Dell (2010), The Persistent Effects of Peru's Mining *Mita*: Replication and Extension to Test Spatial Auto-Correlation

Joshua Bailey

https://github.com/joshua-bailey/plsc536_finalproject.git

December 20, 2023

Abstract

Melissa Dell's seminal paper (Dell 2010) utilises regression discontinuity (RD) to examine the long-run impacts of the *mita*, a forced mining labour system in effect in colonial Peru and Bolivia between 1573 and 1812. The results show a *mita* effect lowers household consumption by around 25% and increases the prevalence of stunted growth in children by around 6 percentage points in subjected districts today. I extend Dell's analysis to explore whether the results are robust to spatial auto-correlation. Through the application of Moran statistic analysis, I show that the results may exhibit the characteristics of spatial auto-correlation and that further analysis is needed to ensure the results are robust to this critique.

1 Introduction

The role of historical institutions in shaping present day outcomes is a central debate across the social sciences. With the advent of the credibility revolution (Angrist and Pischke 2010), social scientists have made significant progress towards quantifying these effects (Acemoglu, Johnson, and Robinson 2001) (Acemoglu, Johnson, and Robinson 2002) (Nunn 2008). Much of the literature focuses on establishing a link between extractive economic institutions and long-run economic prosperity (Acemoglu, Johnson, and Robinson 2002) (Banerjee and Iyer 2005) (Glaeser et al. 2004).

(Dell 2010) expands the literature by focusing on the channels of this historical persistence. To do this, Dell examines the long-run impacts of the mining *mita*, a colonial forced labour system instituted by the Spanish government in Peru and Bolivia at the end of the 16th century, which continued for over 200 years, before being abolished in 1812. The *mita* compelled 200 indigenous communities to send one-seventh of their adults male population to work in the Potosí silver and Huancavelica mercury mines. The Potosí mines, founded in 1545, contained the largest silver deposits in

colonial Spain. The scale of the *mita* system is important. Dell calculates that around 3% adult men who have ever lived within the contemporary boundaries of Peru were conscripted into the *mita* at some point. The boundaries between communities who were subject to this practice and both mines are shown in Figure 2 in the Appendix (5).

The discrete boundary between otherwise similar geographic locations lends itself to analysis using a regression discontinuity (RD) design. Developing traditional RD techniques, Dell using geographic boundaries as a multidimensional discontinuity in longitude-latitude space. The results show a significant negative effect of the *mita* on contemporary outcomes, lowering equivalent household consumption by around 25% in subjected districts. Dell repeats the model for childhood stunting, a key determinant of future health outcomes, and finds it is 6 percentage points higher in subjected districts today.

While not discussed in detail in this paper, Dell goes on to hypothesise around channels for this persistence, including the role of *hacidendas* - rural estates with an attached labour force - which the colonial administration only allowed the development of outside the *mita* areas, to avoid competing for scarce labour resources within the areas. Building on (Engerman and Sokoloff K 1997), who argued that high historical inequality and the lack of a well-established landed class, lowered investment in public goods in Latin America, Dell posits that these institutions formed the basis for future property rights and land reform crucial for development. Dell also argues that the *mita* effect lowered education standards and excluded areas from future road network development.

2 Analysis

2.1 Data

The data used throughout is taken from the replication package for (Dell 2010).[1] The list of districts subjected to the *mita* system comes from (Saignes 1984) and (Amat Y Junient 1947). Dell matches this data to modern district definitions by tracking the historical evolution of sub-districts (*anexos*), many of which exist today.

Living standards are measured using two data sets. Household consumption data are from the 2001 Peruvian National Household Survey (Encuesta Nacional de Hogares (ENAHO)) collected by the National Institute of Statistics (INEI). The consumption data is net of transfers and normalised to the price level in the capital, Lima. The second data set is census data from the Ministry of Education covering heights of all 6- to 9-year-old school children in the region. Stunting is defined as children who are more than two standard deviations below the age-specific median, per

---

1. The full package can be accessed at: https://dell-research-harvard.github.io/projects/498mita

the method used by the WHO. Both data sets are matched geographically to the districts. While both data sets are related (lower levels of consumption will correlate with malnutrition and therefore stunting), the heights sample is much larger than the consumption data.

In addition to the geographic functions (see 2.2 below), Dell includes a series of controls for other geographic factors. This ensures the results are robust to geographic factors, such as weather, that are independent of longitudinal and latitudinal location. These controls are generated by computing the mean weighted elevation over a given area, using map data from NASA's Shuttle Radar Topography Mission (SRTM). The same approach is followed to obtain a value for the weighted mean of the slope of a given area. These controls are important because one of the criteria used to decide which areas should be *mita* districts was elevation (as well as proximity to the two mines), as the colonial authorities believed only highland people could survive the intense physical labour in the mines.

2.2   Model

Dell applies the classic RD approach but expands its application to geographic space. Whether or not an entity is assigned to the *mita* area is a discontinuous function in geographic space, which lends itself to estimation within the RD framework. This is different to traditional RD in that there isn't a single running variable, but rather a two-dimensional *mita* boundary in geographic space. The estimating equation therefore takes the form:

$$c_{idb} = \alpha + \gamma \, mita_d + X'_{id}\beta + f(\text{ geographic location }_d) + \phi_b + \varepsilon_{idb} \quad (1)$$

where c_{idb} is the outcome variable of interest, either consumption or height, for observation i in district d along segment b of the *mita* boundary; $mita_d$ is the independent variable of interest and is a dummy equal to 1 if the district d contributed to the *mita* system and equal to 0 otherwise; X_{id} is a vector of covariates used to control for other effects, including elevation and slope in a given district and demographic variables, including the number of infants, children and adults in the household; $f(\text{geographic location}_d)$ is a polynominal function for geographic location, ϕ_b is a set of boundary segment fixed effects, that indicate which of the boundaries are closest to the observation's district capital. Dell excludes metropolitan Cusco from the analysis (which includes seven non-*mita* and two *mita* districts. Cusco was the capital of the Inca Empire and therefore relatively wealthy throughout the historical period in question.

The geographic function in the models collapses distance in latitudinal and longitudinal space using a cubic polynomial. Dell's main specification uses a cubic polynomial in distance to the *mita* boundary, as per the RD specification. Two other specifications are included as baselines. These function operate less like traditional RD and use geographic distance more as a standalone control. The first of these specifications utilises a cubic polynomial in latitude and longitude. To collapse the two-dimensional

3

latitude and longitude (x and y) into a single dimension, the polynomial takes the form $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. The second is a cubic polynomial in distance to Potosí, which was the largest city in the Western Hemisphere during colonial era. A function of distance to Potosí is therefore likely to capture variation in unobserved agglomeration benefits.

RD, applied in this context, requires two identifying assumptions. First, RD assumes that all relevant factors except the treatment vary continuously at the cutoff. Within the model, this means assuming that c_1 and c_0 (denoting potential outcomes under treatment and control), and therefore $E[c_1|x, y]$ and $E[c_0|x, y]$ are continuous at the cutoff, where x and y are longitude and latitude respectively.

To ensure this assumption holds, Dell focuses on the portion of the *Mita* area that cuts across the Andean range in Southern Peru (the area shown in grey in Figure 2). This segment, unlike much of the *mita* boundary, doesn't follow the edge of the range, and so elevation, and other factors like ethnicity, are statistically identical across the study boundary. Dell further checks for continuity across a range of other variables collected just prior to the institution of the *mita* in 1573, including local tribute (tax) rates and other socio-demographic factors.

I have replicated these results in Table 1, which shows means for a series of geographic and demographic variables either side of the *mita* boundary. Observations are averages calculated for a 20km x 20km grid cell. Robust standard errors for the difference in means are reported in the third column of each specification. These are computed over a range of geographic distances from the *mita* boundary, at <100km, <75km and <50km (distances also used in the analyses). The first row shows that elevation is statistically identical across the boundaries. Slope shows some, small, differences, with *mita* districts being less rugged (slopped). Demographically, there is no statistically significant difference in the concentration of indigenous peoples across the boundary.

Differences in governance, as estimated by rates of tribute tax, also show broad continuity. The fourth row of Table 1 shows the average tribute contribution per adult male. These are important because one explanation that could be offered for differing contemporary outcomes across *mita* boundaries could be differing levels of economic extraction, measured here by taxation. The remaining rows (fifth through eighth) shows how this taxation was allocated to different constituencies, across salaries for Spanish priests, judges and mayors and rents for Spanish nobility. Here Table 1 shows some differences at wider apertures (<100km) but they disappear at the lower resolution.

The second assumption for RD in this context is that there is no selective sorting at the cutoff. Here, this assumption would be violated if there was substantial migration across the *mita* boundary that resulted from the effect of *mita* assignment itself. The absence of good historical migration data makes this assumption difficult to test. Dell instead notes that general migration rates since 1876 have been low and similar across *mita* and non-

mita districts. While this does not exclude the possibility of sorting, Dell argues that the relatively closed indigenous communities of the colonial era further makes substantial sorting unlikely (Morner 1978). Insofar as the *mita* did drive migration patterns, Dell internalises this and argues that it could form part of the effect seen in the analysis.

3   Results

3.1   Main Results

The main results are shown at Table 2. Columns 1 to 3 employ household consumption, measured in log of equivalent household consumption, net of transfers, in 2001, as the dependent variable. Columns 4 to 7 use a dummy variable for childhood growth stunting as the dependent variable. Panel A uses a cubic polynomial in latitude and longitude, Panel B uses a cubic polynomial in distance to Potosí, and Panel C uses a cubic polynomial in distance to the *mita* boundary. The results are reported by different bandwidths, <100km, <75km, and <50km from the *mita* boundary. Robust clustered standard errors at the district level are used, in part to account for the potential effect of spatial auto-correlation (discussed further in 3.3).

In Columns 1 to 3, the overall finding is that the long-run effect of the *mita* is to lower household consumption in 2001 by around 25% in treated districts. Focusing on the main Panel C results, the estimate is relatively consistent across bandwidths. All of the coefficients are significant at the 1 or 5% level. Model fit of 4-4.5% is significant in the context of the persistence literature. An historic institution that hasn't existed for more than 200 years explaining 4% of contemporary living standards demonstrates the substantial effect that such systems had and *continue* to have. The effect can also be seen in Columns 4 to 7, where Panel C shows a higher probability of modern day stunted growth of between 0.055 and 0.073. The larger sample size also facilities the models in Column 7, that restricts the sample to just regions bordering *mita* districts. This gives perhaps the clearest picture of the effect, where the probability of contemporary child growth stunting is 0.055 higher in areas that used to be *mita* districts, compared to those that weren't.

The main results are also shown in Figure 1, which shows a single-dimensional RD plot, with distance to the *mita* boundary as the running variable. The cubic polynomial is shown, as well as quadratic and linear specifications for comparison (and as Dell includes in robustness checks).[2]

---

2. While it appears the data does not run up to the cutoff, this is an artifact of how Dell constructs averages within districts for the purpose of graphing, such that the border districts are measured from a central point within the area.

3.2 Table 1: Summary Statistics

Replicated Table 1 from (Dell 2010)

Geographic and Social Descriptive Statistics

| Variable | Distance from *Mita* boundary | | | | | | | | | | | |
| | $\leq 100km$ | | | $\leq 75km$ | | | $\leq 50km$ | | | $\leq 25km$ | | |
	Inside	Outside	s.e.	Inside	Outside	s.e.	Inside	Outside	s.e.	Inside	Outside	s.e.
Elevation	4042.06	4018.43	86.03	4085.17	4103.66	83.26	4117.11	4096.34	90.24	4135.20	4059.60	116.36
Slope	5.54	7.21	0.49	5.75	7.02	0.52	5.87	6.95	0.59	5.77	7.21	0.80
% Indigenous	0.63	0.58	0.03	0.71	0.65	0.03	0.71	0.65	0.03	0.74	0.63	0.04
Log 1572 tribute rate	1.61	1.60	0.03	1.62	1.61	0.04	1.63	1.60	0.04	1.66	1.60	0.03
Spanish Priests	0.21	0.19	0.01	0.22	0.19	0.01	0.21	0.20	0.01	0.21	0.20	0.01
Spanish Justices	0.13	0.13	0.00	0.13	0.12	0.01	0.13	0.12	0.01	0.13	0.12	0.01
Indigenous Mayors	0.06	0.04	0.01	0.05	0.04	0.00	0.04	0.04	0.00	0.04	0.04	0.00
Spanish Nobility	0.60	0.64	0.01	0.60	0.64	0.02	0.62	0.63	0.01	0.61	0.63	0.02
Observations	63	41		47	37		35	30		18	24	

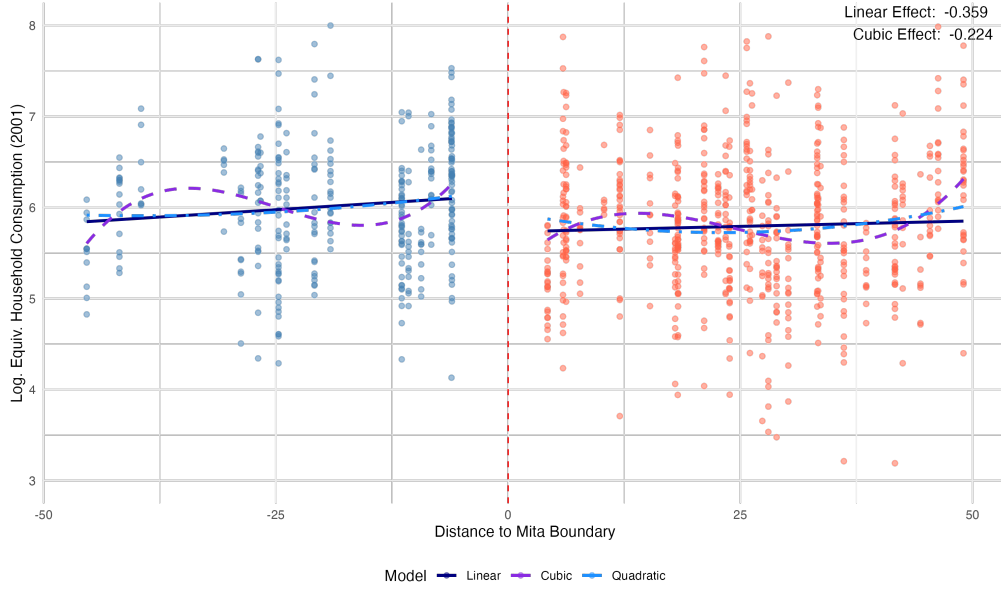Figure 1: Regression Discontinuity Plot, Household Consumption, <50km

Figure 1: RD plot showing the effect of the *mita* on household consumption.

3.3 Extension: Testing for Spatial Auto-correlation

One common (one could say, 'persistent') critique of the historical persistence literature is that the results are not robust to spatial auto-correlation. In the context of this literature, the problem of spatial auto-correlation arises when the spatial arrangement of historical variables (like institutions, events, or policies) influences current outcomes in a correlated manner across geographic areas. This can lead to biased estimates if the spatial dependencies are not appropriately accounted for. In the words of Tobler's First Law of Geography, "everything is related to everything else, but near things are more related than distant things" (Tobler 1970). For instance, given that income is higher in Europe than Africa, any variable that is high in the former and low in the latter will appear to explain growth and poverty. In the context of (Dell 2010), this is critical because the effect Dell argues for is institutional rather than geographic. It is the system of extraction created by the *mita* that drives contemporary outcomes, not where the *mita* were *per se*. Spatial auto-correlation occurs when the regression residuals display a high level of correlation. Spatial auto-correlation could have the effect of driving t statistics high (a common and surprising finding in much of the persistence literature). This critique has been developed in, *inter alia*, (Kelly 2019) and (Kelly 2020).

Dell does go some way to account for spatial auto-correlation. Indeed, the two baseline specifications give some confidence that the results are not the pure result of spatial auto-correlation. Dell also uses robust clustered standard errors at the district level, that help control for similarities be-

3.4 Table 2: Regression Results

Replicated Table 2 from (Dell 2010).

	Log. Equiv. Household Consumption (2001)			Stunted Growth, Children 6-9 (2005)			
	(1)[1]	(2)[1]	(3)[1]	(4)[1]	(5)[1]	(6)[1]	(7)[1]
			Latitude and Longitude				
Mita	-0.284	-0.216	-0.331	0.070	0.084+	0.087+	0.114*
	(0.199)	(0.207)	(0.219)	(0.043)	(0.046)	(0.048)	(0.049)
Num.Obs.	1478	1161	1013	158848	115761	100446	37421
R2	0.059	0.060	0.069	0.051	0.020	0.017	0.050
			Potosí				
Mita	-0.337***	-0.307**	-0.329**	0.080***	0.078***	0.078**	0.063+
	(0.087)	(0.101)	(0.096)	(0.021)	(0.022)	(0.024)	(0.032)
Num.Obs.	1478	1161	1013	158848	115761	100446	37421
R2	0.046	0.036	0.047	0.049	0.017	0.013	0.047
			Mita				
Mita	-0.277***	-0.230*	-0.224*	0.073**	0.061**	0.064**	0.055+
	(0.078)	(0.089)	(0.092)	(0.023)	(0.022)	(0.023)	(0.030)
Num.Obs.	1478	1161	1013	158848	115761	100446	37421
R2	0.044	0.042	0.040	0.040	0.015	0.013	0.043

[1]Robust standard errors, clustered by district, are in parentheses
+ p <0.1, * p <0.05, ** p <0.01, *** p <0.001
All regressions include controls for elevation and slope, as well as boundary segment fixed effects. Columns 1–3 include demographic controls for the number of children, and adults in the household.

tween neighbouring areas. As (Kelly 2019) however notes, this mitigation is only meaningful insofar as the residuals are uncorrelated between clusters, which is not typically the case for spatial data.

One way to test for spatial auto-correlation is to compute the Moran statistic. The Moran statistic (Moran 1950) is a measure used in spatial statistics to determine the degree of spatial auto-correlation. I compute Moran statistics for the three panel specifications, at the <50km bandwidth, and for the household consumption dependent variable, using the following method. The result are included below in Table 3.

(1) Computing the spatial weights matrix: The first step involves defining the spatial relationships between different data points. This is done by creating a 'neighbours list', which identifies which data points are considered neighbours of each other. I compute this by first producing a neighbours list using the latitude and longitude data. The neighbour list computes the k nearest other observations (k neighbours $= 5$, per (Kelly 2020)) in Euclidean space. This is computed using great circle distance between two points i and j such that:

$$d_{ij} = R * arcos[cos(\Delta Lon) \times sin(Lat_{r(j)}) + cos(Lat_{r(i)}) \times cos(Lat_{r(j)})] \quad (2)$$

where subscripts d and r refer to degrees and radians, $\Delta Lon = Lon_r(j) - Lon_r(i)$ and R is the radius of the earth. Each observation is connected to its k closest neighbours, forming a network of spatial relationships. These connections are then used to construct the spatial weights matrix, where the weights represent the strength or presence of a spatial relationship between each pair of observations. An $N \times N$ matrix, W, is formed where the element W_{ij} expresses connectivity between the ith and j^{th} area, defined as a binary if an element is a neighbour (i.e. whether it is on its neighbour list) and zero otherwise. The rows are then normalised. The diagonal elements of W are by definition all zero.

Computing Moran's I: The Moran's I statistic is then computed in the form:

$$I = \frac{n}{W} \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \quad (3)$$

where n is the number of locations, W is the sum of all weights, w_{ij} is the weight between locations i and j, x_i and x_j are the standardized values of the variable at locations i and j, and $\bar{x}$ is the mean of the variable. The formula is essentially a ratio, where the numerator captures the product of the deviation of each observation from the mean and the deviation of its neighbours (weighted by the spatial weights matrix), and the denominator is the total variance of the variable. A positive Moran's I indicates positive spatial auto-correlation (similar values cluster together), while a negative Moran's I indicates negative spatial auto-correlation (dissimilar values cluster together).

Applying Moran statistics: In the context of Dell's analysis, the presence of spatial auto-correlation would suggest that there are spatial

patterns in the data that are not random. This could mean that the effects of the *mita* system are spatially varying and possibly correlated with other regional characteristics or unobserved variables. The results at Table 3 suggest that low-to-moderate spatial auto-correlation is a potential issue with the analysis. I have visualised the result for the 50km bandwidth in Figure 3 contained in the Appendix (5). The figure shows the Moran statistic as the slope of the line that best fits the relationship between each observation's household consumption and spatially lagged counterparts.

Across all three models, the Moran statistics are positive and significant, ranging from 0.093 to 0.125.[3] The expected value is ≈ 0, which aligns with the expectation of random spatial distribution under the null hypothesis. Nearby observations tend to have more similar residual values than would be expected if there were no spatial auto-correlation. This would tentatively support the critique from (Kelly 2019) and (Kelly 2020) that persistence regressions do not contain adequate controls for unobserved spatially-correlated variables, suggesting the causal claim made regarding the institutional effect of the *mita* may not be as significant as first thought.

Spatial auto-correlation does not necessarily invalidate Dell's findings but highlights the need for careful interpretation. It suggests that additional spatial factors might be influencing the outcomes, and these should be accounted for in the analysis. The auto-correlation might occur for a host of reasons, for instance, if regions closer to ancient mines, irrespective of the *mita* influence, have developed distinct economic characteristics due to factors like mineral wealth. This could create a pattern where neighbouring regions have similar economic outcomes, not directly due to the *mita*, but because of their proximity to mines. This is however speculation and lends itself to further investigation.

4   Conclusion

(Dell 2010) provides one of the most important results in the historical persistence literature, demonstrating the long-term and significant effect of extractive economic institutions on contemporary living standards. These results are consistent across household consumption and childhood stunting and are robust to a range of specifications. Building on (Kelly 2019), I extend Dell's analysis to test whether the results are robust to spatial auto-correlation, finding evidence of low-to-medium spatial auto-correlation. This suggests that there may be confounding geographic factors that could at least partially explain Dell's results.

---

3. As weights and the variable of interest are standardised in this analysis, the Moran statistic is bounded $[-1, 1]$, with -1 denoting perfectly negative spatial auto-correlation and 1 perfectly positive spatial auto-correlation. A value of 0 would represent a random geographic distribution.

5 Appendix

Table 3: Moran's I Test Results for Spatial Auto-correlation

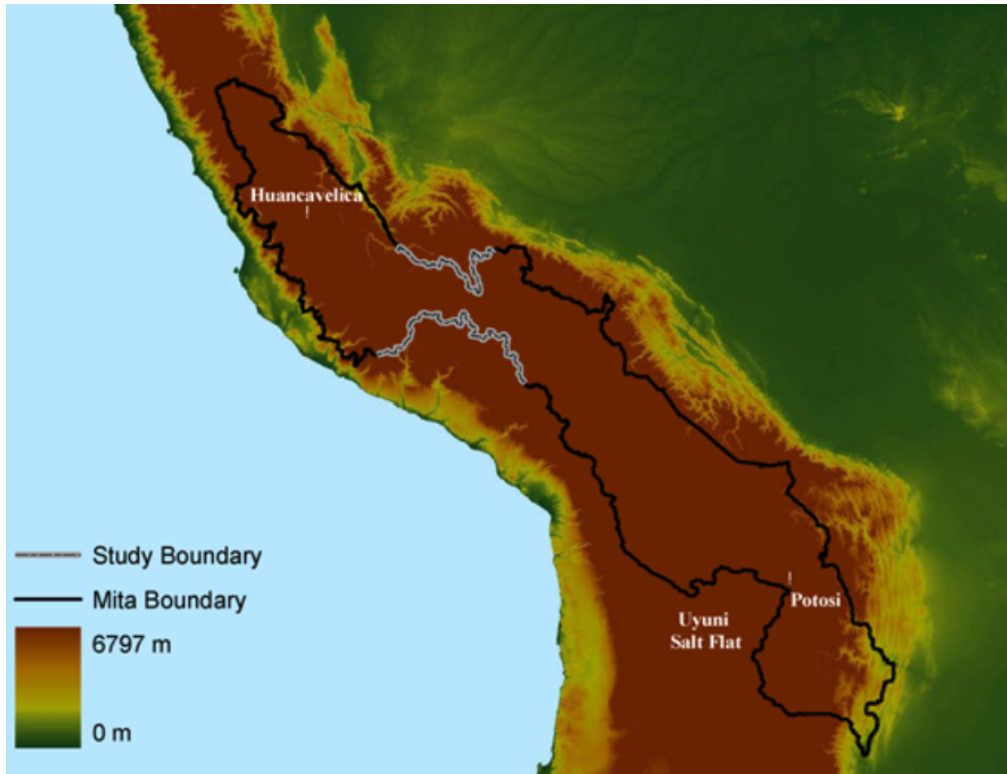| Model	Moran's I	P-value	Expected Value
Latitude-Longitude Model	0.071	0.000	-0.001
Potosi Model	0.097	0.000	-0.001
Mita Model	0.099	0.000	-0.001

Figure 2: The *mita* region of Peru and Bolivia

Figure 2: The *mita* boundary is in black. The study boundary is in light grey. The map also shows elevation across the area of interest. Source: (Dell 2010).

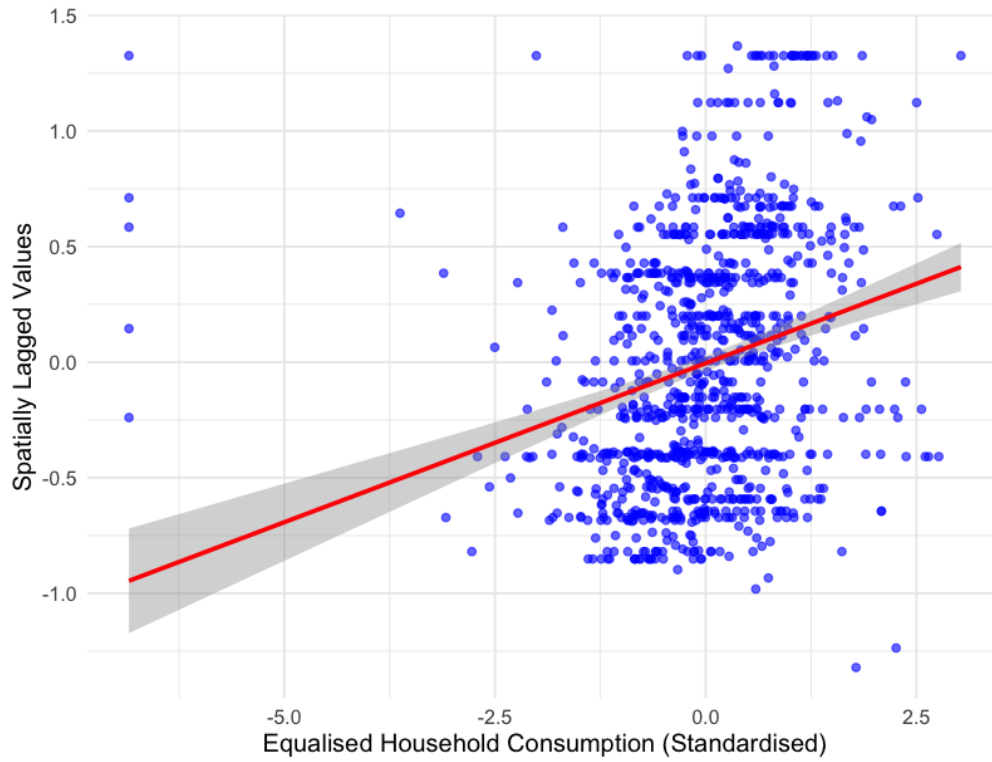Figure 3: Visualising Moran statistics

Figure 3: The plot shows spatially lagged values of standardised household consumption correlated with non-lagged values, for the <50km bandwidth. This visually shows the Moran statastic, which is essentially the slope of this graph.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The colonial origins of comparative development: An empirical investigation." *American Economic Review* 91 (5). ISSN: 00028282. https://doi.org/10.1257/aer.91.5.1369.

————. 2002. "Reversal of fortune: Geography and institutions in the making of the modern world income distribution." *Quarterly Journal of Economics* 117 (4). ISSN: 00335533. https://doi.org/10.1162/003355302320935025.

Amat Y Junient, M. 1947. "Sevilla: Escuela de Estudios Hispano- Americano." *Memoria de Gobierno.*

Angrist, Joshua D., and Jörn Steffen Pischke. 2010. "The credibility revolution in empirical economics: How better research design is taking the con out of econometrics." *Journal of Economic Perspectives* 24 (2). ISSN: 08953309. https://doi.org/10.1257/jep.24.2.3.

Banerjee, Abhijit, and Lakshmi Iyer. 2005. *History, institutions, and economic performance: The legacy of colonial land tenure systems in India,* 4. https://doi.org/10.1257/000282805482574.

Dell, Melissa. 2010. " The Persistent Effects of Peru's Mining Mita." *Econometrica* 78 (6). ISSN: 0012-9682. https://doi.org/10.3982/ecta8121.

Engerman, S, and Sokoloff K. 1997. "Factors Endowments, Institutions, and Differential Paths of Growth Among New World Economies." In *How Latin America Fell Behind,* edited by S Haber, 260–304. Stanford: Stanford University Press.

Glaeser, Edward L., Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2004. "Do institutions cause growth?" *Journal of Economic Growth* 9 (3). ISSN: 13814338. https://doi.org/10.1023/B:JOEG.0000038933.16398.ed.

Kelly, Morgan. 2019. "The Standard Errors of Persistence." *SSRN Electronic Journal,* https://doi.org/10.2139/ssrn.3398303.

————. 2020. "Understanding Persistence." *Cper.*

Moran, P. A. P. 1950. "Notes on Continuous Stochastic Phenomena." *Biometrika* 37 (1/2). ISSN: 00063444. https://doi.org/10.2307/2332142.

Morner, C. 1978. *Perfil de la sociedad rural del Cuzco a fines de la coloni.* Lima: Universidad del Pacifico.

Nunn, Nathan. 2008. "The Long Term Effects of Africa{\textquoteright}s Slave Trades." *Quarterly Journal of Economics* 123 (1).

Saignes, T. 1984. "Las etnias de Charcas frente al sistema colonial (Siglo XVII): Ausentismo y fugas en el debate sobre la mano de obra indígena, 1595-1665." *Jahrbuchfr Geschichte (JAS)* 21:27–75.

Tobler, W. R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46. ISSN: 00130095. https://doi.org/10.2307/143141.