

Kjelsrud, Anders; Kotsadam, Andreas; Rogeberg, Ole

Working Paper

Cooperative Property Rights and Development: Evidence from Land Reform in El Salvador: A Comment

I4R Discussion Paper Series, No. 20

Provided in Cooperation with:

The Institute for Replication (I4R)

Suggested Citation: Kjelsrud, Anders; Kotsadam, Andreas; Rogeberg, Ole (2023) : Cooperative Property Rights and Development: Evidence from Land Reform in El Salvador: A Comment, I4R Discussion Paper Series, No. 20, Institute for Replication (I4R), s.l.

This Version is available at:

<https://hdl.handle.net/10419/268943>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



No. 20
I4R DISCUSSION PAPER SERIES

Cooperative Property Rights and Development: Evidence from Land Reform in El Salvador: A Comment

Anders Kjelsrud

Andreas Kotsadam

Ole Rogeberg

March 2023

I4R DISCUSSION PAPER SERIES

I4R DP No. 20

Cooperative Property Rights and Development: Evidence from Land Reform in El Salvador: A Comment

Anders Kjelsrud¹, Andreas Kotsadam², Ole Rogeberg²

¹Oslo Business School, OsloMet, Oslo/Norway

²The Frisch Centre, Oslo/Norway

MARCH 2023 (updated Version, March 10)

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the [ZBW – Leibniz Information Centre for Economics](#), and [RWI – Leibniz Institute for Economic Research](#), I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur
University of Ottawa

Anna Dreber
Stockholm School of Economics

Jörg Ankel-Peters
RWI – Leibniz Institute for Economic Research

Cooperative Property Rights and Development: Evidence from Land Reform in El Salvador: A Comment*

Anders Kjelsrud[†] Andreas Kotsadam[‡] Ole Rogeberg[§]

March, 2023 (updated version, March 10)**

Abstract

Montero (2022) explores a discontinuity in a land reform in El Salvador and reports two main findings. First, relative to outside-owned haciendas operated by contract workers, the productivity of worker-owned cooperatives is higher for staple crops and lower for cash-crop. Second, cooperative property rights increase workers' incomes and compress wage distributions. In this comment, we show that the latter result rests on two mistakes: three-quarters of the observations are duplicates and income inequality is calculated over too few workers to be meaningful. When corrected, the data sources and research design provide no credible evidence regarding the causal effects of ownership structure on income levels and inequality.

*The comment is accepted at the *Journal of Political Economy*. The issues were identified during the Oslo Replication Games organized by the Institute for Replication. We thank Eduardo Montero for making the replication data and code freely available, and for his admirable response in support of research integrity in his personal communications with our team. We further thank the editor of the JPE, Emir Kamenica, and Abel Brodeur, Kalle Moene, Sahar Parsa and Oddbjørn Raaum for insightful comments. Replication files for this comment can be found here: https://andreaskotsadam.files.wordpress.com/2023/02/replication_montero.zip.

** We made the following change: The standard errors in columns 5 and 6 in table 1 were erroneously reported as 24.79 and 43.09. The standard errors are now reported as 43.09 and 84.53. We thank Rony Rodriguez-Ramirez for pointing out this error.

[†]Oslo Business School, OsloMet, Oslo, Norway. anders.kjelsrud@oslomet.no.

[‡]The Frisch Centre, Oslo, Norway. andreas.kotsadam@frisch.uio.no.

[§]The Frisch Centre, Oslo, Norway. ole.rogeberg@frisch.uio.no.

1 Introduction

Montero (2022) uses a land reform in El Salvador to investigate the long run causal effects of ownership structure on agricultural production and worker outcomes. The paper presents two main empirical findings: 1. Relative to outside-owned haciendas operated by contract workers, the productivity of worker-owned cooperatives is higher for staple crops and lower for cash-crops. 2. The cooperative workers have higher incomes and more compressed wage distributions than those working in the haciendas.

In this comment, we show that the second set of results rests on two mistakes in the analysis: three-quarters of the observations are duplicates and income inequality is calculated over too few workers to be meaningful. When corrected, the data sources and research design provide no credible evidence regarding the causal effects of ownership structure on income levels and inequality.

The identified issues concern a central finding of the paper: the causal effects on income levels and inequality are highlighted in both the abstract and the conclusion, as well as in the author’s own popularization of the work (see Montero, 2017, 2019). The result has also been referenced by others as clear evidence that land reforms lead to “*more equitable rural economies based on staple crops and higher incomes*” (Boberg-Fazlić et al., 2022, p.2), and that cooperative agriculture facilitates “*a more equitable distribution of agricultural revenue*” (Rao and Shenoy, 2022, p.6). Although these empirical claims are intuitively plausible and may well be true, they nonetheless lack empirical support of the kind claimed in Montero (2022).

We stress that the errors are confined to the analyses of income levels and distributions and that the research design employed in Montero (2022) is theoretically appropriate, but the specification described is not the one implemented and the actual data available is insufficient for the analysis. Below we start by describing the original results in more detail. We then describe the errors in turn and provide updated estimates, before we conclude.

2 Reproducing the original results

Montero (2022) estimates the long run causal effects of a property rights reform in El Salvador on workers’ income and inequality. The reform had been developed in secret

and the military moved swiftly to implement the first phase in 1980: seizing agricultural land from owners whose total agricultural land holdings across all plots exceeded 500 hectares (ha), and transferring ownership to agricultural cooperatives run by permanent laborers working the land.

Three separate analyses are used to assess the effects on worker inequality and income. Two of these use a regression discontinuity design (RD) to assess how inequality and income levels differ for properties on different sides of the 500 ha discontinuity threshold, while the last uses a quantile regression to assess differences in the income distribution of haciendas and cooperatives.

The analyses use data from a repeated household survey, *Encuesta de Hogares de Propósitos Múltiples*, covering the period 2000 to 2013. This data includes geographical identifiers for each participant at the level of department, municipality and canton.¹ The survey data also includes information on whether an individual works in agriculture as a cooperative member or as a hacienda laborer. [Montero \(2022\)](#) uses this information to link the household survey data with the other property data, restricting the matching to cantons with only one cooperative/large hacienda (p.60).

Inequality is measured using the interquartile range (IQR) of real per capita household income, defined as the value for the 75th percentile minus the value for the 25th percentile. In the replication code, this calculation is done at the level of property-year. Collapsing the data to this level, the specification used to estimate the effect on inequality can be written as follows:

$$\text{IQR}_{po,t} = \gamma \text{Above500}_o + f(\text{holdings}_o) + \sigma_t + \epsilon_{po,t}, \quad (1)$$

where $\text{IQR}_{po,t}$ is the IQR in year t among workers at plot p owned by owner o before the reform, Above500_o is an indicator variable for whether owner o had over 500 ha in cumulative landholdings before the reform, $f(\text{holdings}_o)$ is a linear control for the forcing variable (total landholdings of pre-reform property owner).

In columns 1 and 2 of table 1, we reproduce the main results on worker equity using this specification. Column 1 displays estimates when limiting the sample to properties within 300 ha of the reform threshold, while column 2 displays estimates when limiting the sample to properties within 150 ha of the reform threshold. The results suggest that

¹These cantons roughly correspond to villages according to footnote 11 in [Montero \(2022\)](#).

the income distribution in cooperatives is more compressed than in haciendas.

The same survey data is used to study the effect of cooperatives on workers' income level, using the following specification:

$$\text{IncomePC}_{ipo,t} = \gamma \text{Above500}_o + f(\text{holdings}_o) + \sigma_t + \epsilon_{ipo,t}, \quad (2)$$

where $\text{IncomePC}_{ipo,t}$ is real income per capita (last month) for the household of worker i . We reproduce these estimates in columns 1 and 2 of table 3, using the broad and narrow bandwidth, respectively. The results suggest that workers at cooperatives have higher incomes on average. Finally, [Montero \(2022\)](#) uses the household survey data to estimate a quantile regression, which investigates how the income premium for cooperative workers varies across the income distribution. We reproduce this in the left panel of figure 1. The figure suggests that the cooperative ownership disproportionately boosts the earnings of the lowest quantiles.

3 Identified issues with the analysis

In this section, we first describe a mistake in the data construction: three-quarters of the worker-level data consist of duplicate observations. We then show that the implemented analysis differs from the description in the paper and that it measures income inequality based on too few workers to be meaningful. Correcting these issues, we conclude that the reported results on income levels and distributions do not hold.

3.1 Duplicate observations

A visual inspection of the replication data indicates a large number of duplicates, as illustrated by the screenshots from the data editor in figure 2.

We use the raw data from the survey to remove duplicates from the author-submitted analysis data file. We first restrict the survey sample to observations with non-missing “property type”, as this variable was used by [Montero \(2022\)](#) to match the survey data to the property data. Next, we identify a set of survey variables that are both retained in the author’s analysis data and sufficient to uniquely identify individuals in the raw

data.² This allows us to remove all duplicates by collapsing the author’s estimation data to unique combinations of the identifying variables.³ For the broad bandwidth sample, this reduces the number of workers by three-quarters, from 4,770 to 1,146.

The duplication issue affects 129 of the 327 property-year observations, but leaves the number of aggregate property-year observations largely unchanged: as long as the duplicated observations of an individual are assigned to the same property, there will be at least one individual-level observation remaining to inform the property-year observation. This is largely the case, so while the duplicates reduce the individual level sample to a quarter of its original size, the number of property-year observations is only reduced by three.⁴⁵

3.2 Measuring inequality with too few observations

Inequality is measured at the property-year level as the inter-quartile range (IQR) of the individual-level survey observations assigned to this property-year cell. As noted in [Montero \(2022\)](#), this measure requires a certain number of observations to be meaningful: *“To construct measures of the income distributions, I limit the sample to cooperatives and haciendas for which there are at least five members represented in the household surveys”* ([Montero, 2022](#), p.75). The sample restriction means that the estimation sample will be more strongly affected by the duplicate issue than the “unrestricted” property-year sample, since the removal of duplicates strongly reduced the number of individual-level observations informing the property-year observations.

In addition, a review of the paper’s replication code revealed that the sample restric-

²The following variables are sufficient to uniquely identify individuals within cantons: age, household size, household income, landholding type and the number of other workers at the property. The only exception is 2 non-unique observations in the survey for 2011. These observations are located in cantons that are not in the sample of [Montero \(2022\)](#).

³We manually checked households with more than one remaining observation against the raw data and removed individuals from these households that only appear in the estimation data. We removed 43 observations in this way. It is unclear to us how these observations appeared.

⁴Six unique individuals were found to have duplicated observations allocated to *different* properties. All observations for these individuals were discarded (as we do not know the correct property assignment). Three property-year observations were based on a single observation from one of these six individuals.

⁵A second data error concerns the variable for real household income, used in the computation of the outcome variables. By reviewing the computer codes, we infer that [Montero \(2022\)](#) used a consumer price index (CPI) published by the World Bank to deflate nominal income from different survey years. However, the computer code does not include the year 2010 and observations from this year are assigned a value of real income equal to zero. This implies that the IQR is calculated to be zero for all properties in 2010. In our corrected dataset, we handle this by including 2010 in the computer code, using the CPI value for this year. The correction does not affect the estimates much.

tion described in the main text was not actually implemented in the code for the RD regression on income inequality. Even without removal of the duplicates, some 29% of the property-year units in the original broad bandwidth sample contain only *one* observation. The interquartile range of a single observation is by necessity zero, and these property-year observations cannot inform analyses of property-level inequality. An additional 32% of the property-year units in the author-submitted data contain two to four observations. In total, 61% of the property-year units in the paper’s own estimation sample should therefore have been excluded according to the criteria defined in the paper.

3.3 Updated results

In columns 3 and 4 of table 1, we show estimates of the effects of property right systems on the IQR once we correct the household data as described in section 3.1. The data corrections alone do not affect the estimates much. As mentioned, even though the number of workers in the sample falls by three-quarters, the number of observations in the regressions barely changes as the sample is collapsed on property-year.

We next impose the sample restriction described in [Montero \(2022\)](#) and remove property-year units with less than five observations. We first impose the restriction to the original estimation data (which includes the duplicates). Estimates are shown in columns 5 and 6 of table 1. We are left with 128 property-year observations in the broad bandwidth sample, only three of which are above the RD threshold. In the narrow bandwidth sample we are left with 45 observations, only two of which are above the RD threshold. None of the results are statistically significant (p-values of 0.243 and 0.295, respectively).

We then impose the sample restriction to the corrected dataset (which excludes the duplicates). Estimates are shown in columns 7 and 8 of table 1. The remaining dataset is insufficient to perform the inequality analyses. The sample with the broad bandwidth has 70 remaining property-year units, but only one observation above the RD threshold; the narrow bandwidth sample has 26 property-year units below the threshold but no observations above the threshold. We thus conclude that the data at hand is insufficient to conduct the analysis as described in the paper.

3.4 Results with alternative sample restrictions

The issues we have identified are clear errors: duplicate observations and a failure to implement an explicitly stated (and methodologically important) sample restriction in the actual code.

Having said that, the number of individual-level observations required to compute a meaningful interquartile income range is not obvious. We find the paper's choice of a five-observation minimum reasonable, but it is worth examining how sensitive the results are to alternative thresholds. We use the corrected (non-duplicate) data.

Estimates are shown in table 2. We first restrict the sample to property-years with more than three observations (columns 1 and 2), and then to property-years with more than two observations (columns 3 and 4). The sample sizes increase somewhat for these specifications but there are still only a few property-year units above the RD threshold. As a consequence, the estimated coefficients vary widely, from +182 in column 1 to -28 in column 2.

These results show that there is insufficient data to implement a regression discontinuity analysis on observations at the property-year level, since many property-year cells have sparse counts and do not allow us to meaningfully compute within-cell income inequality. We thus conclude that the original results on property right system on inequality do not hold for any reasonable sample restriction.

3.5 Results on income levels and quantile regression

Our main focus in this comment is on the inequality analysis, as this was a main outcome of interest in [Montero \(2022\)](#). The household survey data is, however, also used to estimate the effects of ownership structure on income levels and to run a quantile regression to identify the income effect at different points of the income distribution.

Starting with the RD analysis of income levels, we find that the estimates are no longer statistically significant in the corrected data (excluding the duplicates). As shown in columns 3 and 4 of table 3, the point estimate drops by three quarters using the broad bandwidth (p-value of 0.684) and switches sign in the narrow bandwidth sample (p-value of 0.566).

Turning to the quantile regression, our code review found that this analysis imple-

ments the sample restriction described for the inequality analysis. The restriction is not warranted in the present context as the quantile regression uses individual-level observations to compare the *overall* population of cooperative workers to the *overall* population of hacienda workers. Also, the sample restriction is not mentioned or discussed in the context of the quantile regression. In the right panel of figure 1 we present the results from the quantile regression when we disregard the sample restriction and use the corrected data.⁶ The figure is much less informative than the published figure and it no longer suggests that the “*earnings policies within reform cooperatives [...] help workers at the bottom of the income distribution*” (Montero, 2022, p.76).

4 Conclusion

In “*Cooperative Property Rights and Development: Evidence from Land Reform in El Salvador*”, it is claimed that workers of cooperatives have less income inequality and higher average incomes than workers of haciendas. The paper provides empirical evidence on this by combining household survey data with other data on cooperatives and haciendas.

In this comment, we have shown that these conclusions are wrong – in the sense that they are not supported by the data once duplicate observations are removed and the analyses are implemented as described in the paper. While the mistakes are confined to the results on worker income and inequality, these results are a key finding highlighted in the abstract and the conclusion of the paper. Recent papers citing Montero (2022) also highlight these results.

We fully agree with Montero (2022) when arguing that “*property rights institutions are of central importance to understanding economic development*”, and that there is “*limited causal empirical evidence on the impacts of different property rights system*” (p. 87). Unfortunately, we still lack robust and credible empirical evidence regarding the effects of different property rights system on worker equity. We hope that more research will be conducted in this area.

⁶We also add controls for age and age squared for each worker as described in Montero (2022) but not implemented in the submitted computer codes. The inclusion of these controls barely affect the estimated coefficients. The paper states that the quantile regression includes a control for gender as well. We are not able to add this control, as gender varies to little over the income distribution (93% of the sample are men).

References

- Boberg-Fazlić, N., M. Lampe, P. M. Lasheras, and P. Sharp (2022). Winners and losers from agrarian reform: Evidence from Danish land inequality 1682–1895. *Journal of Development Economics* 155, 102813.
- Montero, E. (2017). Workers Unite: Cooperative Property Rights and Development in El Salvador. World Bank Development Impact Blog, December 2017.
- Montero, E. (2019). Workers unite: Cooperative property rights and development in El Salvador. VoxDev, April 2019.
- Montero, E. (2022). Cooperative property rights and development: Evidence from land reform in El Salvador. *Journal of Political Economy* 130(1), 48–93.
- Rao, M. and A. Shenoy (2022). Got (Clean) Milk? Organization, Incentives, and Management in Indian Dairy Cooperatives. Abdul Latif Jameel Poverty Action Lab (J-PAL) working paper 863.

TABLE 1: Household Income per Capita, Interquartile Range

	Original results		Corrected data		Sample restriction		Corrected data and sample restriction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Above500	-42.97** (17.82)	-51.79** (24.79)	-72.24*** (17.37)	-79.78*** (22.67)	-50.91 (43.09)	-90.86 (84.53)	-65.81 (45.36)	–
Observations	327	118	324	116	128	45	70	26
Clusters	98	36	98	36	54	21	31	12
#Treated prop.	30	21	27	19	3	2	1	0
#Workers	4,770	1,583	1,146	420	4,377	1,434	675	254
Bandwidth	300	150	300	150	300	150	300	150

The table shows estimates of the effect of the land reform on inequality. Columns 1 and 2 reproduce columns 3 and 4 of table 5 in [Montero \(2022\)](#). The estimates in columns 3 and 4 are based on the corrected data (excluding duplicates). The estimates in columns 5 and 6 impose the sample restriction described in [Montero \(2022\)](#) to the original estimation data (including duplicates). The estimates in columns 7 and 8 impose the sample restriction to the corrected data. “Above500” is a binary variable equal to 1 if the former owner had over 500 ha in cumulative landholdings in 1980; “Observations” displays the number of property-year units; “Clusters” displays the number of clusters (former land owners); “#Treated prop.” displays the number of property-year units above the RD threshold; “#Workers” displays the number of workers included in IQR calculation; while “Bandwidth” shows the RD bandwidth in ha.

*** $p < .01$, ** $p < .05$.

TABLE 2: Alternative sample restrictions and specifications, Interquartile Range

	Property-year n>3		Property-year n>2	
	(1)	(2)	(3)	(4)
Above500	182.20*** (21.52)	-28.35 (49.07)	157.08*** (21.82)	-3.53 (42.03)
Observations	98	37	136	49
Clusters	46	16	62	20
#Treated prop.	4	3	6	4
#Workers	787	298	901	334
Bandwidth	300	150	300	150

The table shows estimates of the effect of the land reform on inequality using alternative sample restrictions. All estimates are based on the corrected data (excluding duplicates). The sample used in columns 1 and 2 is restricted to property-year units with more than three observations, while the sample in columns 3 and 4 is restricted to property-year units with more than two observations. “Observations” displays the number of property-year units; “Clusters” displays the number of clusters (former land owners); “#Treated prop.” displays the number of property-year units above the RD threshold; “#Workers” displays the number of workers included in IQR calculation.

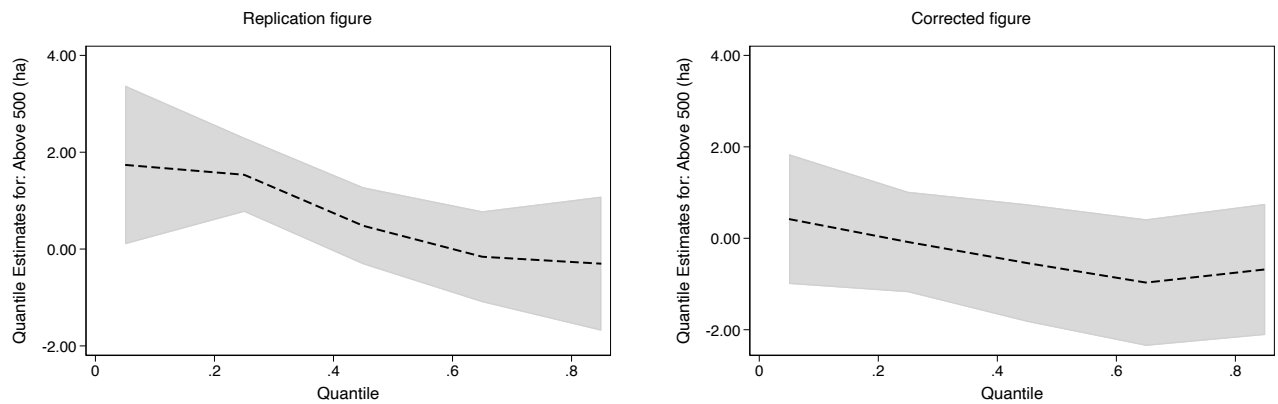
*** $p < .01$

TABLE 3: Household Earnings per Capita (Previous Month), Levels

	Original results		Corrected data	
	(1)	(2)	(3)	(4)
Above500	60.49* (31.03)	52.79 (40.61)	13.21 (32.35)	-24.24 (41.90)
Observations	4770	1583	1146	420
Clusters	98	36	98	36
#Treated	88	48	51	31
Bandwidth	300	150	300	150

The table shows estimates of the effect of the land reform on income levels. Columns 1 and 2 reproduce columns 3 and 4 of table 5 in [Montero \(2022\)](#). The estimates in columns 3 and 4 are based on the corrected data (excluding duplicates). “Observations” displays the number of workers; “Clusters” displays the number of clusters (former land owners); “#Treated” displays the number of workers above the RD threshold.

* $p < .1$.

FIGURE 1: Figure 6 in [Montero \(2022\)](#), replication and corrected

Note: The figure presents estimated quantile RD coefficients, using a bandwidth of 150 ha from the RD threshold. The independent variable is the log of real per capita household income. The regression includes year fixed effects and a linear control for the forcing variable (total landholdings of pre-reform property owner), estimated separately on each side of the RD threshold. Standard errors are clustered at the former-owner level and the gray areas display 95% confidence intervals. The left figure reproduces figure 6 in [Montero \(2022\)](#). The quantile regression shown in the right figure follows the description in [Montero \(2022\)](#): it does not impose any sample restrictions; it is based on the corrected data; and it includes controls for age and age squared of each worker.

FIGURE 2: Screenshots of estimation data

Note: The figure displays four screenshots from the estimation data, labelled “ehpm.incomemodule_wreform.dta” in the replication package of [Montero \(2022\)](#). This datafile is used to study the effect of the land reform on income levels and distributions.