Ludwig-Maximilians-Universität, Munich

Causal Inference in Economics

Replication Study of

The Persistent Effect of Peru's Mining *Mita*

by Melissa Dell
in Econometrica, 2010

Author:

Lolakshi Rajlakshmi

lolakshi.rajlakshmi@campus.lmu.de

Matriculation Number: 12339702

August 19, 2024

Contents

Abstract

This study serves to replicate the main findings of the paper titled "The Persistent Effect of Peru's Mining *Mita*" by Dell (2010). The study by (Dell, 2010) represents a significant contribution to development economics, employing regression discontinuity to examine the long-term economic impact of the Mita, an extensive system of forced labour in the mining industry that operated in Peru and Bolivia between 1573 and 1812. The study demonstrates that the effects of the *mita* endured through its influence on land tenure and the provision of public goods. To this end, data from the Spanish Empire and the Peruvian Republic were employed to trace channels of institutional persistence. My contributions to this body of work provide robustness checks and extensions of the results, relying on advancements in the literature on regression discontinuity designs (RDDs) and spatial RDDs.

1 Introduction

The field of development economics is both fascinating and crucial, as it investigates the mechanisms driving growth and poverty reduction. These studies provide valuable insights into the immediate effects of interventions, guiding policymakers toward more effective strategies. A related and compelling area of research within development economics is the analysis of historical data to understand long-term development patterns. This growing field includes seminal works such as (Acemoglu, Johnson and Robinson, 2001), (Nunn, 2008), and (Ashraf and Galor, 2011), along with other notable studies by the author of this paper like (Dell and Querubin, 2018) and (Dell, Lane and Querubin, 2018).

This field of inquiry leads to an exploration of the historical roots of inequality and the role of colonial institutions in shaping modern economies. A critical question arises: What factors contribute to the economic disadvantages faced by former colonies compared to their colonizers? Literature on colonial institutions, including (Acemoglu, Johnson and Robinson, 2001), (Banerjee and Iyer, 2005), (Michalopoulos and Papaioannou, 2013), and others, illustrates that the governance structures imposed by colonizers have had significant and enduring effects on economic development. Colonial institutions, often designed to benefit the colonizers at the expense of the local population, are frequently cited as key drivers of persistent underdevelopment. But are these historical legacies simply remnants of the past, or do they actively perpetuate economic disparities across regions? This is one of the main questions that (Dell, 2010) tries to answer.

Focusing on the *Mita* system, a coercive labor policy implemented by the Spanish in colonial Peru and Bolivia, Dell's study aims to determine which economic outcomes differ between regions affected by the *Mita* and those that were not. Her analysis reveals that regions subjected to the *Mita* system exhibit significantly lower levels of economic prosperity today compared to neighboring regions that were not affected, suggesting that historical institutions can have enduring effects. Additionally, Dell seeks to understand the underlying reasons for the persistence of these effects by leveraging historical data to explore how institutional legacies have influenced land tenure systems, local governance, and economic opportunities over time.

While many studies in development economics utilize field studies to establish causality, such approaches are not feasible when examining historical institutions. Researchers often turn to methods like Instrumental Variables (IVs) and Regression Discontinuity Designs (RDDs) to validate their findings. As noted by (Cattaneo and Titiunik, 2022), the RDD was introduced by Thistlethwaite and Campbell (1960) as a "method of testing causal hypotheses" in contexts where random assignment of treatment is not possible. The design applies to situations where units receive a score and treatment is assigned based on a specific rule linked to a known cutoff: units with scores above the cutoff receive treatment, while those below the cutoff serve as controls. Under appropriate assumptions, the discontinuous change in the probability of receiving treatment at the cutoff can help identify the causal effects of the treatment on an outcome of interest.

Dell's study utilizes a geographical or spatial RDD, arguing that the random assignment of districts to the administrative *Mita* boundary resulted in a quasi-random experiment. By isolating the causal effect of this colonial institution, Dell's methodology provides insights into the enduring impacts of colonial legacies. I replicate part of Dell's main results, focusing on the impact on current living standards and testing various functional forms for robustness. My findings broadly align with Dell's results. I further extend the study by conducting additional balance tests that match treatment and control units, examining mass points in the score (latitude and longitude), and re-estimating results using non-parametric methods. These methods employ local polynomial regressions with bias-corrected robust inference procedures, as proposed by (Calonico, Cattaneo and Titiunik, 2014).

The rest of this paper proceeds as follows. In the next section, I discuss the setting of the study, followed by an overview of the data and methodology in Section 3. Section 4 presents the replications, while Section 5 explores the extensions. Finally, Section 6 concludes.

2 The Mining *Mita*: Historical Context and Assignment Criteria

The *Mita* system, introduced by the Spanish in 1573, was a forced labor program that required indigenous communities in Peru and Bolivia to provide one-seventh of their adult male population to work in silver and mercury mines. This system, which operated until its abolition in 1812, was crucial for the Spanish Empire's mining operations, particularly in the lucrative Potosí and Huancavelica mines. The geographical scope of the *Mita* was not arbitrary; it was determined by several logistical and environmental factors. Indigenous villages within a designated region were compelled to participate, with the administrative costs of maintaining the *Mita* increasing with the distance from the mines. Additionally, the elevation of the region played a significant role in the assignment process. Spanish officials believed that only highland inhabitants, accustomed to living at high altitudes, could endure the strenuous physical demands of mining work. **??**, derived from the original source text, illustrates the Mita boundary and the study boundary.

The assignment of districts to the *Mita* was influenced by both distance from the mines and elevation. Regions closer to Potosí and Huancavelica, with more manageable distances and lower administrative costs, were more likely to be included in the *Mita*. Conversely, areas farther away faced higher costs and logistical challenges, which contributed to their exclusion from the program. Furthermore, highland areas were preferred for their presumed resilience to the harsh conditions of high-altitude mining. This selective assignment helps to explain the geographic distribution of *Mita* regions, as illustrated in Figure **??**.

Despite these criteria, there are concerns that other, less documented factors may have

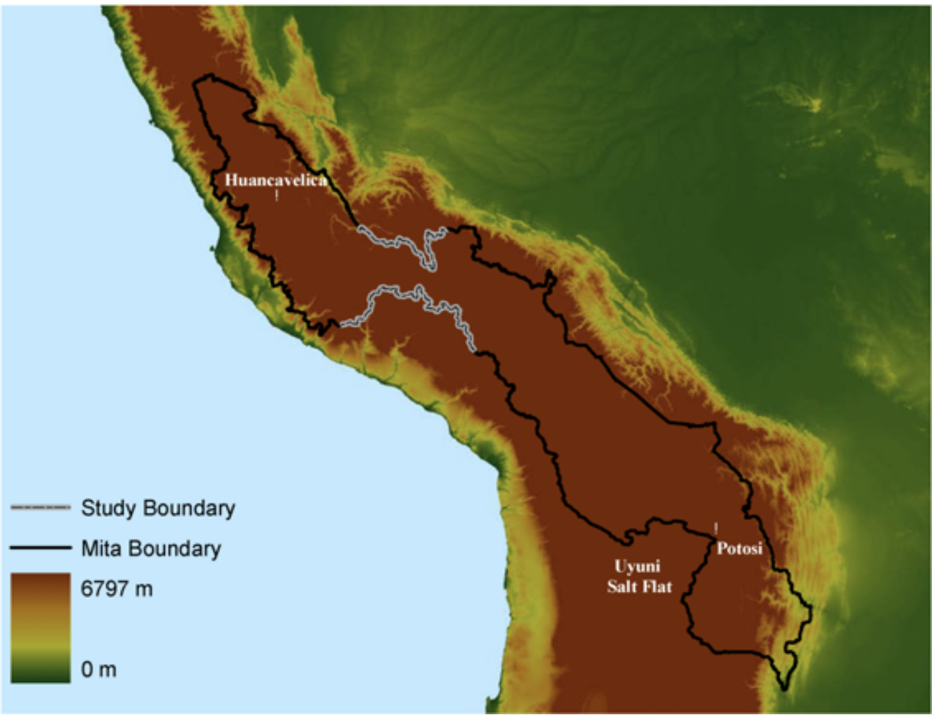


FIGURE 1.—The *mita* boundary is in black and the study boundary in light gray. Districts falling inside the contiguous area formed by the *mita* boundary contributed to the *mita*. Elevation is shown in the background.

influenced the assignment of the *Mita*. For example, local economic conditions, political factors, or pre-existing social structures could have played a role in determining which communities were subjected to forced labor. (Dell, 2010) provides evidence that the treatment and control groups are balanced across the border, and I bolster this evidence with additional tests. This is crucial to argue for the validity of the study design. In the next section, we will look at the data and the empirical strategy used in original study, and discuss the underlying assumptions.

3 Data and Empirical Strategy

3.1 Data

To evaluate the long-term impact of the *Mita* system on contemporary economic development, the study utilizes a range of historical and modern data sources. Historical data on *Mita* assignment, detailing which districts were subjected to forced labor, is obtained from Saignes (1984) and Amat y Junient (1947) and matched to modern districts. Living standards are measured using the 2001 Peruvian National Household Survey (Encuesta

Nacional de Hogares, ENAHO) for household consumption, adjusted to Lima metropolitan prices, and a microcensus dataset from the Ministry of Education, which records the heights of 6- to 9-year-old children, is used to assess nutritional status.

Geographic controls such as mean area-weighted elevation and slope for each district are obtained using NASA's Shuttle Radar Topography Mission (SRTM) data. These geographic metrics are also combined with district maps.

3.2 Empirical Strategy

The impact of the *Mita* system is assessed using a regression discontinuity design (RDD), capitalizing on the deterministic and discontinuous nature of *Mita* treatment as a function of known covariates, specifically geographic coordinates. Dell estimates causal effects by using a basic regression equation with different functional forms to provide robustness. We will first look at the RDD in greater detail, before examining the estimation strategy employed by the study paper.

3.2.1 RDD and underlying assumptions

The RDD, a popular research design that allows for the rigorous study of non-experimental interventions, is defined by three fundamental ingredients: a score (also known as a running variable, forcing variable, or index), a cutoff, and a treatment rule that assigns units to treatment or control based on a hard-thresholding rule. All units receive a score, and the treatment is assigned to units whose value of the score exceeds the cutoff and not assigned to units whose value of the score is below the cutoff. This assignment rule implies that the probability of treatment assignment changes abruptly at the known cutoff. If units are not able to perfectly determine or manipulate the exact value of the score that they receive, the discontinuous change in the probability of treatment assignment can be used to study the effect of the treatment on outcomes of interest, at least locally, because units with scores barely below the cutoff can be used as comparisons or "counterfactuals" for units with scores barely above it.

Building on this, the basic RDD, or the sharp RDD has three main features: (i) the score is continuously distributed and has only one dimension, (ii) there is only one cutoff, and (iii) compliance with treatment assignment is perfect. Over time, there have been various extensions of this design which modify one or more of these features. The study uses a geographic RDD, a special case of the multi-score RDD.

In the Multi-Score RD Design, two or more running variables determine the treatment assignment. With multiple running variables, there is no longer a single cutoff value at which the treatment status of units changes from control to treated; instead, the discontinuity in the treatment assignment occurs along a boundary of points. In the Geographic RD design, the score is a vector of two coordinates such as latitude and longitude that determine the exact geographic location of each unit. In practice, this score is calculated using Geographic Information Systems (GIS) software, which allows researchers to obtain the coordinates corresponding to the geographic location of each unit in the study, as well as to locate the entire treated and control areas, and all points on the boundary between them.

What are the assumptions that are required for identification of a parameter in this setup? First, we need a score/running variable that uniquely represents each unit's geographic location and allows us to compute its distance to any point along the RDD boundary, and this score is usually the coordinates (latitude and longitude) of the unit. We also assume that there are enough units in the neighbourhood of the boundary, an assumption that is sometimes restrictivein spatial RDDs. An important assumption is that of continuity in the two-dimensional score - the average potential outcomes under treatment and control should be continuous at all points on the boundary. This allows us to identify the treatment effect at each boundary point. Let us now look at the study paper's estimation.

3.2.2 Estimation Equation and Functional Forms

Dell (2010) uses a geographic discontinuity to study the effect of a forced labor mining system in Peru and Bolivia (the mita) on household consumption and other measures. Her approach incorporates geographic coordinates into the analysis, but is different from the current literature. Importantly, her estimation strategy uses a two-dimensional score (latitude and longitude) that is the same for all units contained in the same cluster: the specifications include a cubic polynomial in the latitude and longitude of each observation's district capital, but the unit of observation is the individual or household, with many individual observations contained in each district. It is important to note here that with clustered geographic coordinates one cannot define the individual-level treatment assignment as a deterministic function of this score, as in the traditional sharp RD, since there are no individual-level scores available, only cluster-level scores. Additionally, while Dell (2010) captures geographic treatment effect heterogeneity with boundary segment fixed effects, current literature focuses on the identification and estimation of treatment effects at every boundary point, which captures geographic heterogeneity in a more general way, as these effects for specific boundary points can always be integrated to obtain the effect for any larger segment.

Dell (2010) uses the following regression model to identify treatment effects:

$$c_{idb} = \alpha + \gamma Mita_d + X'_{id}\beta + f(geographic\ location_d) + \phi_b + \epsilon_{idb}, \tag{1}$$

where c_{idb} is the outcome variable for observation i in district d along segment b of the *Mita* boundary, and $Mita_d$ is an indicator for whether district d was subjected to *Mita*. X_{id} includes covariates such as mean elevation and slope, while $f(geographic\ location_d)$ represents the RD polynomial controlling for smooth variations in geographic location. Boundary segment fixed effects ϕ_b are included to account for variations within the *Mita* boundary. [1]

Continuity is one of the assumptions that the study paper spends a considerable time on, ensuring that districts just inside and outside the *Mita* boundary are comparable. Potential confounders, including elevation, terrain ruggedness, soil fertility, rainfall, ethnicity, preexisting settlement patterns, and local tribute rates, are examined in Table 1 for smoothness around the boundary, with results showing generally minimal differences in these characteristics.

To address the multidimensional nature of the *Mita* boundary, three RD specifications are used: a cubic polynomial in latitude and longitude, a cubic polynomial in Euclidean distance to Potosí, and a cubic polynomial in distance to the *Mita* boundary. This multidimensional RD approach allows for flexibility in modeling spatial effects, but raises concerns about overfitting and precise specification. Additionally, the study assumes that there is no selective sorting across the treatment threshold - if the *Mita* effect led to significant migration of productive individuals, it could influence the results. However,

[1] The analysis excludes metropolitan Cusco due to its distinct historical and economic context.

historical evidence suggests limited migration, and existing data show stable population distributions over time, which mitigates this concern. In the next section, I will go over the results that the study paper obtains for outcomes on living standards.

4 Results and Replications

In this section, I will discuss the main results of the paper using my replications. Broadly, I obtained the same results as the study paper, with some small differences in coefficient size and standard errors [2].

4.1 Balancing Tests

The study paper identifies the condition that all the characteristics of the treatment and control group vary smoothly at the *mita* cutoff as a critical assumption for the identification strategy, because this assumption ensures that individuals just oustisde the *mita* catchment can be an apporporiate counterfactual for those inside it. To assess the plausibility of this assumption, the paper presents summary statistics for sample characteristics within and outside of the mita boundaries for a variety of outcomes. Different distances from the *mita* boundaries are used to show that if there are any statistically significant differences, these disappear as we approach the boundary. I replicate this table and present my results below. To examine elevation and terrain ruggedness, the paper uses geographic data and divide the study region into $20 * 20$ km grid cells, approximately equal to the mean size of the districts in their sample, and calculate the mean elevation and slope within each grid cell[3]. Since these geographic data are spatially correlated, the study reports standard errors corrected for spatial correlation. I, however, report robust standard errors (see footnote 2).

The first row from the replicated Table 1 shows that elevation is statistically identical across the *mita* boundary and the second row shows that there are some statistically significant, but relatively small, differences in slope, with *mita* districts being less rugged.

[2]The study paper runs its analysis in Stata and I have replicated this analysis in R. There are some cases, for example when running weighted regressions, where it was not possible to exactly replicate the same result, or where the method for estimating standard errors differs.

[3]They report that all results are similar if the district is used as the unit of observation instead of using grid cells.

Table 1: Summary Statistics, Replicated

Sample Falls Within:	<100 km of Bound.			<75 km of Bound.			<50 km of Bound.			<25 km of Bound.		
	Inside	Outside	s.e.	Inside	Outside	s.e.	Inside	Outside	s.e.	Inside	Outside	s.e.
GIS Measures												
Elevation	4042.0555	4018.429	113.5467	4085.166	4103.661	109.0625	4117.112	4096.335	118.5089	4135.203	4059.598	148.6373
Slope	5.541806	7.206142	0.6556570***	5.747340	7.021311	0.6909635**	5.869653	6.947026	0.7702902*	5.771129	7.208883	1.0281249*
Observations	177	95		144	86		104	73		48	52	
% Indigenous	63.579	58.743	9.76	71.480	64.848	8.14	71.449	64.848	8.43	74.772	63.745	10.55
Observations	1112	366		831	330		683	330		329	251	
% 1572 tribute to												
Spanish Nobility	-1.313	-2.728	1.531***	-1.316	-2.96	1.682***	-1.623	-2.693	1.973***	-1.191	-3.091	2.468***
Spanish Priests	0.2105	0.1910	0.0108*	0.2190	0.1945	0.0115*	0.2059	0.1993	0.0109	0.2145	0.1998	0.0145
Spanish Justices	0.1336	0.1257	0.0052	0.1331	0.1246	0.0064	0.1281	0.1248	0.0061	0.1306	0.1237	0.0086
Indigenous Mayors	0.0567	0.0440	0.009	0.046	0.0429	0.004	0.0442	0.0447	0.004	0.04478	0.04428	0.0058
Observations	63	41		47	37		35	30		18	24	

Notes: The unit of observation is grid cells for the geospatial measures, the household for % indigenous, and the district for the 1572 budget data. In the first three columns, the sample includes only observations located less than 100 km from the boundary, and this threshold is reduced to 75, 50, and finally 25 km in the succeeding columns. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

The third row shows that are no statistically significant differences in ethnic identification across the *mita* boundary.

To argue that the *mita* boundary was determined as-if randomly, the paper provides historical evidence from 1572 showing good matching on tribute and budget characteristics across the border. I replicate the results only for budget characteristics, and in contrast to the first three rows, obtain very different results. My replication shows that not only are the coefficient estimates I obtain quite different, but also statistically significant for the tribute paid to the Spanish nobility. However, all other tributes are not different across the border, providing overall support to the continuity assumption. This divergence in results also provides another reason for additional balance tests, discussed in the next section on replication.

4.2 Results on Living Standards

The study paper is interested in the impact of the *Mita* today on both economic outcomes (proxied with household consumption from 2001) and outcomes in terms of health (proxied with stunted growth in children from 2005). These results are presented in Table 2 of the paper, which I replicate here. I obtain broadly the same coefficients and significance levels[4] and therefore can support the conclusions from the paper.

Columns 1–3 of Table II estimate that a long-run *mita* effect lowers household consumption in 2001 by around 25 - 30% in subjected districts. The point estimates remain fairly stable as the sample is restricted to fall within narrower bands of the *mita* boundary.

[4] I obtain numerically equal coefficients for all models, but more of my estimates are significant than the study paper and sometimes at a higher level

Moreover, the *mita* coefficients are economically similar across the three specifications of the RD polynomial. All of the *mita* coefficients in panels B and C, which report the single-dimensional RD estimates, are statistically significant at the 1% level. In contrast, the point estimates using a cubic polynomial in latitude and longitude (panel A) are not strongly significant until quite close to the boundary. The study paper points to the relative flexibility of the specification, the small number of observations and clusters (the household survey samples only around one-quarter of districts), and measurement error in the dependent variable as probable reasons.

Columns 4–7 of Table II examine census data on stunting in children, which offers a substantially larger sample. When using only observations in districts that border the *mita* boundary, point estimates of the *mita* effect on stunting range from 0.055 to 0.114 percentage points. This compares to a mean prevalence of stunting of 40% throughout the region examined. All the 12 point estimates reported in Table II for height are statistically significant.

Table 2: Living Standards, Replicated

	Dependent Variable						
	Log Equiv. Household Consumption (2001)			Stunted Growth, Children 6-9 (2005)			
Sample Within:	<100 km of Bound. (1)	<75 km of Bound. (2)	<50 km of Bound. (3)	<100 km of Bound. (4)	<75 km of Bound. (5)	<50 km of Bound. (6)	Border District (7)
Panel A: Cubic Polynomial in Latitude and Longitude							
Mita	-0.284*	-0.216	-0.331**	0.070***	0.084***	0.086***	0.114***
	(0.122)	(0.124)	(0.125)	(0.005)	(0.006)	(0.006)	(0.008)
R^2	0.059	0.060	0.069	0.051	0.019	0.017	0.050
Panel B: Cubic Polynomial in Distance to Potosí Mita							
Mita	-0.337***	-0.307***	-0.329***	0.080***	0.078***	0.078***	0.063***
	(0.069)	(0.071)	(0.068)	(0.004)	(0.004)	(0.004)	(0.006)
R^2	0.046	0.036	0.047	0.049	0.017	0.013	0.047
Panel C: Cubic Polynomial in Distance to Boundary							
Mita	-0.277***	-0.230***	-0.224***	0.073***	0.061***	0.064***	0.055***
	(0.067)	(0.067)	(0.064)	(0.003)	(0.004)	(0.004)	(0.006)
R^2	0.044	0.042	0.040	0.040	0.015	0.013	0.043
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Boundary F.E.s:	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clusters:	71	60	52	289	239	185	63
Observations:	1478	1161	1013	158848	115761	100446	37421

Notes: The unit of observation is the household in columns 1–3 and the individual in columns 4–7. Robust standard errors, adjusted for clustering by district, are in parentheses. The dependent variable is log Mita equivalent household consumption (ENAHO (2001)) in columns 1–3, and a dummy equal to 1 if the child has stunted growth and equal to 0 otherwise in columns 4–7 (Ministro de Educación (2005a)). *Mita* is an indicator equal to 1 if the household's district contributed to the and equal to 0 otherwise. Panel A includes a cubic polynomial in the latitude and longitude of the observation's district capital, panel B includes a cubic polynomial in Euclidean distance from the observation's district capital to Potosí, and panel C includes a cubic polynomial in Euclidean *Mita* distance to the nearest point on the boundary. All regressions include controls for elevation and slope, as well as boundary segment fixed effects. Columns 1–3 include demographic controls for *Mita* the number of infants, children, and adults in the household. In columns 1 and 4, the sample includes observations whose district capitals are located within 100 km of the boundary, and this threshold is reduced to 75 and 50 km in the succeeding columns. Column 7 includes only observations whose districts border the boundary. 78% of the observations are in districts in column 1, 71% in column 2, 68% in column 3, 78% in column 4, 71% in column 5, 68% in column 6, and 58% in column 7. Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

4.3 Prosperity, Institutions and Demographics in 1572

To provide support for the RDD estimation strategy identifying the *mita's* long-run effect as opposed to some other underlying difference, the study paper checks whether economic prosperity, institutions, or demographics prior to the mita's enactment differ across the border. In a series of specification checks, they examine the shares of 1572 tribute revenues allocated to rents for Spanish nobility, salaries for Spanish priests, salaries for local Spanish administrators, and salaries for indigenous mayors to account for differences in prosperity levels and the presence of institutions. They also examine whether demographics differ using the population shares of tribute paying males (those aged 18–50), boys, and women as the dependent variables. These regressions, replicated in Table V, do not show statistically significant differences across the *mita* boundary, and the estimated *mita* coefficients are small. I obtain different coefficients for the share of tributes to Spanish nobility, for which I also had results deviating from the study paper in Table 1, pointing to a difference in the definition of the variable. However, this does not affect results[5].

Table 3: 1572 Tribute and Population, Replication of Table V

	Dependent Variable							
	Share of tribute revenues					Percent		
	Log Mean Tribute (1)	Spanish Nobility (2)	Spanish Priests (3)	Spanish Justices (4)	Indig. Mayors (5)	Men (6)	Boys (7)	Females (8)
	Panel A: Cubic Polynomial in Latitude and Longitude							
Mita	0.020	2.628	0.004	0.003	0.003	0.006	0.011	-0.009
	(0.027)	(2.50)	(0.018)	(0.010)	(0.005)	(0.008)	(0.011)	(0.014)
R^2	-	0.163	0.090	0.228	0.266	-	-	-
	Panel B: Cubic Polynomial in Distance to Potosí							
Mita	0.019	0.294	0.008	0.006	-0.001	-0.012	0.005	-0.011
	(0.027)	(1.870)	(0.013)	(0.008)	(0.005)	(0.008)	(0.009)	(0.011
R^2	-	0.100	0.073	0.151	0.132	-	-	-
	Panel C: Cubic Polynomial in Distance to Boundary							
Mita	0.040	1.262	0.005	0.003	0.001	0-0.011	0.001	-0.008
	(0.027)	(1.601)	(0.012)	(0.007)	(0.004)	(0.007)	(0.007)	(0.010)
R^2	-	0.114	0.039	0.118	-	-	-	
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Boundary F.E.s	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	65	65	65	65	Yes	65	65	65

Notes: The dependent variable in column 1 is the log of the district's mean 1572 tribute rate. In columns 2–5, it is the share of tribute revenue allocated to Spanish nobility, Spanish priests, Spanish justices, and indigenous mayors (caciques), respectively. In columns 6–8, it is the share of 1572 district population composed of males (aged 18–50), boys, and females (of all ages), respectively. Panel A includes a cubic polynomial in longitude and latitude, panel B includes a cubic polynomial in Euclidean distance from the observation's district capi- tal to Potosí, and panel C includes a cubic polynomial in Euclidean distance to the nearest point on the mita boundary. All regressions include geographic controls and boundary segment fixed effects. The samples include districts whose capitals are less than 50 km from the mita boundary. Column 1 weights by the square root of the district's tributary population and columns 6–8 weight by the square root of the district's total population. 66% of the observations are from mita districts Coefficients that are significantly different from zero are denoted by the following system: *10%, **5%, and ***1%.

4.4 Other Results from the Paper

Going beyond these findings, the paper also explores additional channels of persistence related to land tenure and public goods provision which I do not replicate here. No-

[5] Note that I do not provide R-squared measures for all models estimated here. This is because it was not possibly to obtain these directly from the model output and the ones I did obtain did not differ from the study paper.

tably, the *Mita* system had a significant effect on the formation and concentration of large land estates, or haciendas, which were statistically less prevalent in *Mita* districts compared to non-*Mita* districts in 1689, 1845, and 1940. This persistent impact is also evident in educational outcomes. Results from 1876, 1940, and 2001 show lower literacy rates and mean years of schooling in *Mita* districts, although these differences are not highly persistent over time. Additionally, the study finds that *Mita* districts exhibit lower integration into road networks, with a significantly reduced density of local and regional roads, including paved and gravel roads, observed as late as 2006. Lastly, the paper discusses the effects on proximate determinants of consumption, such as the labor force, market participation, and the prevalence of subsistence farming, highlighting the broader economic and social consequences of the *Mita* system.

5 Extensions

In this section, I extend the validity of the RDD framework in several ways. First, I include additional falsification tests with a different underlying methodology and compare them to results obtained using the same method as in the study paper. I then re-estimate coefficients from 2 using local linear polynomial regressions. I also include histograms of chordal distances near the boundary and end with a discussion on the presence of mass polynomials in the data.

5.1 Balancing tests or Falsification tests

It is common in the spatial RDD literature to provide empirical evidence that the continuity assumption holds. Most common is a series of tests on predetermined or pretreatment outcomes, but also common are balance tests which investigate whether the mean (or other feature of the distribution) of the covariates is statistically indistinguishable between treated and control units near the boundary. However, as (Keele and Titiunik, 2015) note, assessing covariate balance in a spatial RDD may require a different strategy since covariate balance is a spatial construct. To this end, they suggest matching each treated unit with its closest/nearest control unit in terms of geographic distance[6] and

[6]Ties are broken randomly

Table 4: Balancing Table, Pre-matching

Variable	MeanT	MeanC	Pval ttest	Var ratio	Pval KS	QQ med diff
Elevation	3.867	3.741	0	0.539	0	0.131
(std. error)	(0.01)	(0.019)	NA	NA	NA	NA
Slope	6.277	8.238	0	0.801	0	0.124
(std. error)	(0.112)	(0.182)	NA	NA	NA	NA
Infants	0.51	0.42	0.013	1.302	0.105	0.023
(std. error)	(0.022)	(0.028)	NA	NA	NA	NA
Children	1.221	1.072	0.024	1.24	0.05	0.015
(std. error)	(0.04)	(0.052)	NA	NA	NA	NA
Adults	2.549	2.643	0.187	0.804	0.238	0.006
(std. error)	(0.037)	(0.061)	NA	NA	NA	NA
Sample size (T,C)	1112	526	NA	NA	NA	NA

Notes: The table presents balance test results for various variables between treatment and control groups. - **Pval ttest**: p-value from a t-test comparing means of the variable between treatment and control groups. A value less than 0.05 indicates a significant difference in means. - **Var ratio**: Ratio of the variances between the treatment and control groups. A ratio close to 1 suggests similar variability. - **Pval KS**: p-value from the Kolmogorov-Smirnov test comparing the distributions of the variable between treatment and control groups. A value less than 0.05 indicates a significant difference in distributions. - **QQ med diff**: Median difference from the quantile-quantile (QQ) plot, reflecting differences in the distribution's central tendency. NA indicates values that are not applicable or not provided for specific rows.

then applying standard balance tests to this spatially matched data. This overcomes a shortcoming of the balance tests that compare averages based on distances around the discontinuity - this balance may change as we move along the border i.e. while balance may hold along one section of the border, it may not hold along another section. Such geographic heterogeneity will be missed in a geographically naive balance test. Using the same covariates as those used in in 2, I perform both the naive and post-matching balancing tests as suggested by (Keele and Titiunik, 2015). Tables 4 and 5 summarise the results.

From the tables, we conclude that both in the pre-matching and post-matching balance tests, there is no significant difference in means and distributions in elevation and infants across groups, but slope, children and adults have significant differences in both means and differences. This casts some doubts on the RDD, but it is important to explore this further as his balance test, like all balance tests, is not a strict hypothesis test based on statistical significance. Moreover, there were only 164 unique control units matched to treatment groups, which could bias results.[7]

[7] (Keele and Titiunik, 2015)'s paper had just 6 matched units with a similar sample size, but perhaps then this indicates the possibility for improvement in matching criteria

Table 5: Balance Table, Post-Matching

Variable	**MeanT**	**MeanC**	**Pval ttest**	**Var ratio**	**Pval KS**	**QQ med diff**
Elevation	3.867	3.842	0.235	0.257	0	0.168
(std. error)	(0.01)	(0.019)	NA	NA	NA	NA
Slope	6.277	9.239	0	0.454	0	0.145
(std. error)	(0.112)	(0.167)	NA	NA	NA	NA
Infants	0.51	0.514	0.879	1.308	0.01	0.024
(std. error)	(0.022)	(0.019)	NA	NA	NA	NA
Children	1.221	1.397	0.001	1.231	0	0.012
(std. error)	(0.04)	(0.036)	NA	NA	NA	NA
Adults	2.549	2.271	0	1.079	0	0.005
(std. error)	(0.037)	(0.036)	NA	NA	NA	NA
Sample size (T,C)	1112	526	NA	NA	NA	NA

Notes: The table presents balance test results for various variables between treatment and control groups. - **Pval ttest**: p-value from a t-test comparing means of the variable between treatment and control groups. A value less than 0.05 indicates a significant difference in means. - **Var ratio**: Ratio of the variances between the treatment and control groups. A ratio close to 1 suggests similar variability. - **Pval KS**: p-value from the Kolmogorov-Smirnov test comparing the distributions of the variable between treatment and control groups. A value less than 0.05 indicates a significant difference in distributions. - **QQ med diff**: Median difference from the quantile-quantile (QQ) plot, reflecting differences in the distribution's central tendency.
NA indicates values that are not applicable or not provided for specific rows.

5.2 Re-estimating results with local linear polynomial regression

Following (Cattaneo, Idrobo and Titiunik, 2023)'s discussion on state of the art estimation techniques for RDD, I attempt re-estimating the effect of the *Mita* on household consumption and the stunting of growth in children using a non-parametric estimation.

Tables 6 and 7 show the results for a local linear regression of **consumption** on the distance from the *mita* boundary. Conventional estimates are based on local linear regression without bias correction, while robust estimates adjust for bias through data-driven bandwidth selection. The results show a negative treatment effect that is statistically significant at the cutoff, even more so after including covariates. In both estimations we see that while the estimated effect remains negative, the result becomes less significant as we move towards robust estimates, guarding against false positives.

Table 6: RD Estimates using Local Polynomial Regression

Method	Coefficient	Std. Error	z-Value	P-Value	95% CI Lower	95% CI Upper
Conventional	-0.709	0.223	-3.173	0.002	-1.146	-0.271
Bias-Corrected	-0.629	0.223	-2.817	0.005	-1.067	-0.191
Robust	-0.629	0.330	-1.908	0.056	-1.275	0.017

Tables 8 and 9 show the same results but for height and stunting in children. What is interesting here is that the estimates suggest that being in a *Mita* district decreases stunting and increases the height of children, which is opposite to what the study paper's

Table 7: Covariate-Adjusted RD Estimates using Local Polynomial Regression

Method	Coefficient	Std. Error	z-Value	P-Value	95% CI Lower	95% CI Upper
Conventional	-1.193	0.293	-4.069	0.000	-1.767	-0.618
Bias-Corrected	-1.065	0.293	-3.633	0.000	-1.640	-0.491
Robust	-1.065	0.397	-2.681	0.007	-1.844	-0.286

estimation strategy reports. However, these estimates are insignificant and thus further validate the study paper's results.

Table 8: Covariate-adjusted RD estimates using local polynomial regression for Height of Children

Method	Coefficient	Std. Error	z	P-value	[95% C.I.]
Conventional	0.306	0.270	1.134	0.257	[-0.223 , 0.836]
Bias-Corrected	0.162	0.270	0.600	0.549	[-0.367 , 0.691]
Robust	0.162	0.319	0.508	0.611	[-0.463 , 0.787]

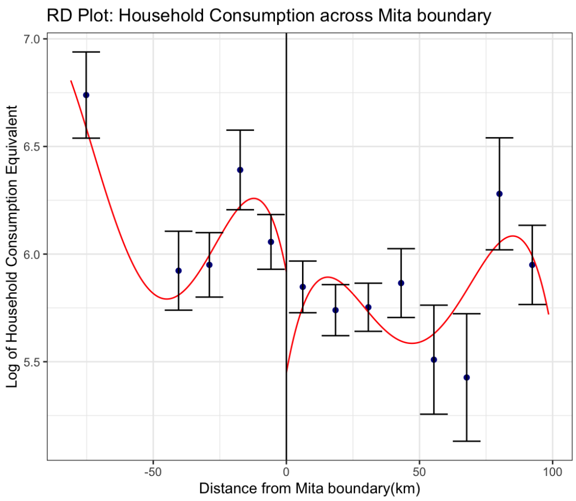
Table 9: Covariate-adjusted Sharp RD estimates using local polynomial regression for Stunting in Children

Method	Coefficient	Std. Error	z	P-value	[95% C.I.]
Conventional	-0.015	0.018	-0.848	0.396	[-0.049 , 0.020]
Bias-Corrected	-0.012	0.018	-0.673	0.501	[-0.046 , 0.023]
Robust	-0.012	0.021	-0.577	0.564	[-0.052 , 0.028]

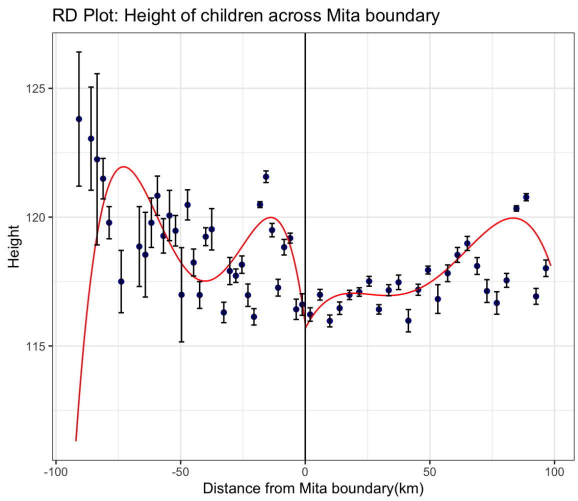
The accompanying RD plots (Figures 1a,1b and 1c) show the variation in outcome variable for the local linear polynomial regressions.

5.3 Presence of mass points

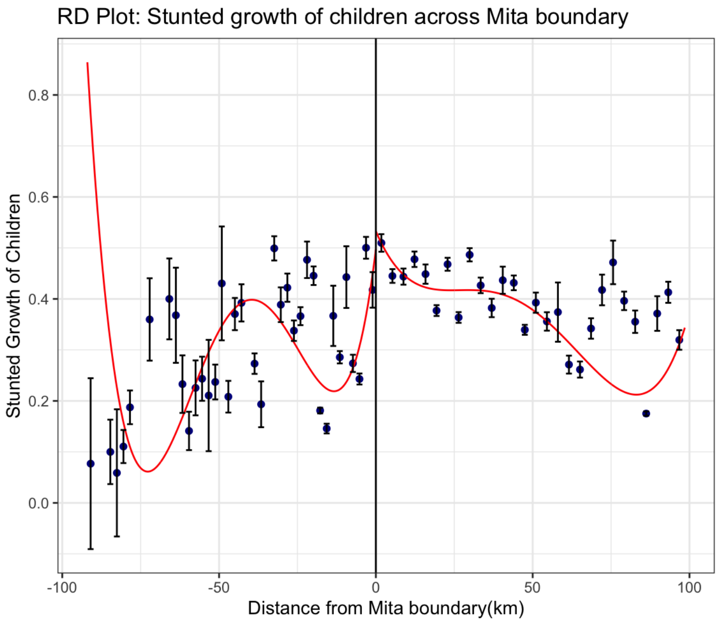
Because our running variable, co-ordinates, are discrete, they have mass points i.e. more than one unit is located at the same point. This reduces the effective number of observations. For example, in the consumption regression with covariates, the effective number of observations reported in the results decreases from 526 to 189 control units and from 1112 to 140 treatment units. For the stunting and height samples, this reduction is stronger, with control units going from 61047 to 11805 (5-fold reduction) and treatment units from 124614 to 11523 (almost 11 times). The differences in point estimates have already been discussed.



(a) Variation of household consumption near *Mita boundary*



(b) Variation in heights of children near *Mita boundary*



(c) Probability of stunting in children near *Mita boundary*

6 Conclusion

In this term paper, I have replicated the results of Dell (2010), a seminal paper in the development economics and RDD literature which estimates the causal effects of a colonial institution, the *Mita* on present day outcomes and channels of persistence. I focused on present day outcomes in affected areas and in general, my replication results were aligned with the main results of the paper. Given the advancements in the field since the publication of this paper, my extensions focused on new methods.

My first extension extended balance tests along another dimension by creating spatially matched treatment - control units. The results of this extension were mixed, perhaps because of incomplete matching, with some important covariates being statistically same across the treatment boundary and others not. My second extension re-estimated the results on consumption and stunting, included actual height of children as a new dependent variable. Broadly, these re-estimations support the study paper's results, but do point towards some intricacies in the spread of the variables across the distance from the boundary. It could also be interesting to look at the channels of persistence that the paper uncovers through the lenses of these extensions.

References

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91(5):1369–1401.

Ashraf, Quamrul and Oded Galor. 2011. "Dynamics and Stagnation in the Malthusian Epoch." *American Economic Review* 101(5):2003–2041.

Banerjee, Abhijit and Lakshmi Iyer. 2005. "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review* 95(4):1190–1213.

Calonico, Sebastian, Matias D Cattaneo and Rocio Titiunik. 2014. "Robust data-driven inference in the regression-discontinuity design." *The Stata Journal* 14(4):909–946.

Cattaneo, Matias D, Nicolás Idrobo and Rocio Titiunik. 2023. "A practical introduction to regression discontinuity designs: Extensions." *arXiv preprint arXiv:2301.08958* .

Cattaneo, Matias D and Rocio Titiunik. 2022. "Regression discontinuity designs." *Annual Review of Economics* 14(1):821–851.

Dell, Melissa. 2010. "The persistent effects of Peru's mining mita." *Econometrica* 78(6):1863–1903.

Dell, Melissa, Nathan Lane and Pablo Querubin. 2018. "The Historical State, Local Collective Action, and Economic Development in Vietnam." *Econometrica* 86(6):2083–2121.

Dell, Melissa and Pablo Querubin. 2018. "Nation Building through Foreign Intervention: Evidence from Discontinuities in Military Strategies." *Quarterly Journal of Economics* 133(2):701–764.

Keele, Luke J and Rocio Titiunik. 2015. "Geographic boundaries as regression discontinuities." *Political Analysis* 23(1):127–155.

Michalopoulos, Stelios and Elias Papaioannou. 2013. "Pre-colonial Ethnic Institutions and Contemporary African Development." *Econometrica* 81(1):113–152.

Nunn, Nathan. 2008. “The Long-Term Effects of Africa’s Slave Trades.” *Quarterly Journal of Economics* 123(1):139–176.

Voigtländer, Nico and Hans-Joachim Voth. 2013. “How the West ’Invented’ Fertility Restriction.” *American Economic Review* 103(6):2227–2264.