Human Capital Persistence and Development

Rudi Rocha (UFRJ)
Claudio Ferraz (PUC-Rio)
Rodrigo R. Soares (EESP-FGV)

WIDER Development Conference
June 7, 2016

Motivation

- Human capital plays a key role in explaining modern cross-region and cross-country differences in income (Acemoglu and Dell 2010, Gennaioli et al. 2012).

- However, identifying human capital as a determinant of long-term development remains a challenge. Correlated mechanisms:

  - Incentives of the elites (i.e. inequality and institutions).

  - Characteristics of the population (i.e. ethnicity, culture, or religion).

  - Geography.

- Dense literature: North and Thomas 1973; Acemoglu et al. 2002, Galor et al. 2009; Engerman et al., 2012; Easterly and Levine 2012; Acemoglu et al. 2014.

This Paper...

- In a controlled empirical setting: we isolate the role of initial HK as a determinant of long-term development.

- More specifically, we explore variation in HK in late 19th/early 20th century induced by state sponsored settlements in SP.

  - Between 1877-1915 the government of SP created 30 official settlement colonies across the state.

  - Rural villages formed by small plots of land occupied by European immigrants with relatively more education.

  - Regions with/out settlement colonies with similar pre-conditions. Education was the only observable dimension to change.

- We follow regions over time, document HK persistence, and long-run development; we provide evidence on likely mechanisms.
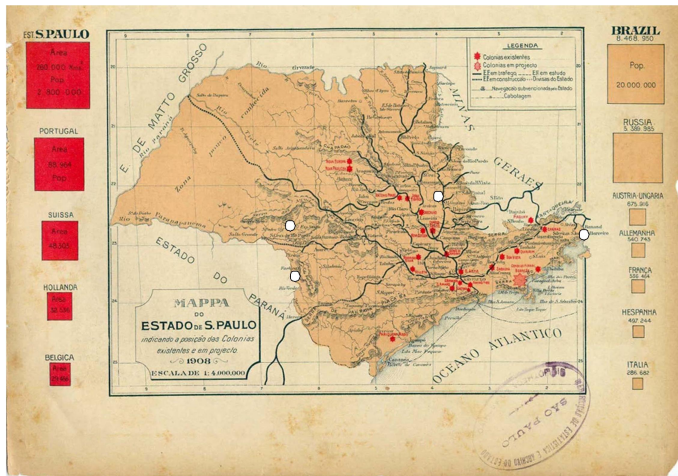
Historical Background

- The context:

    - Coffee expansion generated food shortage in urban centers.

    - Slavery abolishment (1888) generated labor supply shortage at the agricultural frontier.

    - Unclaimed public lands in backward areas and the in frontier.

- Among the political elite, there was a view that Europeans were more educated and productive than native workers.

- The main goal of the colonies was thus the attraction of European immigrants: a hook in the official propaganda.

- Because of different motivations, the same colonization framework was distributed across different regions of SP.

Historical Background

- The framework:

  - After legislative approval, public employees searched for locations mainly among unclaimed public lands.

  - Exact location was chosen based on two conditions: soil fertility and proximity to transportation.

  - Small plots of land were then designated and separated into different types (rural and urban). Plots of land were offered by Brazilian representatives in European cities.

  - The land was not freely-given, it had to be paid in installments after the first harvest. After the payment of installments, the settler could claim the property right.

  - The settlement was emancipated from public administration when all the plots were paid for.

Data
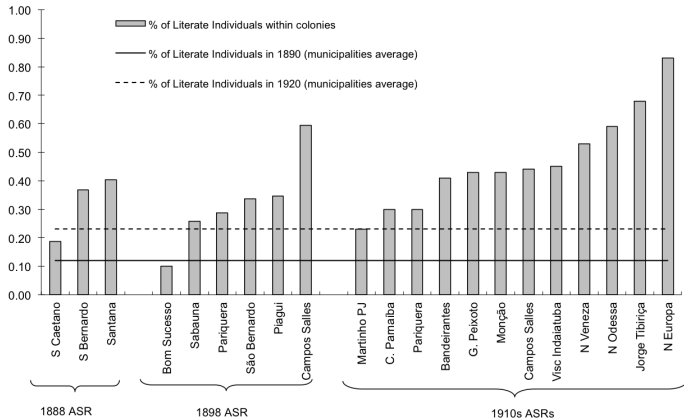
- Settlements' location: ASRs + case studies + ArqPubSP.

- Census (1872, 1920, 1940, 2000) + ASRs.

  - Pre-existing conditions, 88 municipalities.

  - Short-term: 1920, 202 municipalities (main division).

  - Mid-term: 1940, 270 municipalities adjusted to 202 AMCs.

  - Long-term: 2000, 645 municipalities adjusted to 202 AMCs.

- Controls:

  - Geography (lat, long, soil quality etc).

  - Coded railways across municipalities and over time.

Empirical Models

- Estimating short-term effects:

$$\mathrm{HK}_{i,1920} = \alpha + \beta_1 \mathrm{Settlement}_i + \mathrm{X}'_{i,1872}\beta_2 + \mathrm{Geo}'_i\beta_3 + \epsilon_i$$

- Estimating mid/long-term effects, analogously:

$$\mathrm{HK}_{i,1940/2000} = \alpha + \beta_1 \mathrm{Settlement}_i + \mathrm{X}'_{i,1872}\beta_2 + \mathrm{Geo}'_i\beta_3 + \epsilon_i$$

- HK long-term effects on income, via 2SLS:

$$Y_{i,2000} \text{ on } \mathrm{HK}_{i,1920},$$
$$\text{instrumented by } \mathrm{Settlement}_i$$

Table 3 - Summary Statistics for Geographic Characteristics and for Socioeconomic Variables in 1872 for Municipalities With and Without Settlements After 1872

Variables in 1872	Settlements		No Settlements		Difference
	Mean	Std. Dev.	Mean	Std. Dev.	
Log distance to capital (in km)	4,92	0,73	4,87	0,58	0,053
Latitude	-22,81	0,76	-22,94	0,90	0,126
Longitude	-47,14	1,49	-46,84	1,20	-0,299
Altitude (in 100m)	5,74	1,53	5,96	2,57	-0,217
Latosol (0/1)	0,53	0,38	0,44	0,44	0,091
% Literacy rate	0,19	0,07	0,17	0,10	0,019
% Children attending school	0,15	0,15	0,16	0,13	-0,015
% Foreigners	0,02	0,02	0,01	0,01	0,010*
% Slaves	0,20	0,10	0,17	0,10	0,025
Population density	9,72	7,49	9,47	6,88	0,249
Railway	0,11	0,32	0,01	0,12	0,091
Public administration	1,15	1,05	1,01	1,11	0,137
Legal professionals	0,95	0,57	0,86	0,94	0,087
% Emp agriculture	0,61	0,10	0,59	0,12	0,014
% Emp manufacturing	0,11	0,04	0,10	0,05	0,012
% Emp services and retail	0,28	0,09	0,31	0,11	-0,027

Short-Term Effects on Human Capital

Table 5 - The Short-Term Effects of Settlements on Literacy Rates

Dependent Variable: Literacy rate (%) in 1920	No controls	Control for Geography	Control for Charact. 1872	No Controls	Control for Geography	Control for Charact. 1872
	(1)	(2)	(3)	(4)	(5)	(6)
Settlement	0.104 (0.030)***	0.102 (0.028)***	0.081 (0.027)***			
% Population in Settlement				0.748 (0.310)**	0.676 (0.320)**	0.484 (0.260)*
% Foreigners 1872			-0.06 (0.604)			0.021 (0.535)
% Slaves 1872			0.319 (0.090)***			0.333 (0.090)***
% Literate 1872			0.063 (0.060)			0.077 (0.061)
% Agriculture 1872			1.579 (2.417)			0.962 (2.398)
% Manufacturing 1872			1.865 (2.417)			1.239 (2.402)
% Services and Retail 1872			1.611 (2.401)			0.989 (2.385)
Railway 1872			0.115 (0.024)***			0.091 (0.026)***
Geography controls	No	Yes	Yes	No	Yes	Yes
Observations	202	202	202	199	199	199
Adjusted R-squared	0.10	0.31	0.38	0.04	0.29	0.38

Table 6 - Mid and Long-Term Effects of Settlements on Human Capital

Dependent variable:	1940 Census		2000 Census		
	Literacy rate (aged 5+)	Literacy rate (aged 15-19)	Literacy rate (aged 5+)	Literacy rate (aged 15-19)	Years of Schooling (aged 5+)
	(1)	(2)	(3)	(4)	(5)
Settlement	0,107	0,135	0,015	0,001	0,519
	(0.034)***	(0.040)***	(0.005)***	(0.002)	(0.138)***
Geographic controls	Yes	Yes	Yes	Yes	Yes
Characteristics 1872	Yes	Yes	Yes	Yes	Yes
Observations	202	202	202	202	202
Adjusted R-squared	0,45	0,45	0,41	0,16	0,42

Table 7 - Persistent Effects of Settlements on Years of Schooling, by Cohorts of Individuals Born in the Municpality

Dependent variable: Years of schooling for individuals born in municipality	All Cohorts	1920-29	1930-39	1940-49	1950-59	1960-69
	(1)	(2)	(3)	(4)	(5)	(6)
Settlement	0,883	0,655	0,759	0,941	0,927	1,022
	(0.229)***	(0.236)***	(0.225)***	(0.277)***	(0.303)***	(0.200)***
Mean Dependent Variable	6,63	2,73	3,54	4,64	6,10	7,04
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Characteristics 1872	Yes	Yes	Yes	Yes	Yes	Yes
Observations	202	202	202	202	202	202
Adjusted R-squared	0,50	0,39	0,46	0,45	0,43	0,50

Table 9 - The Effects of Settlements on School Inputs and School Attendance

Dependent variable:	Schools per school aged child	Teachers per school aged child	% Children attending school
	(1)	(2)	(3)
Panel A: 1920 Census			
Settlement	0.222	5,053	NA
	(0.108)**	(1,874)***	
R-squared	0.163	0,423	
Panel B: 1940 Census			
Settlement	-0,251	NA	0.08
	(0.294)		(0.025)***
R-squared	0.202		0.381
Panel C: 2000 Census			
Settlement	-0.622	11,235	0.006
	(0.721)	(4,601)**	(0.003)*
R-squared	0.575	0,121	0.405
Geographic controls	Yes	Yes	Yes
Characteristics 1872	Yes	Yes	Yes
Observations	202	202	202

Alternative Stories

Table 10 - The Short-Term Effects of Settlements on Other Demographic and Economic Characteristics

Dependent variable in 1920	% Foreigners	% Literate foreigners	Population density	% Small farms	Coffee productivity	Log land prices	Log wages construction	Log wages agriculture
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Settlement	0.025	0.051	14.226	0.046	-0.008	0.185	0.04	0.031
	(0.011)**	(0.027)*	(15.060)	(0.029)	(0.007)	(0.144)	(0.062)	(0.056)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Characteristics 1872	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	202	172	202	202	202	202	148	125
R-squared	0.63	0.33	0.30	0.25	0.27	0.51	0.44	0.61

Alternative Stories

Table 11 - The Effects of Settlements on Literacy Rates, Conditional on National Identities and Religion

Dependent Variable:	Literacy Rate in 1920			Literacy Rate in 1940	Years of Schooling in 2000
	Control for country of origin	Control for religion	Control for country of origin and religion	Control for country of origin and religion	Control for country of origin and religion
	(1)	(2)	(3)	(4)	(5)
Settlement	0,087	0,085	0,091	0,121	0,521
	(0,027)***	(0,025)***	(0,026)***	(0,032)***	(0,149)***
% Germans	0,224		0,206	0,182	0,061
	(0,081)***		(0,080)**	(0,100)*	(0,528)
% Spaniards	0,067		0,066	0,113	0,602
	(0,070)		(0,066)	(0,066)*	(0,356)*
% Italians	0,128		0,129	0,211	0,54
	(0,055)**		(0,055)**	(0,062)***	(0,332)
% Portuguese	0,102		0,104	0,242	1,081
	(0,059)*		(0,059)*	(0,086)***	(0,378)***
% Japanese	0,019		0,02	0,034	0,462
	(0,074)		(0,070)	(0,077)	(0,411)
% Catholics		0,083	0,081	0,064	0,128
		(0,023)***	(0,024)***	(0,039)	(0,307)
Geographic controls	Yes	Yes	Yes	Yes	Yes
Characteristics 1872	Yes	Yes	Yes	Yes	Yes
Observations	202	202	202	202	202
Adj R-squared	0,46	0,45	0,48	0,44	0,37

Table 12 - Long-Run Effects of Settlements on Income per Capita

Dependent variable: Log income per capita in 2000	OLS (1)	OLS (2)	IV (3)	IV (4)
Settlement	0,153 (0.054)***			
Literacy rate (%) in 1920		1,122 (0.131)***	1,884 (0.434)***	0,274 (1.252)
Years of schooling in 2000				0,148 (0.074)**
F-test excluded instrument			9,34	2,31
Geographic controls	Yes	Yes	Yes	Yes
Characteristics 1872	Yes	Yes	Yes	Yes
Observations	202	202	202	202
Adjusted R-squared	0,44	0,53		

Notes: Robust standard errors in parentheses, clustered at the 1872 census boundaries: *** p<0.01, ** p<0.05, * p<0.1. Dependent

Table 13 - The Long-Run Effects of Settlements on Employment and Structural Transformation

Dependent variable:	% Emp. Agriculture	% Emp. Manufacturing	% Emp. Retail and Services
	(1)	(2)	(3)
Panel A: 1920 Census			
Settlement	-0.121	0.072	0.048
	(0.039)***	(0.024)***	(0.017)***
Adjusted R-squared	0.457	0.390	0.503
Panel B: 1940 Census			
Settlement	-0.165	0.092	0.073
	(0.046)***	(0.027)***	(0.028)**
Adjusted R-squared	0.496	0.472	0.396
Panel C: 2000 Census			
Settlement	-0.053	0.006	0.048
	(0.019)***	(0.016)	(0.017)***
Adjusted R-squared	0.391	0.315	0.280
Observations	202	202	202
Geographic controls	Yes	Yes	Yes
Characteristics 1872	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at the 1872 census boundaries: *** p<0.01, ** p<0.05, * p<0.1.

Table 14 - The Effects of Settlements on Years of Schooling, by Cohorts of Individuals Not Born in the Municipality

Dependent variable: Years of Schooling of immigrants	All Cohorts	1920-29	1930-39	1940-49	1950-59	1960-69
	(1)	(2)	(3)	(4)	(5)	(6)
Settlement	0,544	0,333	0,285	0,534	0,542	0,766
	(0.147)***	(0.174)*	(0,185)	(0.191)***	(0.198)***	(0.177)***
Mean Dependent Variable	5,79	3,02	3,61	4,56	5,85	6,51
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Characteristics 1872	Yes	Yes	Yes	Yes	Yes	Yes
Observations	202	202	202	202	202	202
R-squared	0,33	0.28	0.27	0.27	0,21	0.20

Notes: Robust standard errors in parentheses, clustered at the 1872 census boundaries; *** p<0.01, ** p<0.05, * p<0.1. In all columns

Concluding Remarks

- We find that an increase in human capital due to state-sponsored settlements changed the level of education in treated regions; this increased level of HK persisted over time.

- Immigrants demanded schools and inputs in settlement areas, but still unclear why inputs persist. Intergenerational transmission of the value of education? Political economy mechanism?

- With high levels of internal migration, agglomeration and concentration of skills in more skill-intensive regions may have played an important role in the long-run.