0.134	-0.134	-0.134
	(0.193)	(0.147)	(0.179)
Number of households (log)	0.039	0.039	0.039
	(0.034)	(0.036)	(0.031)
Constant	-0.287	-0.287	-0.287
	(0.223)	(0.244)	(0.230)
Province Fixed Effects	yes	yes	yes
Observations	376	376	376
R^2	0.141	0.141	0.141
Adjusted R^2	0.058	0.058	0.058
Residual Std. Error (df = 342)	0.341	0.341	0.341
F Statistic (df = 33; 342)	1.701^{**}	1.701^{**}	1.701^{**}
Note:			*p<0.1; **p<0.05; ***p<0.01

As shown in the table, the standards error were generally rather similar, suggesting that the OLS model that Mattingly use fit the data well and that his results were accurate. However, it is interesting to note that the indicator for the lineage leader being a cadre member was only significant at the $p < 0.1$ level when calculating the model with robust or cluster-robust standard errors. This could possibly suggest that there are other forces at play besides lineage leaders that could lead to increased land expropriation. We present possible alternatives later on in this

paper. Another very significant covariate is the distance to township, which is appropriate in this situation, as land further away from the township would be seen as less desirable, and therefore would have a negative coefficient.

Another check for model fit that we conducted was to create a logit model using the same covariates as the original model, then calculate the robust and cluster-robust standard errors for that model. The results are shown in Table 6.

Table 6: Standard Errors, Robust Standard Errors, and Cluster-Robust Standard Errors (at the province level) for a logit model using the same covariates as the OLS model by Mattingly

	Dependent variable:		
	Land expropriations		
	(SE)	(RSE)	(CRSE)
Lineage leader is cadre	1.374^{*}	1.374	1.374
	(0.777)	(0.842)	(0.947)
Active lineage	-0.445	-0.445	-0.445
	(0.701)	(0.748)	(0.632)
Distance to county seat	0.005	0.005	0.005
	(0.010)	(0.012)	(0.011)
Terrain roughness	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)
Agricultural suitability index	0.253	0.253	0.253
	(0.242)	(0.332)	(0.463)
Wealth (nighttime lights proxy)	0.024	0.024	0.024
	(0.023)	(0.029)	(0.026)
Township control over elections	0.995	0.995	0.995^{*}
	(0.870)	(0.969)	(0.549)
Distance to township	-0.142^{***}	-0.142^{**}	-0.142^{**}
	(0.052)	(0.056)	(0.059)
Surname fragmentation index	1.331	1.331	1.331
	(0.838)	(1.099)	(1.110)
Ethnic fragmentation index	-1.553	-1.553	-1.553
	(2.424)	(2.541)	(3.108)
Number of households (log)	0.414	0.414	0.414
	(0.323)	(0.362)	(0.325)
Constant	-5.924^{***}	-5.924^{**}	-5.924^{**}
	(2.200)	(2.633)	(2.316)
Province Fixed Effects	yes	yes	yes
Observations	376	376	376
Log Likelihood	-123.497	-123.497	-123.497
Akaike Inf. Crit.	314.994	314.994	314.994
Note:			*p<0.1; **p<0.05; ***p<0.01

Again, while the standard errors are rather similar, suggesting a decent model fit, the indicator variable for a lineage leader being a cadre member is not significant at all for robust and cluster-robust standard errors. This further supports our hypothesis that this may possibly not be the only significant factor at play in the increase of land expropriations. In addition, the surname

fragmentation index ceases to be significant at any level with this model, suggesting that it has little to no real significant effect on whether the land is seized by the village leadership or not.

We also attempted to match on the data to check for robustness. Mattingly used propensity scoring to match his data, which is not very reliable. We instead use Mahalanobis distance to match the data. The results before and after matching, as well as the Average Treatment Effect estimate are shown in Tables 7-9.

Table 7: Balance before matching the data

	mean.Tr	mean.Co	sdiff	sdiff.pooled	var.ratio	T pval
clan.any.max	1.00	0.15	Inf	340.68	0.00	0.00
county.distance	22.17	31.01	-43.32	-41.35	0.84	0.01
terrain	253.65	224.49	11.65	11.62	0.99	0.44
soil.quality	1.42	1.64	-36.65	-30.47	0.53	0.02
night.lights	3.50	3.85	-5.34	-5.18	0.89	0.73
election.township.mean	0.16	0.15	3.61	3.41	0.80	0.81
township.distance	5.21	5.78	-11.96	-11.18	0.78	0.44
surname.index	0.68	0.72	-19.21	-19.61	1.09	0.20
ethnic.index	0.03	0.04	-0.44	-0.43	0.97	0.98
log.hh	6.02	6.02	-0.74	-0.76	1.09	0.96
province	13.12	12.38	10.54	11.45	1.44	0.47

Table 8: Balance after matching the data

	mean.Tr	mean.Co	sdiff	sdiff.pooled	var.ratio	T pval
clan.any.max	1.00	0.15	Inf	340.68	0.00	0.00
county.distance	22.17	31.01	-43.32	-41.35	0.84	0.01
terrain	253.65	224.49	11.65	11.62	0.99	0.44
soil.quality	1.42	1.64	-36.65	-30.47	0.53	0.02
night.lights	3.50	3.85	-5.34	-5.18	0.89	0.73
election.township.mean	0.16	0.15	3.61	3.41	0.80	0.81
township.distance	5.21	5.78	-11.96	-11.18	0.78	0.44
surname.index	0.68	0.72	-19.21	-19.61	1.09	0.20
ethnic.index	0.03	0.04	-0.44	-0.43	0.97	0.98
log.hh	6.02	6.02	-0.74	-0.76	1.09	0.96
province	13.12	12.38	10.54	11.45	1.44	0.47

Table 9: ATT after balancing

Estimate	0.21154
AI SE	0.085808
T-stat	2.4653
p.val	0.013691
Original number of observations	359
Original number of treated obs	52
Matched number of observations	52
Matched number of observations (unweighted).	52

From the tables, we can see that the treatment effect is indeed significant and is positive, showing that the treatment variable, which is the indicator variable for the lineage leader being a cadre member, leads to the land being more likely to be expropriated. It is also important to note that the matching did virtually nothing to make the data more balanced, meaning that the data was already very balanced originally.

To further test for robustness, we calculated the frontier plot to check what the ideal number of observations being pruned would be. While our personal intuition is that the ideal number of observations to be pruned would be around 150-200, looking at the frontier and estimated effect plots. However, the matching pruned far more than that, around 300, which we can see on the estimated effect plot leads to a very accurate estimate, most likely due to the one-to-one matching of observed to treated variable that occurs. However, this large of a pruning would lead to much larger variation in the data.

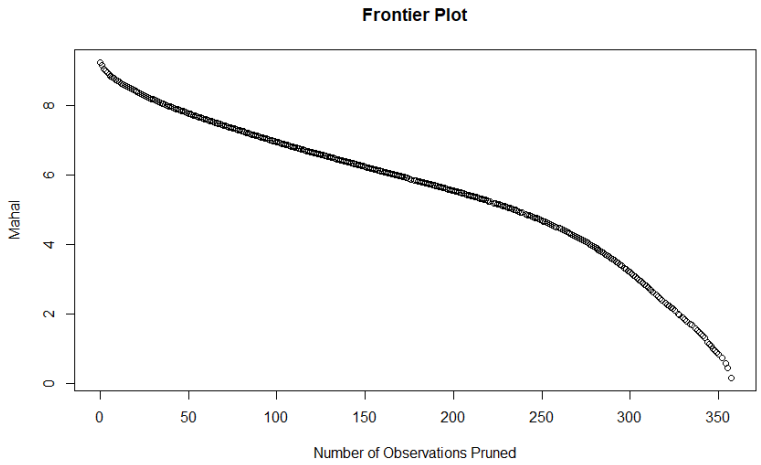

Figure 2: Frontier Plot for Mattingly Data

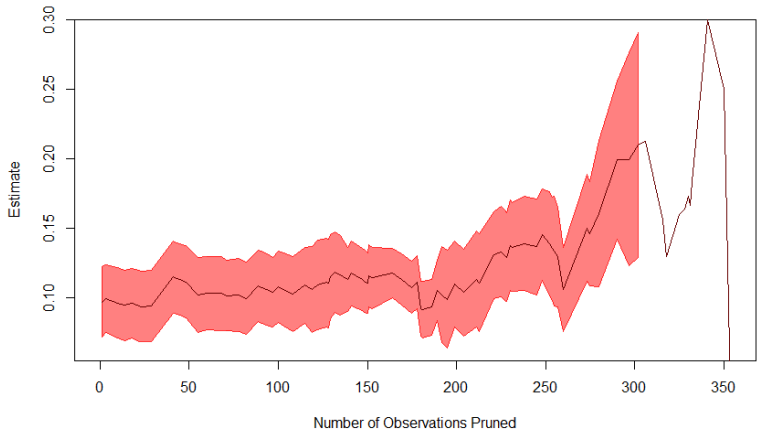

Figure 3: Estimated Effects Plot for Model Dependence on Mattingly Data

We proceeded to generate a new dataset with around 175 observations pruned, but the new model created with that dataset did not differ significantly from the original model. As such, we

have placed it in the appendix under Table 17.

Extension

3 Background and Hypothesis: Economic Rationale of Land expropriation

The extension section is driven by challenging the author's argument on lineage elites eliciting compliance from local population and extracting spoil of their power – which is potentially insufficient to explain the persistence of "exploitation" throughout the examined time period. (Mattingly, 2016) By providing a brief review of the two waves of "land contracting," which drove most of the land expropriation in rural China during the time period of interest, we propose a new perspective in interpreting the role of embedded village cadres: promoting local industrial enterprises. We will also test our hypothesis OLS model, and conduct robustness checks using robustness standard error, the result for matching is also included in the appendix. We found that the number of local enterprises is a crucial factor driving land expropriation, and embedded village cadres exacerbates the effect of enterprise on land expropriation, possibly due to economic consideration. This finding poses challenge to Mattingly's theory by introducing the rational calculation into village cadre election and social governance in China, which points to an explanation of persistence of the embedded cadres. (Mattingly, 2016)

Why do villagers in 2015 still choose to trust their local cadres when they suffer from land expropriation due to their overconfidence in embedded cadres in the past decade? Can lineage norms override economic calculations, especially when the author himself acknowledges petition surges when land expropriation occurs? The underlying assumption that peasants could be myopic, lacking information and forgetful as opposed to their lineage leaders and local cadres are partially true, but it does not necessarily apply to the case of land expropriation. As a result of the Household Registration System, holding land in rural area is a crucial right generated by rural household status, which is associated with receptions of social welfare of all aspects. (Oi, 1992) Therefore, land serves as an insurance of welfare and social relief when other income resources become unavailable to a rural household, which makes peasants exceptionally risk averse when facing land expropriation. Village cadres, according to 1982 Constitution, are elected by villagers and face fixed terms as regulated by Organic Law of Village of P.R.C enacted in 1998. (Congress, 1982) (Committee, 1998) Therefore, a village cadre not only faces competition when he runs for office, but could also be punished for failing in reelection when he harms the villagers' long-term

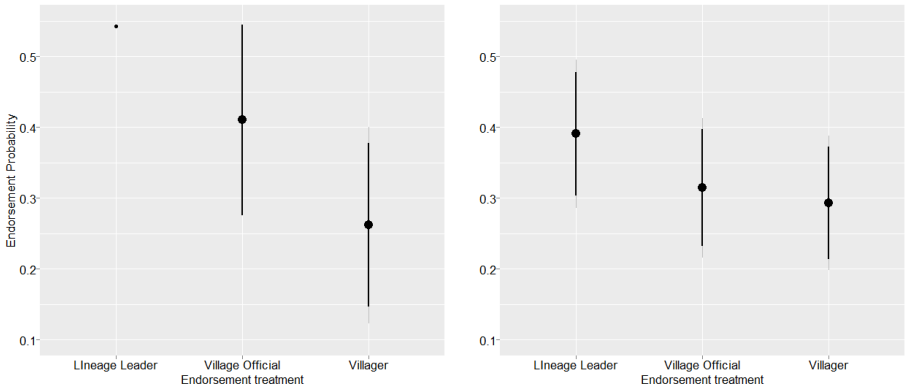

Figure 4: Survey experiment results for villagers with lineage elites being cadres vs. without.

benefit. What's more, the embedded cadres not only receive trust from his relatives, but could also be held more accountable due to internal pressure within the family. (Oi, 1985) How can a lineage cadre survive in the village after the negative effects of land expropriation, if truly undesirable as Mattingly argued, become salient as time goes by?

What confused the picture even more is the survey experiment conducted in 2015 by the author himself. A decade after the CGSS survey was conducted and theoretically all harms for long-term benefit should have exposed, the villagers still showed trust for their lineage cadres (see Figure 4). (Mattingly, 2016) In fact, if we subset the survey experiment into villagers who have a lineage leader as the local cadre as opposed who do not, those who have lineage leaders as village cadres trust their cadres more, while those who do not have lineage cadres trust the cadres as little as they do for generic villagers. Mattingly mentioned this phenomenon as a result of "elite-led persuasion", which enabled the elites maintain control over the local society while continue to reap spoil from holding office.(Mattingly, 2016)

But why would the spells of elites persist despite the fact that the peasants were "deprived" of their major source of income? Why do the villagers still choose to trust their lineage cadre when most lineage cadre, according to Mattingly 2016, takes advantage of their trust and cause more land expropriation to his village? (Mattingly, 2016) Are the villagers as deceitful but also forgetful as Mattingly argued? Is there no economic rationale behind land expropriation at household level at all? We have no measure for the villagers' intelligence level, nor is it necessary to assume the influence of lineage norms can override personal economic motivations – there's no reason to believe the assumption of rational individual deviates from the reality of Chinese peasants more than the rest of the world. A possibly fruitful path is to reflect on the time period of interest and ask: what could be possibly driving the land expropriation process? Why are the land appropriated? Who

Table 10: Reasons for Land Expropriation

Reason for expropriation	Count
Government expropriation	102
Collective Enterprise	91
Second-round Land Contracting	923
No Land expropriation	1080

(Yuan, 1998)

benefits from land expropriation?

The responses in CGSS 2005 rural households indicate that most of land expropriation were carried out under the name of "second-round land contracting" (*di'erlun tudi chengbao*), as shown in Table 10. [1] With the aim of developing collective economy and promoting rural enterprises, the second-round land contracting acknowledges the right of legal transfer of land for economic purposed among the villagers, which served to enable land aggregation for increasing farming scale and also developing local enterprises by introducing market economy in land resource distribution. (CPCDI, 2017) Similar concerns for developing collective enterprises were also displayed in the first-round land contracting. Yet without the acknowledgement of right of transfer, aggregation was carried out by direct expropriation by the village collective. For example, Fu'an village (*fu'an cun*) devoted 90.6 out of 131.6 acres appropriated land to developing collective enterprise that employed local villagers and provided for public goods, which is a common phenomenon in the late 80s and early 90s. (Yuan, 1998) Therefore, the development of local enterprises that provides alternative sources of income is one of the driving forces that's motivating land expropriation and redistribution.

The role of an embedded village cadre could be more complicated than a non-embedded local cadre. For those villages that cannot form an agreement on the choice of village cadre, they are appointed with cadres from the county government. While all village cadres are charged with the responsibility to develop village economy, an embedded village cadre is possibly more capable of mobilizing land investment due to trust from his lineage relatives. (Oi, 1985) Therefore, an embedded village cadre might result in even more land expropriation at a given number of local

[1] Land, just like any other resources, is theoretically collectively owned in PRC. In rural China, the village collective claims ownership of all land resources and other production resources. The attempt of collective farming and distribution – People's Commune and its primitive forms (1958-1984) – saw a bitter failure by causing the Great Famine of 1959-1961, and economic slow-down. As People's Commune proved the failure of collective rural economy and finally collapsed, rural China carried out "land contracting" in 1984, which served essentially as privatization of land resources. Household sign contracts with village collective for franchise of the land for fixed term. The first-round land contracting set the term at 15 years, and 1998 witnessed the launch of the second round land contracting, which was designed to set land issue for the coming three decades. In 2017, the contract was uniformly prolonged for another 30 years. (Xinhua, 2017)

enterprises. Correspondingly, we could also expect that an embedded village cadre devotes more time and energy in managing or providing service for the village enterprises as compared to the non-embedded cadres. (Walder, 1995) Given the above discussion, we come up with the following hypothesis:

1. Villages with more local enterprises undergo more land expropriation as compared to villages with fewer local enterprises.

2. Villages with embedded local cadres undergo more land expropriation as compared to villages without embedded local cadres, given the number of local enterprises.

3. Embedded local cadres devote more time/energy in managing the local enterprise.

4 Hypothesis Evaluation

We continue to rely on the data sets provided generously by Professor Mattingly. The data sets were aggregated from the rural subset of CGSS 2005 A and B survey. While we did acquire the CGSS 2005 A data set from Renmin University of China, the geographical information which would have allowed us to distinguish interviewees from the same villages was muted as according to the regulation; CGSS 2005 B also became inaccessible to the public. As a result, we were unable to aggregate the dataset ourselves from the original 2005 A data set. Our evaluation of hypothesis will rely on the data set provided by Professor Mattingly.

We will evaluate our hypothesis using linear probability model by including the original variable of interest – "embedded", which is a dummy variable that denotes whether the village cadre is a lineage leader, and the measure of existing clan. We also kept the control variables that measure the economic status of the village, solidarity of village, as well as the fixed effect at the provincial level. The control variables largely remained the same: ethnic index, surname index, distance to county, logged household, mean of township election, terrain and soil quality. (Mattingly, 2016) We replaced one measure of economic development – the nightlight index from 1992 – with logged mean income at village level, as the nightlight index was collected 13 years prior to the CGSS 2005 survey. The possibly outdated measurement might introduce more disturbance to the model than actually make contributions. And since we are not able to relocate the village and thus match the nightlight luminosity index ourselves, we are also more confident with the data derived from CGSS 2005.

We also introduce the variable of interest for hypothesis 1: the total number of private and collective enterprise of a village. The reason to include private enterprises is that during early 1990s when collective enterprises at the verge of bankruptcy were mostly sold to better-off households or the factory managers, thus changing the registration status into privately-owned enterprise. However, the enterprises still employed local villagers and shared the obligation for providing welfare for the village, and expands by aggregating land obtained from the local villagers. (Yuan, 1998) (Lan, 2005) Therefore both enterprises should be regarded as potential driving force for land expropriation, and the coefficient of interest should be positive. The model we adopt is expressed as follows

$$y_i = \alpha + \beta_1 x_i + \beta_2 z_i + \gamma \boldsymbol{W_i} + \theta_j + \epsilon_i \tag{1}$$

We maintain the general form of model adopted by Mattingly with slight modifications, where the y_i is the dummy variable for land expropriation, x_i is the dummy variable for whether the local cadre is a lineage leader, and z_i is the number of local enterprises. The control variables are included in the matrix $\boldsymbol{W_i}$ which are included in most specifications, and θ_j denotes the fixed effect. The coefficient of interest for our model is β_2. The regression result is presented in Table 11.

As we could see from the regression result, the coefficient of interest is statistically significant throughout different model specification, while the original coefficient of interest (lineage elite being cadre) is still statistically significant. However, the coefficient for lineage elites are smaller than the original regression results, which shifted from 0.14 to around 0.10. [2] Which means holding all other covariates at the same level, having a lineage elite results in about 10% increase in probability of land expropriation. While holding all other covariates at the same level, one more village enterprise results in 5% increase of probability of land expropriation. At the same time, note that distance to township is negatively correlated with land expropriation probability. This result is also consistent with our hypothesis: the closer it is a village to the township, the more likely it becomes involved in non-agricultural economic activities and develop enterprises. Thus the regression result supports our hypothesis that village enterprises are a driving force for land expropriation.

The regression result from Table 11 still suggests that it is possible that village cadres who are lineage leaders could be more capable of persuading their villagers despite the number of

[2] See Mattingly 2016, pp401, Table 3. (Mattingly, 2016)

Table 11: Evaluation of hypothesis 1

	Dependent variable:			
	Land Expropriation			
	(1)	(2)	(3)	(4)
Lineage leader is cadre	0.110^{*}	0.098^{*}	0.107^{*}	0.111^{*}
	(0.056)	(0.058)	(0.058)	(0.059)
Total village enterprises	0.010^{***}	0.008^{***}	0.007^{**}	0.005^{*}
	(0.002)	(0.003)	(0.003)	(0.003)
Active lineage	−0.163	−0.046	−0.068	−0.060
	(0.129)	(0.158)	(0.158)	(0.160)
Average income (logged)		0.033	0.034	0.023
		(0.045)	(0.045)	(0.047)
Distance to county seat (km)		−0.0004	0.0003	0.0005
		(0.001)	(0.001)	(0.001)
Terrain roughness		0.00002	0.00004	0.00003
		(0.0001)	(0.0001)	(0.0001)
Agricultural suitability index		0.052^{*}	0.044	0.036
		(0.028)	(0.028)	(0.030)
Township control over election			0.114	0.087
			(0.094)	(0.097)
Distance to township (km)			-0.010^{***}	-0.010^{**}
			(0.004)	(0.004)
Surname fragmentation index				0.125
				(0.091)
Ethnic fragmentation index				−0.131
				(0.194)
Number of households (logged)				0.026
				(0.034)
Province fixed effects	No	Yes	Yes	Yes
Constant	0.109^{***}	−0.322	−0.310	−0.420
	(0.021)	(0.430)	(0.428)	(0.466)
Observations	389	387	387	373
R^2	0.054	0.145	0.164	0.152

Note: *p<0.1; **p<0.05; ***p<0.01
From aggregated CGSS 2005 A and B data set.

local enterprises. Therefore, to further stress our argument that economic benefit is one of the driving forces for land expropriation, we extend our model into model (2) to evaluate hypothesis 2. Hypothesis 2 points directly at our argument that economic motivations rather than coercive power of lineage norms is the facilitated by lineage leaders holding powers.

For hypothesis 2, we include the interaction between number of enterprises and the dummy variable "embedded", which denotes whether the village cadre is a linage leader or not. For this hypothesis to hold, we expect the coefficient for the interaction term to be positive and statistically significant. The model is expressed as follows:

$$y_i = \alpha + \beta_1 x_i + \beta_2 z_i + \beta_3 (x_i * z_i) + \gamma \boldsymbol{W_i} + \theta_j + \epsilon_i \tag{2}$$

The model is mostly the same as the previous model, except that we include the interaction term of the number of enterprises and whether the village cadres are lineage leaders. The coefficient of interest for this step is β_3. We continue to evaluate this hypothesis using CGSS 2005 aggregated dataset provided by the author, the regression result is presented in Table 12.

As we can see from the regression result, the interaction term is statistically significant, while

Table 12: Evaluation of hypothesis 2

	Dependent variable:			
	Land Expropriation			
	(1)	(2)	(3)	(4)
Lineage leader is cadre	0.022	0.007	0.018	0.029
	(0.063)	(0.065)	(0.065)	(0.066)
Enterprises * lineage leader is cadre	0.028^{***}	0.028^{***}	0.028^{***}	0.026^{***}
	(0.009)	(0.009)	(0.009)	(0.010)
Total village enterprises	0.008^{***}	0.006^{**}	0.005^{*}	0.003
	(0.002)	(0.003)	(0.003)	(0.003)
Active lineage	−0.133	−0.048	−0.069	−0.058
	(0.128)	(0.157)	(0.157)	(0.159)
Average income (logged)		0.017	0.019	0.011
		(0.045)	(0.045)	(0.047)
Distance to county seat (km)		−0.0004	0.0003	0.001
		(0.001)	(0.001)	(0.001)
Terrain roughness		0.00003	0.0001	0.00004
		(0.0001)	(0.0001)	(0.0001)
Agricultural suitability index		0.046^{*}	0.039	0.030
		(0.028)	(0.028)	(0.030)
Township control of election			0.108	0.079
			(0.093)	(0.096)
Distance to township (km)			-0.010^{***}	-0.010^{**}
			(0.004)	(0.004)
Surname fragmentation index				0.118
				(0.090)
Ethnic fragmentation index				−0.116
				(0.193)
Number of households (logged)				0.028
				(0.034)
Province fixed effects	No	Yes	Yes	Yes
Constant	0.115^{***}	−0.176	−0.164	−0.300
	(0.021)	(0.429)	(0.427)	(0.464)
Observations	389	387	387	373
R^2	0.076	0.166	0.184	0.169

Note: *p<0.1; **p<0.05; ***p<0.01
From aggregated CGSS 2005 A and B data set.

at the same time the original coefficient of interest becomes insignificant and the coefficient itself decreases. Which means when the lineage leader is the village cadre, one more local enterprise is associated with the increase of about 2.6% increasing in land expropriation probability. When the village cadre is not from the local kin groups, the increase of enterprises results in a 0.3% increase of land expropriation probability, which is also statistically insignificant when we control for all covariates. This speaks to our argument that the effect of local lineage group is demonstrated when there are more local enterprises, which could be a result of the lineage leaders who are embedded have more at stake for running a successful enterprise which results in more expansion of local enterprises; or it could also be that when the persuasive power is more likely to take effect when more enterprises exists, and therefore the cadres receive more pressure from the local enterprises. At this point, it is hard to rule out entirely the explanation of "persuasive power" of the lineage elites, but our hypothesis is that local enterprises is necessary for lineage elites to exert influence on land appropriation.

Before we move on to making further claims on our regression findings, we would like to examine the model specification using robust standard errors and clustered robust standard errors, clustering at the province level. (King and Roberts, 2014) When the model is well-specified, the RSE's and cluster RSE should be rather close to the standard error extracted from the classic OLS regression. We test the model using model (4) in table 12, which includes most control variable. The RSE's and clustered SE are plotted in the following graph. [3] As the plot shows, the robust standard errors are very similar to the classic standard error, as well as the clustered robust standard errors. Which means our OLS model captured most of the variation for most of the covariates. (King and Roberts, 2014) The variables with relatively large deviation among these standard errors are intercept and ethnic fragmentation index.

Therefore, our argument that economic incentives is one likely force to drive lad expropriation is supported by the regression and robust to model specification tests. This result is also congruent with the fact that villagers maintained trust for their linage leaders who are village cadres, and provides an explanation for the phenomenon described by Mattingly: the peasants resist the land expropriation as it launches, while as time goes by the economic benefit demonstrated by running local enterprise appease the disagreement and result in further investment in land resources.

Although the current data set might not allow testing mechanism, we could still show suggestive

[3] We also attempted to rule out model dependence using Mahalanobis matching by checking the pruning choices using Match Frontier. The matching tests is beased on the idea of challenging the original model presented by Mattingly, where we include the number of local enterprise will change his model robustness in matching tests.However, as Mahalanobis matching did not significantly improve balance, this part is presented in the appendix.

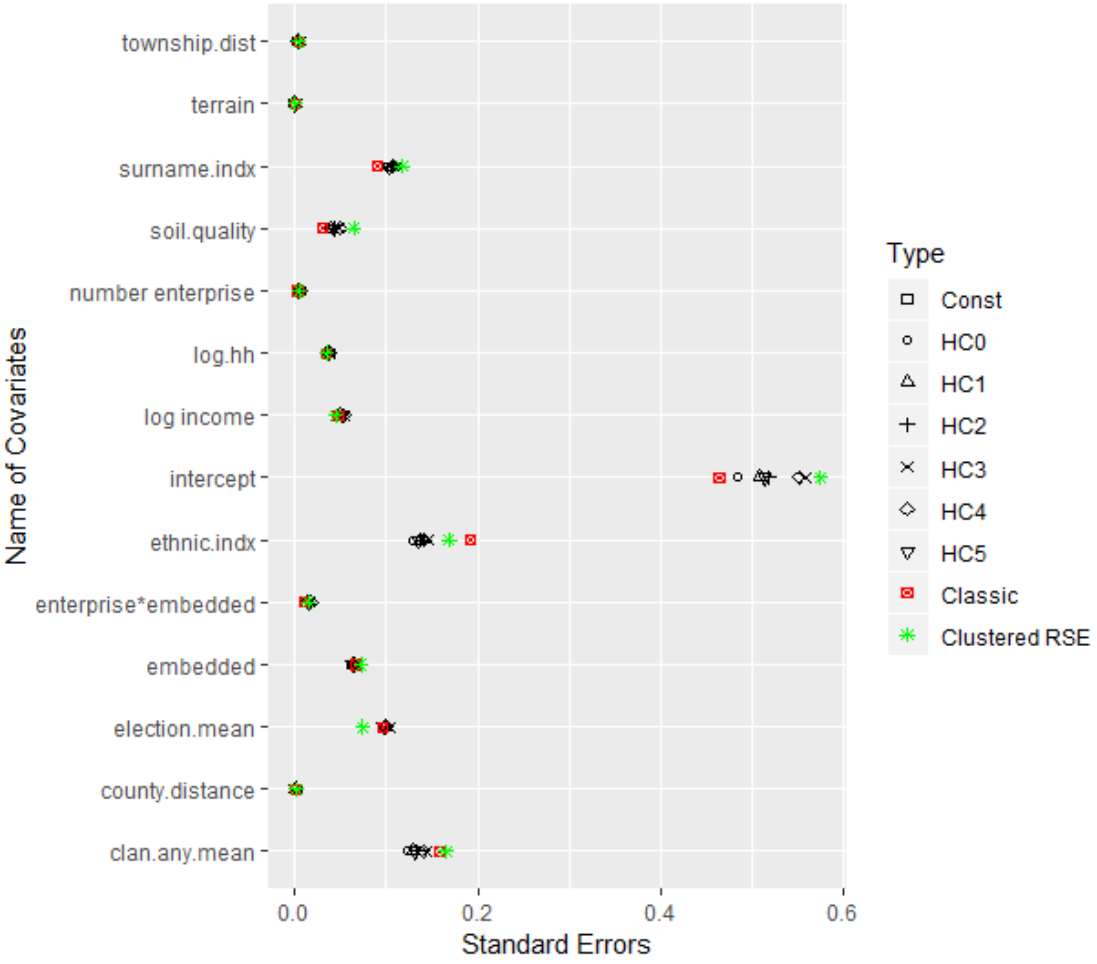

Figure 5: Robust Standard Error and Clustered Standard Error

evidences for traces of causal chains. The incentives of peasants to trust their lineage leader could be a result of lineage leaders are bonded by pressure within the kinship group, and thus are incentivized to promote local economy by devoting more energy into local enterprises. Therefore, we move on to hypothesis 3 using the following model.

$$y_i = \alpha + \beta_1 x_i + \beta_2 z_i + \gamma \boldsymbol{W_i} + \theta_j + \epsilon_i \tag{3}$$

The dependent model y_i is the proportion of time investment in managing local enterprises. We include similar control variables as model 1, changing the dependent variable into time investment proportion of the village cadres. The regression result is presented in Table 13.

Table 13: Evaluation of hypothesis 3

	Dependent variable:			
	Time Investment in Local Enterprise Management			
	(1)	(2)	(3)	(4)
Lineage leader is cadre	0.629	1.936*	2.001*	1.967*
	(0.838)	(1.051)	(1.063)	(1.000)
Number of Local Enterprise	0.183***	0.194***	0.200***	0.127***
	(0.036)	(0.043)	(0.043)	(0.045)
Active lineage	−0.536	−0.303	−0.612	0.707
	(1.933)	(3.056)	(3.081)	(2.902)
Average income (logged)		1.183	1.140	1.377*
		(0.775)	(0.776)	(0.753)
Distance to county seat (km)		−0.007	−0.014	−0.019
		(0.017)	(0.018)	(0.016)
Terrain roughness		0.00004	−0.00000	−0.0005
		(0.002)	(0.002)	(0.002)
Agricultural suitability index		−0.436	−0.444	−0.191
		(0.609)	(0.609)	(0.576)
Township control of election			1.408	0.819
			(1.565)	(1.508)
Distance to township (km)			0.079	0.063
			(0.066)	(0.064)
Surname fragmentation index				1.730
				(1.558)
Ethnic fragmentation index				−0.041
				(3.250)
Number of households (logged)				−1.632***
				(0.564)
Province fixed effects	No	Yes	Yes	Yes
Constant	1.124***	−11.151	−11.186	−4.782
	(0.321)	(7.376)	(7.387)	(7.650)
Observations	385	287	287	277
R^2	0.064	0.223	0.230	0.202

Note: *p<0.1; **p<0.05; ***p<0.01
From aggregated CGSS 2005 A and B data set.

The regression result supports hypothesis 3: at a given level of local enterprises and holding all other covariates, being a lineage leader cadre means more time investment in local enterprises as compared to non-lineage leader cadres. The number of enterprise also has a positive coefficient, which means more local enterprises will result in more time spent on providing service for local enterprises. The number of households is negatively correlated with time investment on enterprise management, which does not contradict our theory, as more households means more time spent

on conflict resolutions and daily affairs on managing the village.

The finding of our extension work are that economic considerations is likely a reason for embedded cadres being able to mobilize more land resources in developing local village economy. This points to overcoming the coordination problem in developing local enterprises by taking advantage of existing informal institutions, and adds in rationale for persistent compliance of the villagers. While Confucianism norms are believed to promote compliance, they need not override self-interested calculation in modern Chinese rural village, which more or less sets China as an exception and shields off challenge from the conventional wisdom.

References

Committee, National People's Congress Standing. 1998. "The Law of Land Administration of the People's Republic of China.".

Congress, National People's. 1982. "1982 Constitution of People's Republic of China.".

CPCDI. 2017. "100 QA after The 19th Communist Party of China National Congress." *http://www.ccdi.gov.cn/special/zmsjd/zm19da$_z$m*19*da*/201712/*t*20171213$_1$13732.*html*.

King, Gary and Margaret Roberts. 2014. "How Robust Standard Error Exposes Methodological Problems They Do Not Fix, and What to Do About it." *Political Analysis* 23:159–179.

Lan, Yuyun. 2005. *Villages in Urban Townships – A Field Research on "Neo-Village Community" (dushi li de cunzhuang yige xin cunshe gongtongti de shidi yanjiu).* Beijing: Sanlian Bookshop.

Mattingly, Daniel. 2016. "Elite Capture: How Decentralization and Informal Institutions Weaken Property Rights in China." *World Politics* 68:383–412.

Oi, Jean. 1985. "Communism and Clientelism: Rural Politics in China." *World Politics* 37:238–266.

Oi, Jean. 1992. "Fiscal Reform and the Economic Foundations of Local State Coorporatism in China." *World Politics* 45:99–126.

Walder, Andrew. 1995. "Local Government as Industrial Firms: An Organizational Analysis of China's Transitional Economy." *American Journal of Sociology* 101:263–301.

Xinhua, Press. 2017. "Premiere Li: Ensuring Prolonging Second-round Land Contracting Terms for Another 30 Years After the Current Term (*likeqiang: luoshi dierlun tudi chengbao daoqi hou zai yanchang 30 nian zhengce*)." *http://www.xinhuanet.com/politics/2018lh/2018-03/05/c*$_1$37017009.*htm*.

Yuan, Peng. 1998. *Hundred Role Models of Rural Economic Development – Fu'an Village (quanguo baijia xiangcun jingji dianxing diaocha – fu'an cun).* Beijing: Comtemporary China Press.

5 Appendix

1. `Matching` results.

   Mahalanobis Matching and MatchFrontier for Extension Model 2.

Table 14: Before Match

Covariate	mean.Tr	mean.Co	sdiff	sdiff.pooled	var.ratio	T pval
Land expropriation	0.20	0.13	16.83	18.22	1.42	0.24
Active lineage (mean)	0.23	0.03	79.90	102.09	4.44	0.00
Distance to county seat (km)	22.38	30.66	-41.24	-39.19	0.82	0.01
Terrain roughness	257.91	227.63	12.45	12.19	0.92	0.40
Agriculture suitability index	1.40	1.66	-43.14	-35.42	0.51	0.01
Township control of election	0.16	0.16	-0.30	-0.28	0.76	0.98
Distance to township (km)	5.24	5.83	-12.80	-11.85	0.75	0.39
Surname fragmentation index	0.67	0.72	-20.95	-21.49	1.11	0.15
Ethnic fragmentation index	0.03	0.03	-1.69	-1.67	0.94	0.91
Number of households (logged)	6.02	6.04	-2.99	-2.99	1.01	0.84
Province ID	12.62	12.48	1.97	2.15	1.46	0.89
Average income	9893.83	9543.04	6.44	6.27	0.90	0.66
Private enterprise	0.20	0.36	-29.54	-11.56	0.08	0.21
Collective enterprise	2.62	2.90	-5.56	-4.71	0.56	0.72

Table 15: After Match

Covariate	mean.Tr	mean.Co	sdiff	sdiff.pooled	var.ratio	T pval
Land expropriation	0.20	0.04	40.54	40.54	4.57	0.01
Active lineage (mean)	0.23	0.10	52.65	52.65	1.11	0.00
Distance to county seat (km)	22.38	22.13	1.27	1.27	1.57	0.89
Terrain roughness	257.91	171.82	35.39	35.39	1.53	0.00
Agriculture suitability index	1.40	1.42	-3.05	-3.05	0.90	0.76
Township control of election	0.16	0.13	11.91	11.91	2.02	0.24
Distance to township (km)	5.24	4.98	5.46	5.46	1.21	0.53
Surname fragmentation index	0.67	0.71	-14.48	-14.48	0.97	0.09
Ethnic fragmentation index	0.03	0.02	14.61	14.61	2.67	0.03
Number of households (logged)	6.02	6.08	-8.64	-8.64	0.96	0.26
Province ID	12.62	12.25	5.11	5.11	1.13	0.36
Average income	9893.83	9666.68	4.17	4.17	1.09	0.66
Private enterprise	0.20	0.18	3.26	3.26	1.65	0.86
Collective enterprise	2.62	2.31	6.18	6.18	1.84	0.41

The ATT after matching is as presented in Table 16.

As we can see from the two tables of balance condition of pre-matching and post-matching results, one-on-one Mahalanobis matching did not improve balance condition for most of the variables. The variables are mostly balanced prior to matching. While after match, the ATT becomes statistically insignificant. Which serves to support our observation that taking into

Table 16: After balance ATT

Estimate	0.16364
AI SE	0.059102
T-stat	2.7687
p.val	0.0056282
Original number of observations	373
Original number of treated obs	55
Matched number of observations	55
Matched number of observations (unweighted).	55

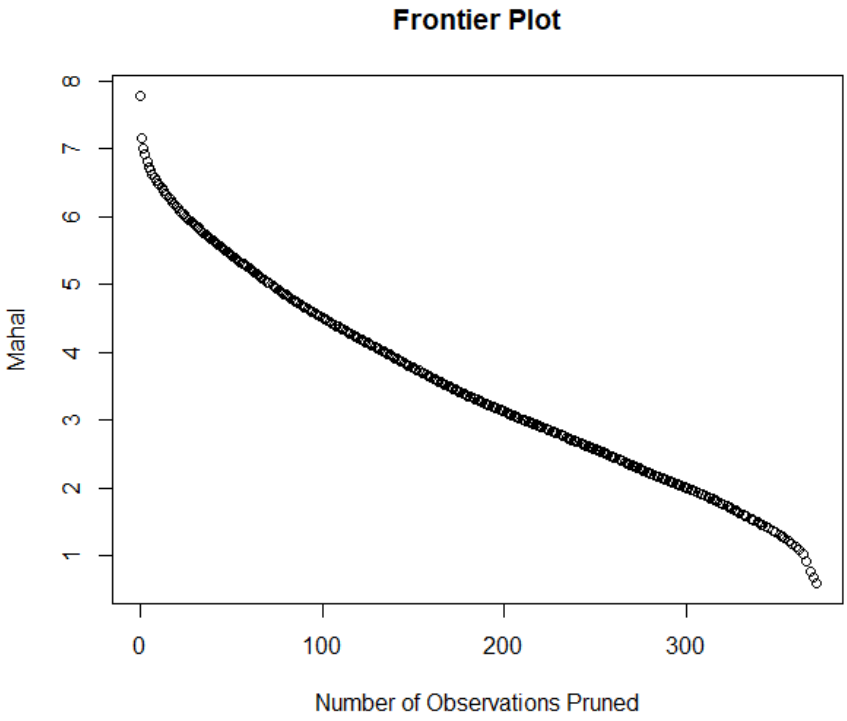

Figure 6: Frontier Plot

consideration of local enterprises will makes the original model of Mattingly 2016 no longer robust to matching tests

2. `MatchFrontier`

   We also check our pruning decisions using `MatchFrontier`, which generates the following frontier plot as shown in figure 3, and the expected treatment effect with varying number of prune, as shown in figure 7. Therefore as we can see, the confidence interval covers 0 when we vary the number of prune, which suppports our challenge to the original model.

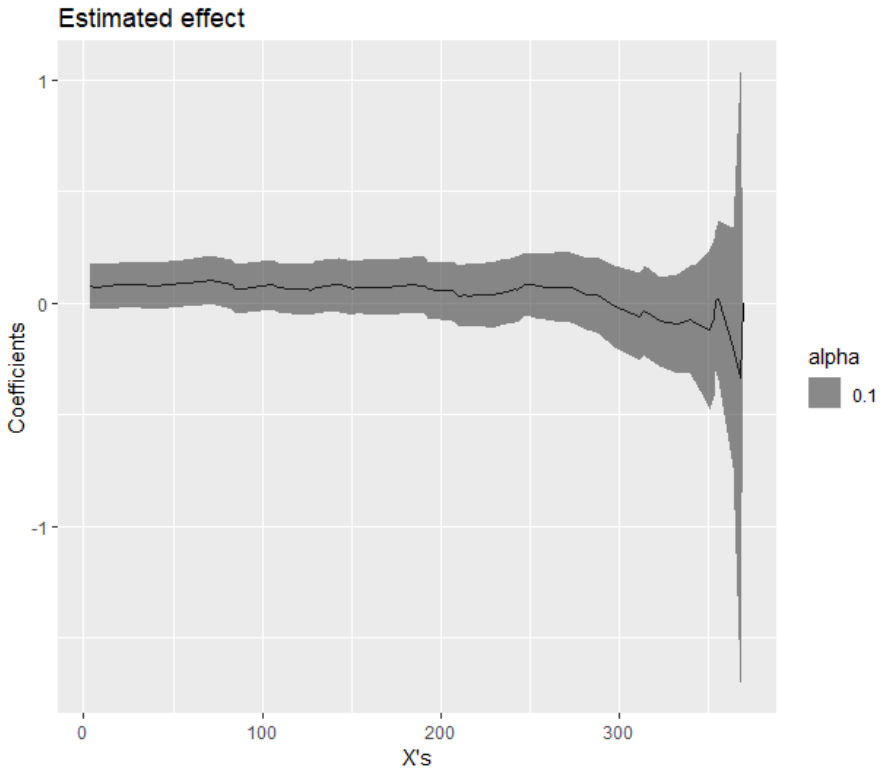

Figure 7: Expected treatment effect

3. Other robustness checks

We can see from the new model of the pruned dataset in Table 17 that the indicator variable for cadre lineage leaders is once again significant at the $p<0.05$ level, while the distance to township and surname fragmentation index are significant to the same levels that they were in the original OLS model. This verifies the robustness of the original model.The standard errors do not vary much for any of the models in Tables 18-20 either, suggesting that the model fit is acceptable for the data.

Table 17: Model created using pruned dataset

	Dependent variable:
	Land expropriations
Lineage leader is cadre	0.173^{**}
	(0.073)
Active lineage	-0.067
	(0.058)
Distance to county seat	0.001
	(0.001)
Terrain roughness	-0.00002
	(0.0001)
Agricultural suitability index	0.034
	(0.024)
Wealth (nighttime lights proxy)	0.002
	(0.003)
Township control over elections	0.137
	(0.090)
Distance to township	-0.009^{**}
	(0.004)
Surname fragmentation index	0.159^{*}
	(0.085)
Ethnic fragmentation index	-0.249
	(0.184)
Number of households (log)	0.019
	(0.028)
Constant	-0.013
	(0.177)
Observations	359
R^2	0.082
Adjusted R^2	0.050
Residual Std. Error	0.349 (df = 346)
F Statistic	2.565^{***} (df = 12; 346)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 18: Table 4 in Mattingly (Strength of Lineage Institutions) with Standard Errors, Robust Standard Errors and Cluster-Robust Standard Errors

	Dependent variable:		
	Land expropriations occurring in the village		
	(1)	(2)	(3)
Lineage leader is cadre	0.207**	0.207**	0.207**
	(0.087)	(0.089)	(0.083)
Active lineage with ancestral hall or zupu	−0.080	−0.080	−0.080*
	(0.063)	(0.050)	(0.047)
Economic controls	yes	yes	yes
Political controls	yes	yes	yes
Social controls	yes	yes	yes
Province fixed effects	yes	yes	yes
Observations	359	359	359
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 19: Table 5 in Mattingly (Placebo Tests) with Standard Errors, Robust Standard Errors and Cluster-Robust Standard Errors

	Dependent variable:		
	Land expropriations before the current leadership		
	(1)	(2)	(3)
Lineage leader is cadre	0.001	0.001	0.001
	(0.040)	(0.045)	(0.044)
Active lineage	0.030	0.030	0.030
	(0.032)	(0.033)	(0.028)
Economic controls	yes	yes	yes
Political controls	yes	yes	yes
Social controls	yes	yes	yes
Province fixed effects	yes	yes	yes
Observations	359	359	359
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 20: Table 6 in Mattingly (Petitioning and Land Expropriations) with Standard Errors, Robust Standard Errors and Cluster-Robust Standard Errors

	Dependent variable:		
	Village participation in collective petitioning		
	(1)	(2)	(3)
Lineage leader is cadre	−0.003	−0.003	−0.003
	(0.054)	(0.061)	(0.054)
Land seizure	−0.028	−0.028	−0.028
	(0.053)	(0.043)	(0.040)
Lineage leader × Land seizure	0.269**	0.269	0.269
	(0.115)	(0.172)	(0.188)
Constant	−0.016	−0.016	−0.016
	(0.157)	(0.159)	(0.112)
Observations	359	359	359
Economic controls	yes	yes	yes
Political controls	yes	yes	yes
Social controls	yes	yes	yes
Province fixed effects	yes	yes	yes
Note:			*p<0.1; **p<0.05; ***p<0.01