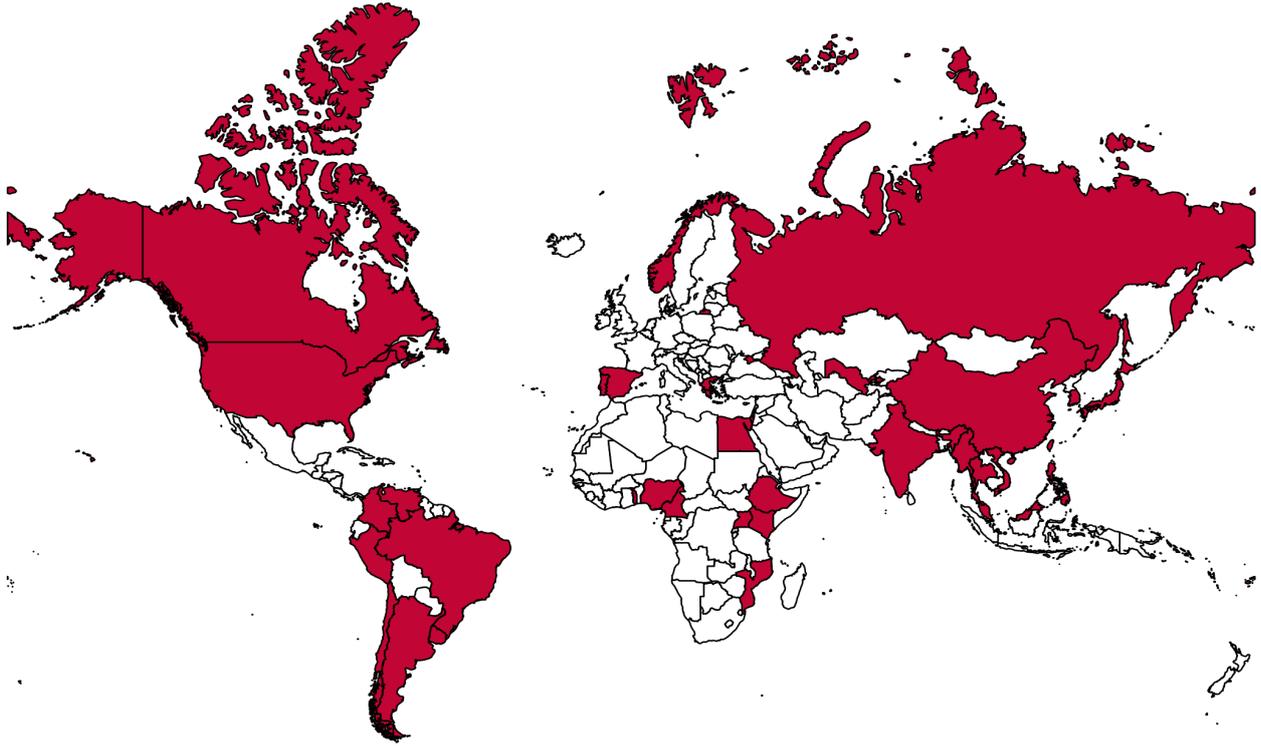


Online Appendix

Social Exclusion and Social Preferences: Evidence From Colombia's Leper Colony

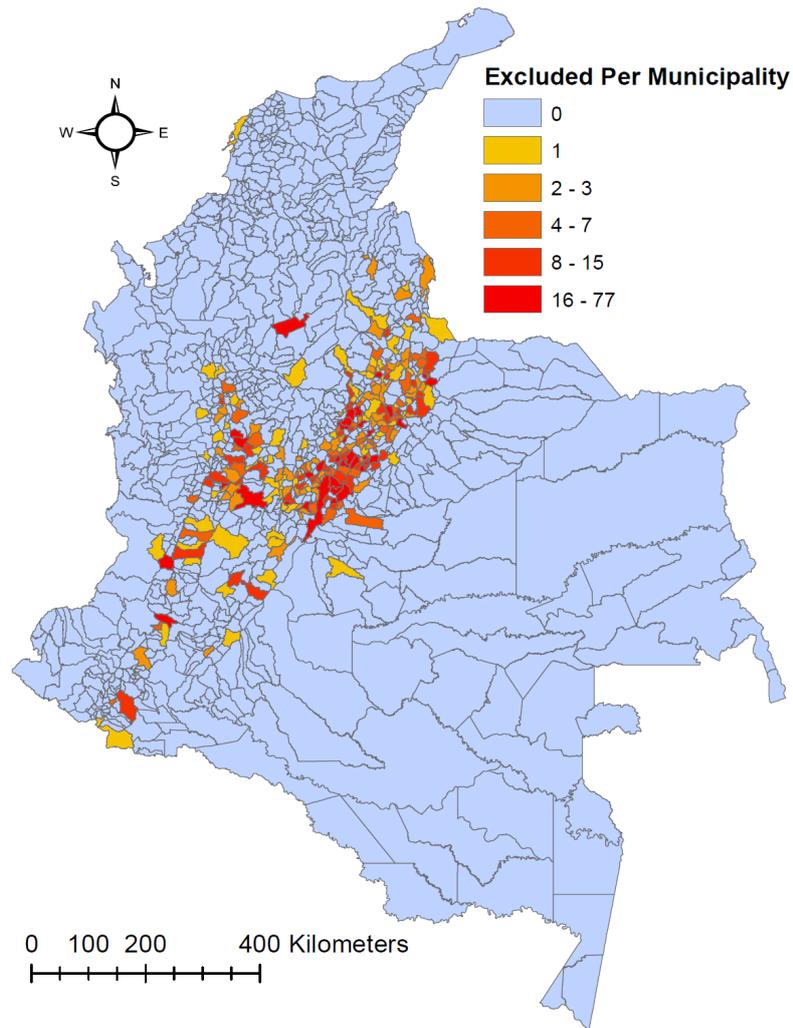
Diego Ramos-Toro
Dartmouth College

Appendix A.1 - Countries with Leper Colonies Similar to Agua de Dios



The map shows the geographical distribution of countries that had leper colonies with features similar to those of the leper colony of Agua de Dios. The countries in red are those where a leper colony of such characteristics has been described by at least one source. Sources: For countries outside Latin America, the information comes from the International Leprosy Association (retrieved from https://leprosyhistory.org/geographical_region/country on Apr 26 2022), which contains a non-exhaustive list of countries around the globe with leper colonies similar to Agua de Dios, plus the leper colonies discussed in Zamparoni (2017), Kula and Robinson (2013), and Sprawson (1939). For Latin America, the countries were identified following various sources (Senado y Cámara de la Nación Argentina, 1928; Imprenta Nacional República Oriental de Uruguay, 1938; Romero Salazar, 2001; Mizell-Nelson, 2003; Cueto, 2004; Modelli, 2016).

Appendix A.2 - Municipalities of Origin of Patients in the Leper Colony in 1920



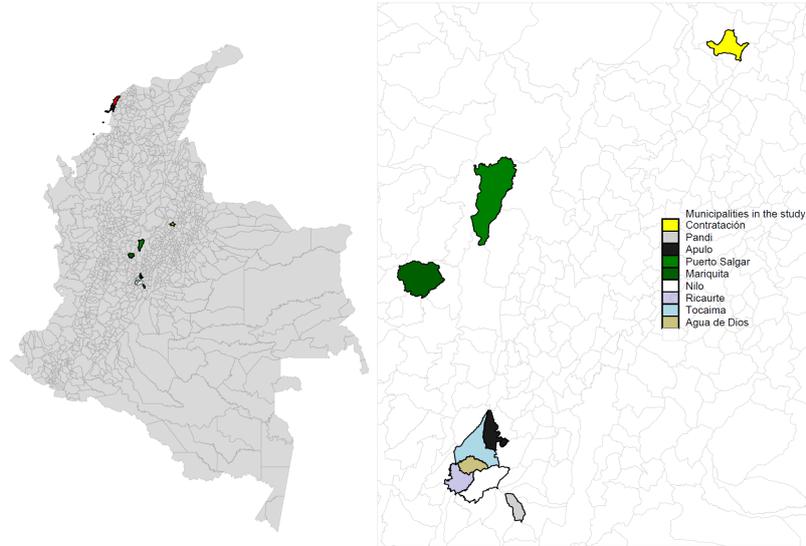
Note: The map shows the municipalities in Colombia where leprosy sufferers were identified and forcefully removed before being sent to the leper colony of Agua de Dios. Darker shades of red correspond to municipalities where a larger number of people with leprosy were identified and forcefully sent to the colony. Data comes from information recorded by the Chief Doctor in the colony in 1920.

Appendix A.3 - Historical Correlates of Municipality-Level Incidence of Leprosy in 1920

	(1)	(2)	(3)	(4)	(5)
	Population Density 1918	Population Density 1928	Male Share 1918	Male Share 1928	Conflict 1900-1917
<i>Panel A - No Controls</i>					
Leprosy Incidence	0.187 (0.148)	0.167 (0.199)	-0.000 (0.000)	0.000 (0.000)	-0.003 (0.001)
R-squared	0.004	0.002	0.025	0.000	0.014
<i>Panel B - Controlling for Climatic Characteristics</i>					
Leprosy Incidence	0.259 (0.163)	0.205 (0.207)	-0.000 (0.000)	0.001 (0.001)	-0.002 (0.001)
R-squared	0.027	0.022	0.145	0.090	0.116
<i>Panel C - Controlling for Climatic and Historic Characteristics</i>					
Leprosy Incidence	0.259 (0.164)	0.205 (0.207)	-0.000 (0.000)	0.001 (0.001)	-0.002 (0.001)
R-squared	0.027	0.023	0.146	0.093	0.120
Mean Dep. Var	49.906	64.672	0.485	0.393	0.094
Observations	254	254	254	254	254

Note: The table suggests that the historical incidence of leprosy at the municipality level was not associated with historical economic development, gender composition, or conflict. The unit of observation is a municipality, and data comes from all municipalities in which at least one case of leprosy was identified and sent to the colony of Agua de Dios (i.e., municipalities for which there is higher certainty that leprosy cases were accurately recorded and monitored; for other municipalities, it is unknown if no patients were sent to the colony because of no leprosy cases or due to lack of presence by the state and by medical entities). Each panel-column reports results from a single OLS regression of the dependent variable described in the column heading on the variable(s) described in the row heading(s). Leprosy Incidence is the number of cases in 1918 in each municipality per 10,000 people, constructed based on the recorded number of people from each municipality sent to the colony (as recorded by the Chief Doctor in the colony in 1920) and municipalities' populations as they appear in the Population Census of 1918 (Uribe-Castro, 2018). Columns 1 and 2 have as outcome variables the population density (people per squared kilometer) in 1918 and 1928, respectively. Columns 3 and 4 have as outcome variables the percentage of the population that was male in 1918 and 1928, respectively. Column 5 is an indicator variable that equals 1 if there was a recorded land conflict in the municipality between 1900 and 1917. Panel A shows the estimates of Leprosy Incidence based on bivariate regressions (i.e., without controls). Panel B shows the estimates of Leprosy Incidence, controlling for precipitation and its seasonality, temperature and its seasonality, and ruggedness. Panel C shows the estimates of Leprosy Incidence, controlling for the same climatic characteristics as Panel B and for an indicator variable that equals 1 if there was an indigenous population recorded in the municipality in the early sixteenth century (a proxy for pre-colonial economic development). Heteroskedastic-robust standard errors are in parenthesis.

Appendix A.4 - Other Municipalities Relevant To This Study - Location



Note: The map shows the geographical location of municipalities outside the region where the study was conducted, which are nonetheless important in this paper for various reasons. These municipalities can be separated into four categories: first, municipalities that were selected as out-groups in the Dictator Games (Apulo, in black, selected as the outgroup in the first round of fieldwork; and Pandi, in grey, selected as the out-group in the second round of fieldwork). Second, municipalities that are similar to Agua de Dios in terms of climatic attributes (Mariquita and Puerto Salgar, in dark and light green, respectively, which served as a comparison group in Appendix Sections A.13 and A.19 - see Appendix Section A.8 for more details on the data that supports why these municipalities serve as a 'climatically-similar' comparison group). Third, Nilo (in white), which is the third municipality that adjoins the former leper colony, where no data was collected due to prevailing systematic historical differences. Fourth, the municipalities of Contratación (yellow) and Caño Loro (red), where the other two leper colonies in Colombia were located. As discussed in Section 2, Agua de Dios was the largest leper colony, and by the 1940s it was the only functioning colony that was receiving patients from throughout Colombia.

Appendix A.5 - Characteristics of Individuals in Adjoining Municipality That Was Not Included, as Well as Those From Outgroups in Dictator Games

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Sample Averages			Difference between region of study and excluded adjoining municipality			Difference between region of study and outgroup municipality in 1st fieldwork			Difference between region of study and outgroup municipality in 2nd fieldwork			
Region of Study Mean	Nilo (Excluded Mun) Mean	Apulo (Outgrp 1) Mean	Pandi (Outgrp 2) Mean	Nilo-Region OLS	Nilo-Region SE	Nilo+Region Obs	Apulo-Region OLS	Apulo-Region SE	Apulo+Region Obs	Pandi-Region OLS	Pandi-Region SE	Pandi+Region Obs	
Panel A - Sociodemographic Characteristics - Census 1973													
Female	0.511	0.486	0.499	0.497	-0.029	0.016	57,016	-0.013	0.018	61,402	-0.014	0.014	59,898
Age	27.468	24.135	23.093	27.802	-2.727	2.116	57,016	-4.648	1.471	61,402	3.352	3.155	59,898
Marital Status	0.272	0.229	0.229	0.350	-0.039	0.039	57,016	-0.047	0.032	61,402	0.079	0.057	59,898
Primary	0.753	0.682	0.736	0.752	-0.080	0.035	51,135	-0.021	0.045	55,144	-0.001	0.051	53,987
Secondary	0.179	0.098	0.166	0.160	-0.085	0.032	51,135	-0.016	0.041	55,144	-0.019	0.038	53,987
Labor Force Participation	0.351	0.337	0.357	0.394	-0.008	0.044	56,086	-0.002	0.043	60,430	0.044	0.071	58,952
Children	3.918	4.134	3.877	4.571	0.229	0.280	15,684	0.018	0.581	16,579	0.646	0.482	16,597
Deceased Children	0.327	0.353	0.316	0.331	0.024	0.026	12,291	-0.012	0.038	12,983	0.004	0.031	13,189
Literacy	0.766	0.696	0.738	0.766	-0.076	0.039	51,136	-0.033	0.048	55,145	-0.001	0.058	53,988
Income	1,293.364	904.907	1,047.961	1,251.146	-403.687	158.098	16,534	-219.071	179.848	17,627	-44.280	276.246	17,870
Panel B - Housing Characteristics - Census 1973													
Has Sewerage	0.452	0.189	0.326	0.218	-0.263	0.116	8,539	-0.126	0.184	9,086	-0.234	0.156	8,444
Has Potable Water	0.435	0.378	0.410	0.404	-0.058	0.075	8,539	-0.025	0.167	9,086	-0.031	0.110	8,444
Has a Kitchen	0.774	0.712	0.856	0.800	-0.062	0.024	8,539	0.082	0.028	9,086	0.026	0.026	8,444
Has a Toilet	0.302	0.155	0.320	0.174	-0.147	0.067	8,539	0.018	0.129	9,086	-0.128	0.106	8,444
Has Electricity	0.599	0.244	0.403	0.236	-0.355	0.129	8,539	-0.196	0.199	9,086	-0.363	0.135	8,444
Has No Ground Slab	0.407	0.167	0.494	0.256	-0.239	0.060	8,539	0.087	0.114	9,086	-0.150	0.075	8,444
Walls Made out of Bricks	0.472	0.362	0.417	0.273	-0.110	0.025	8,539	-0.054	0.046	9,086	-0.199	0.026	8,444
Roof with Concrete Slabs	0.063	0.006	0.071	0.022	-0.056	0.016	8,539	0.008	0.035	9,086	-0.041	0.020	8,444
Own House	0.424	0.323	0.431	0.443	-0.102	0.021	8,539	0.006	0.016	9,086	0.019	0.025	8,444
Credit-Funded	0.036	0.080	0.076	0.046	0.043	0.014	8,539	0.039	0.017	9,086	0.010	0.013	8,444

Note: The table uses individual-level data (Panel A) and housing-level data (Panel B) from the 1973 Census to compare the socioeconomic characteristics of the region where the study was conducted with those of different municipalities. Overall, the table suggests that (1) the adjoining municipality that was not included in the study (Nilo) was systematically poorer than the study's region (Agua de Dios, Ricaurte, and Tocaima), and (2) the two municipalities that served as outgroups in the dictator games (Apulo and Pandi) are similar to the region where the study was conducted across several socioeconomic characteristics. Column 1 presents the average characteristics of people who report being from the region of study. Column 2 presents such characteristics for those who report being from the adjoining municipality that was not included in the study (Nilo). Moreover, Columns 3 and 4 show these averages for those who report being from the municipalities that served as outgroups in the first (Apulo) and the second (Pandi) rounds of fieldwork, respectively. For each of the socioeconomic outcomes in the row headings, Columns 5-7 show the OLS estimates of the difference between the adjoining municipality not included in the study and the region of study, the standard error of the difference (clustered at the municipality-of-residence level), and the number of observations. Columns 8-10 show these results for the difference between the first outgroup and the region of study, while Columns 11-13 show them for the difference between the second outgroup and the region of study. The observations vary across variables in the row headings due to the different number of responses that were recorded for each variable in the 1973 census.

Table A.6: Demographics of the Leper Colony and Its Surrounding Region Circa 1920

	(1)	(2)	(3)
	Colony	Contiguous	Predicted by Origin
Male share	0.471	0.488	0.475 (0.002)
Population Density	61.496	62.214	62.006 (2.521)

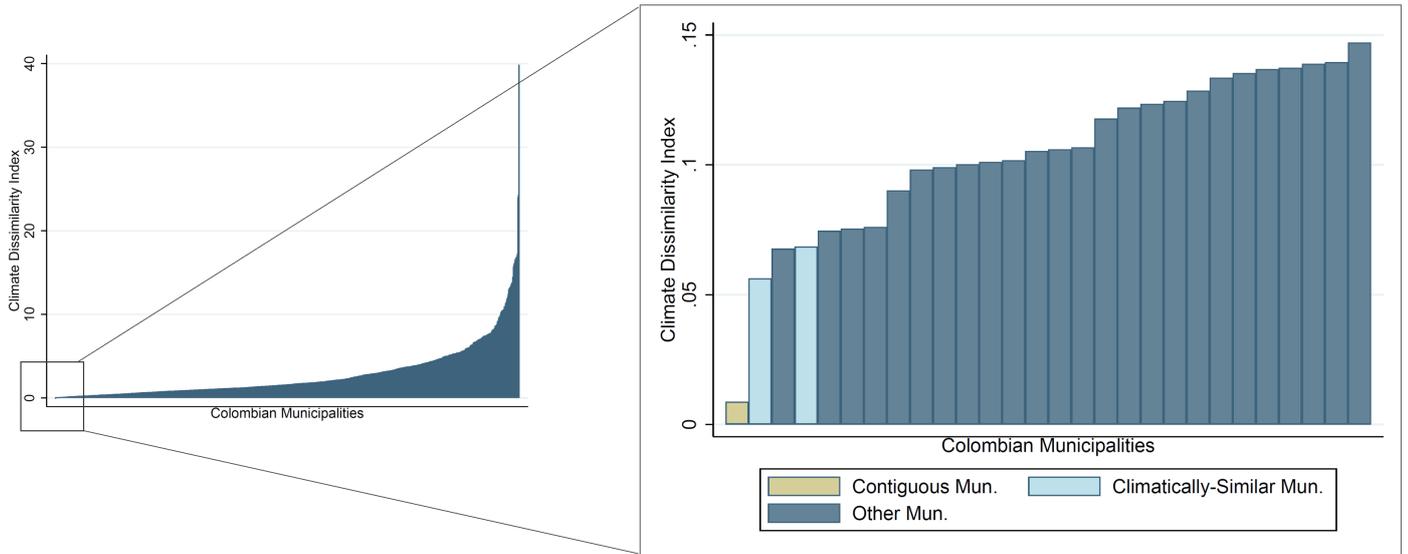
Note: The table examines census data from 1918 to document the similarity of the characteristics of inhabitants of the leper colony, those in the surrounding municipalities, and those predicted by the municipalities of origin of those who were patients of the colony circa 1920. Census data comes from Uribe-Castro (2018), which has municipality-level data on male population and total population for all municipalities in Colombia in 1918. Column 1 presents the male share and population density in the leper colony, which were computed with such census data and the area (in squared kilometers) of Agua de Dios. Column 2 presents the male share and population density computed for the contiguous municipalities. The variables in Column 2 were constructed with data from Tocaima, Ricaurte, and Girardot (i.e., not only Tocaima and Ricaurte), given that Ricaurte was a part of the municipality of Girardot in 1918 (Ricaurte was not formally a municipality until 1968). Column 3 shows the male share and population density that would be predicted by (a) these variables in the municipalities where patients came from and (b) the number of patients in Agua de Dios from each of these municipalities. Patients' municipalities of origin come from a full count of patients in the leper colony, as recorded by the Chief Doctor circa 1920 (mapped in Appendix A.3). I combine this data with male shares and population densities in such municipalities of origin recorded in the census, which allows computing weighted averages for these variables, using the number of patients that came from each municipality as weights. Column 3 also presents the standard errors of the weighted averages for each variable, in parentheses. Overall, the table suggests that the leper colony and its contiguous municipalities were comparable across these two outcomes, and these outcomes closely follow what would be predicted by the municipalities of origin and the number of patients from each of these places.

Appendix A7: People Born in Leper Colony Compared to Non-Excluded People Born in Contiguous Region

	(1)	(2)	(3)	(4)	(5)
	Colony = 1 Mean	Colony = 0 Mean	OLS Coefficient on <i>Colony</i>	SE	Obs.
Female	0.517	0.522	-0.005	(0.031)	10,279
Age	16.263	16.987	-0.724	(0.405)	10,279
Marital Status	0.060	0.083	-0.023	(0.021)	10,279
Primary	0.938	0.908	0.031	(0.017)	10,279
Secondary	0.327	0.252	0.075	(0.064)	10,279
Labor Force Participation	0.377	0.454	-0.077	(0.044)	10,274
Children	0.537	0.699	-0.162	(0.090)	2,380
Deceased Children	0.115	0.096	0.019	(0.032)	895
Literacy	0.947	0.917	0.030	(0.019)	10,279
Income (in COP)	644.237	719.601	-75.363	(113.011)	2,893

Note: The table examines further the validity of the findings presented in Table 1, showing the historical balance between those who almost certainly experienced the exclusion and those from the region of study who were almost certainly not excluded. Specifically, it examines differences in socio-demographic characteristics between those born in the leper colony before 1962 and those born before such a year in the municipalities surrounding the colony. Importantly, Agua de Dios was formally a leper colony until 1961, so those born in the colony until then would have to be direct descendants of the excluded, whereas those suffering from leprosy would not be able to exit the colony, let alone have children outside it. The data is thus restricted to Census respondents who were born between 1950 and 1961. For such a sub-sample, Column 3 shows the coefficients from regressions of each variable described in the row heading on an indicator that equals 1 for individuals born in the colony, while Column 4 presents the corresponding standard errors clustered at the municipality of residence in 1973. Column 5 presents the observations of each regression. Observations vary across rows due to different numbers of responses collected for different variables in the Population Census of 1973.

Appendix A.8 - Climatic Similarity Between the Former Leper Colony and Surrounding Municipalities



Note: The figure illustrates the climate dissimilarity indices for municipalities in Colombia, which summarize the climatic differences between each municipality in Colombia and Agua de Dios. The Climate Dissimilarity Index $CDI(c)$ for municipality c is defined as the sum of the squared percentage differences across different geo-climatic attributes of the municipality and those of Agua de Dios. Hence, $CDI(c) = \sum_{i=1}^N \left(\frac{x_i^c - x_i^{AD}}{x_i^{AD}} \right)^2$, where N is the number of climatic attributes, $c \in C$ are municipalities in Colombia, x_i^c is the average value of each attribute i in municipality c , and x_i^{AD} is the average value of each attribute i in Agua de Dios (so that a higher CDI entails a larger overall climatic difference between Agua de Dios and municipality c). The geo-climatic attributes included in the calculation of the CDI are temperature, precipitation, ruggedness, precipitation seasonality, and temperature seasonality. Bioclimatic variables were taken from WorldClim.org, while terrain ruggedness was computed with data from Nunn and Puga (2012). By construction, smaller indices correspond to municipalities that are more similar to Agua de Dios, CDI is positive, and a CDI equal to 0 can only happen if the municipality has the same geographical attributes as Agua de Dios. The figure graphs the CDI for all municipalities in Colombia (1121 municipalities). Out of all municipalities in such a country, the contiguous municipality of Tocaima has the smallest climatic differences with respect to Agua de Dios. The municipalities that serve as a comparison group in a "climatic matched" analysis (Appendix Sections A.13 and A.19) are in light blue: San Sebastián de Mariquita (second from left to right) is the municipality outside the region of study that is climatically closest to Agua de Dios, while Puerto Salgar (fourth from left to right) is the third most similar out of all municipalities outside the region of study and the most similar out of all the municipalities that belong to the sub-national unit of the study's region (Cundinamarca).

Appendix A.9 - Definition of Variables

- Age: Number of years of the participant since birth. Collected in all rounds of fieldwork.
- Age (Census): Number of years that the census respondent claims to have at the time of the Census. Taken from the 1973 and 2005 Colombian censuses.
- Alloc In-Group: Value that the participant chose to allocate in a Dictator Game to an anonymous receiver from their same municipality, in thousands of Colombian Pesos. Collected in rounds 1 and 2 of fieldwork. In the first round of fieldwork, the allocation was based on COP 16,000 given to participants in cash. Participants in the first round then decided how much of that they wanted to allocate to an anonymous recipient from their municipality. In the second round of fieldwork, participants were automatically enrolled in a lottery for COP 800,000 and were asked how much of that prize they would give to an anonymous recipient in the same municipality. Participants in the second round were warned that the reported ingroup (as well as the outgroup) allocation would be deducted from their prize if they were selected as the winner.
- Alloc Out-Group: Value that the participant chose to allocate in a Dictator Game to an anonymous receiver from a distant municipality, in thousands of Colombian Pesos. Collected in rounds 1 and 2 of fieldwork. In the first round of fieldwork, the allocation was based on COP 16,000 given to participants in cash. Participants in the first round then decided how much of that they wanted to allocate to an anonymous recipient from Apulo, a municipality outside the study's region. In the second round of fieldwork, participants were automatically enrolled in a lottery for COP 800,000 and were asked how much of that prize they would give to an anonymous recipient in the municipality of Pandi. Participants in the second round were warned that the reported outgroup (as well as the ingroup) allocation would be deducted from their prize if they were selected as the winner.
- Alloc Total: Sum of the values that the participant chose to allocate to ingroup and to outgroup members, in thousands of Colombian Pesos. Collected in rounds 1 and 2 of fieldwork. In the first round of fieldwork, the total allocation was based on COP 32,000 given in cash to participants in both dictator games. In the second round of fieldwork, participants were automatically enrolled in a lottery for COP 800,000 and were asked how much of that prize they would give to anonymous ingroup and outgroup recipients (i.e., total allocation was based on a total of COP 800,000).
- Alloc Difference: Difference between the ingroup allocation and the outgroup allocation selected by each participant, in thousands of Colombian Pesos. Collected in rounds 1 and 2 of fieldwork.

- Altruism (1-6 Scale): Response to the question “How much like you is someone who claims that it is important to help people living nearby; to care for their needs?” Data taken from Colombian responses in round six of the World Values Survey.
- Altruism (Binary): Indicator variable that equals 1 for respondents who said ‘very much like me’ or ‘like me’ when asked: “How much like you is someone who claims that it is important to help people living nearby; to care for their needs?” Variable constructed with data from Colombian responses in round six of the World Values Survey.
- Altruism Large Business Owners: Response to the question “On a scale from 1 to 10, how much would you agree with using government funds to support large business owners who have struggled because of the pandemic?” Collected in round 3 of fieldwork.
- Altruism Venezuelan Refugees: Response to the question “On a scale from 1 to 10, how much of a priority do you think the Colombian Government should give to policies making sure that the needs of Venezuelan refugees are met?” Collected in round 1 of fieldwork.
- Ancestry: Indicator variable that equals one if the participant has an ancestor (parent or grandparent) who was secluded in the leper colony during the epoch of isolation. The information that serves as an input for this variable was recorded in slightly different ways in each round of fieldwork: In the first round, participants were asked if any of their grandparents lived in the former leper colony. Ancestry thus equals 1 in the first round if the respondent reports at least one grandparent who lived in the former leper colony. In the second round, participants were asked if any of their parents or grandparents lived in the former leper colony. Ancestry thus equals 1 in the second round if the respondent reports having at least one parent or grandparent who lived in the former leper colony. Lastly, participants in the third round of fieldwork were asked if they had a relative that lived in Agua de Dios during the epoch of isolation. Those who stated that such was the case were then asked which specific relative had to live in isolation. Ancestry thus equals 1 in the third round of fieldwork if the respondent said that they had relatives who lived in isolation and then reported that the relatives who endured that were parents, grandparents, or great-grandparents.
- Antiparasitic Take-up: Indicator variable that equals one if the participant used a voucher to redeem a free dose of antiparasitic offered to them. Collected in round 2 of fieldwork.
- Be Rich and Have Money: Indicator variable that equals one if the respondent claims that it is very important to be rich and have money. Variable constructed with data from second-generation

migrants (i.e., respondents born in a country different than their parents' countries) in the European Social Survey.

- Care for Others' Well-being: Indicator variable that equals one if the respondent claims that it is very important to care for others' well-being. Variable constructed with data from second-generation migrants (i.e., respondents born in a country different than their parents' countries) in the European Social Survey.
- Children: Number of biological or adopted children that the participant reports having. Collected in all rounds of fieldwork.
- Children (Census): The total number of children that the census respondent reports that she had throughout her life up until the time of the census. Taken from the 1973 and 2005 Colombian censuses.
- Climatic Attributes - Precipitation: Annual average precipitation. Variable calculated with data from WorldClim 2.1 climate data.
- Climatic Attributes - Precipitation Seasonality: Coefficient of variation of annual precipitation. Variable calculated with data from WorldClim 2.1 climate data.
- Climatic Attributes - Temperature: Mean annual temperature in degrees Celsius. Variable calculated with data from WorldClim 2.1 climate data.
- Climatic Attributes - Temperature Seasonality: Standard deviation of mean annual temperature, times 100. Variable calculated with data from WorldClim 2.1 climate data.
- Climatic Attributes - Ruggedness: Average terrain ruggedness index in millimeters in the 30 arc-seconds grid. Variable calculated with data from Nunn and Puga (2012).
- Climate (Motive for out-migration): Indicator variable that equals 1 if the census respondent out-migrated from their former municipality of residence, and the reported reason for emigrating is climatic hazards. Variable constructed with data from the Colombian population census of 2005
- Conflict 1900-1917: Indicator variable that equals one if there was a land conflict in the municipality during the 1900-1917 period. Taken from Panel Municipal-CEDE, Universidad de Los Andes.
- Credit-Funded: Indicator variable that equals one if the people that own the housing unit funded the purchase with credit (regardless of the type of entity that issued the credit). Taken from the 1973 Colombian housing census.

- **Deceased Children (Census):** Indicator variable that equals one if the total number of children that the census respondent reported is different from the total number of children that were still alive at the time of the Census. Taken from the 1973 and 2005 Colombian censuses.
- **Direct Ancestors:** Indicator variable that equals one if, when asked from who or where did they learn the historical information they reported, participants mentioned parents or grandparents. Variable constructed with information collected in round 3 of fieldwork.
- **Discriminated Against:** Indicator variable that equals one if the respondent claims they belong to a group discriminated against in their country of residence. Variable constructed with data from second-generation migrants (i.e., respondents born in a country different than their parents' countries) in the European Social Survey.
- **Doctors Uninformed:** Indicator variable that equals one if, when asked about the historical role of physicians in the leper colony, the participant mentioned that doctors were uninformed. Specifically, the indicator equals one if the participant's response includes one of the following root words in Spanish: "mentira" "descono" "desinf" "ignora" "creen" "experim" "ensaya" "no sabía" "no entendía" "creía" "poco conocimiento", or "falta de conocimiento". Variable constructed with information collected in round 3 of fieldwork.
- **Economic (Motive for out-migration):** Indicator variable that equals 1 if the census respondent out-migrated from their former municipality of residence, and the reported reason for emigrating is either education or employment. Variable constructed with data from the Colombian population census of 2005
- **Family Reasons (Motive for out-migration):** Indicator variable that equals 1 if the census respondent out-migrated from their former municipality of residence, and the reported reason for emigrating is related to family matters. Variable constructed with data from the Colombian population census of 2005
- **Family Separation:** Indicator variable that equals one if, when asked about the historical role of physicians in the leper colony, the participant mentioned that those in the colony were separated from their families. Specifically, the indicator equals one when the response has one of the following root words in Spanish: "separa", "famili", or "hij". Variable constructed with information collected in round 3 of fieldwork.

- Female: Binary variable that takes the value of one if the respondent is female. Collected in all rounds of fieldwork.
- Female (Census): Binary variable that takes the value of one if the census respondent is female and zero otherwise. Taken from the 1973 and 2005 Colombian censuses.
- Follow Rules: Indicator variable that equals one if the respondent claims that it is very important to follow rules. Variable constructed with data from second-generation migrants (i.e., respondents born in a country different than their parents' countries) in the European Social Survey.
- Forced Displacement: Indicator variable that equals one if, when asked about the historical role of physicians in the leper colony, the participant mentioned that those with the disease were forcefully displaced to the leper colony. Specifically, the indicator equals one when the response has one of the following root words in Spanish: “desplaz”, “traía”, “a la fuerza”, or “traslad”. Variable constructed with information collected in round 3 of fieldwork.
- Has a Kitchen: Indicator variable that equals one if the housing unit has a kitchen (regardless of whether the kitchen is private or shared with another housing unit). Taken from the 1973 Colombian housing census.
- Has a Toilet: Indicator variable that equals one if there is a toilet inside the housing unit. Taken from the 1973 Colombian housing census.
- Has Electricity: Indicator variable that equals one if the housing unit has an electrical connection. Taken from the 1973 Colombian housing census.
- Has No Ground Slab: Indicator variable that equals one if the housing has no foundation slab but is instead built on top of the existing ground. Taken from the 1973 Colombian housing census.
- Has Potable Water: Indicator variable that equals one if the housing unit has access to potable water via a private pipeline. Taken from the 1973 Colombian housing census.
- Has Sewerage: Indicator variable that equals one if the housing unit is connected to the sewerage system. Taken from the 1973 Colombian housing census.
- Have a Good Time: Indicator variable that equals one if the respondent claims that it is very important to have a good time in life. Variable constructed with data from second-generation migrants (i.e., respondents born in a country different than their parents' countries) in the European Social Survey.

- Health (Motive for out-migration): Indicator variable that equals 1 if the census respondent out-migrated from their former municipality of residence, and the reported reason for emigrating is health. Variable constructed with data from the Colombian population census of 2005
- Healthy Arms: Indicator variable that equals 1 if the census respondent reports that they have no health issues that impede arm movement. Variable constructed with data from the Colombian population census of 2005
- Healthy Ears: Indicator variable that equals 1 if the census respondent reports that they have no health issues that affect hearing. Variable constructed with data from the Colombian population census of 2005
- Healthy Eyes: Indicator variable that equals 1 if the census respondent reports that they have no health issues that affect vision. Variable constructed with data from the Colombian population census of 2005
- Health Independence: Indicator variable that equals 1 if the census respondent reports that they have no issues that impede health independence from others. Variable constructed with data from the Colombian population census of 2005
- Healthy Voice: Indicator variable that equals 1 if the census respondent reports that they have no health issues that affect speaking. Variable constructed with data from the Colombian population census of 2005
- Healthy Legs: Indicator variable that equals 1 if the census respondent reports that they have no health issues that impede walking. Variable constructed with data from the Colombian population census of 2005
- Honesty: Variable that captures the intensity with which honesty was inculcated in children of a given ethnicity, on a 1-10 scale. Variable taken from the Standard Cross-Cultural Sample.
- ICW Index - Altruism: ICW Index (Anderson, 2008) that captures the latent altruistic trait of participants in the first round of fieldwork. The index was constructed with the two experimental measures of altruism (ingroup and outgroup allocations) and the two survey measures of altruism (self-reported altruism and reported solidarity with Venezuelan refugees) collected in the first round of fieldwork.

- In Stigmatizing: Indicator variable that equals one if, when asked about the historical role of physicians in the leper colony, the participant mentioned that those in the colony were stigmatized or mistreated. Specifically, the indicator equals one when the response has one of the following root words in Spanish: “maltra”, “discrim”, “stigm”, “desprec”, “preso”, “prejuic”, or “cruel”. Variable constructed with information collected in round 3 of fieldwork.
- Income (Census): Total income in Colombian pesos of 1973 earned the month before the census. Taken from the 1973 Colombian census.
- Industriousness: Variable that captures the intensity with which industriousness was inculcated in children of a given ethnicity, on a 1-10 scale. Variable constructed with data from the Standard Cross-Cultural Sample.
- Institutional: Indicator variable that equals one if, when asked from who or where did they learn the historical information they reported, participants mentioned school, teachers, radio, museums, politicians, government entities, research, or universities. Variable constructed with information collected in round 3 of fieldwork.
- Isolation of the Ill: Indicator variable that equals one if, when asked about the historical role of physicians in the leper colony, the participant mentioned that those with the disease were isolated. Specifically, the indicator equals one when the response has one of the following root words in Spanish: “aisla”, “interna”, “encierro”, “exclu”, or “aparta”. Variable constructed with information collected in round 3 of fieldwork.
- Labor Force Participation (Census): Indicator variable that equals one if the census respondent who reported not being retired mentioned that in the past week she (1) worked, (2) did not work but had a job, (3) sought a job in the past week and had held a job before, or (4) sought a job for the first time in the past week. Taken from the 1973 and 2005 Colombian censuses.
- Leishmaniasis: Indicator variable that equals 1 if leishmaniasis was widespread and prevalent in the ethnicity. Variable constructed with data from the Standard Cross-Cultural Sample.
- Leprosy: Indicator variable that equals 1 if leprosy was widespread and prevalent in the ethnicity. Variable constructed with data from the Standard Cross-Cultural Sample.
- Leprosy Incidence: Number of leprosy cases per municipality that were sent to the leper colony of Agua de Dios in 1920, divided by the municipality’s population in 1918, times 10,000. Leprosy cases

per municipality come from a full count of the patients coming from each municipality, as recorded by the colony's chief doctor in 1920. Population per municipality comes from the population census of 1918 in Colombia (Uribe-Castro, 2019).

- Literacy (Census): Indicator variable that equals one if the census respondent reports knowing how to read and write. Taken from the 1973 and 2005 Colombian censuses.
- Local Trust (SCCS): Variable that captures the intensity with which trust in community members was inculcated in children of a given ethnicity, on a 1-10 scale. Variable taken from the Standard Cross-Cultural Sample.
- Loyal to Close Contacts: Indicator variable that equals one if the respondent claims that it is very important to be loyal to people that are close. Variable constructed with data from second-generation migrants (i.e., respondents born in a country different than their parents' countries) in the European Social Survey.
- Malaria: Indicator variable that equals 1 if malaria was widespread and prevalent in the ethnicity. Variable constructed with data from the Standard Cross-Cultural Sample.
- Male Share 1918: Municipality's male population in 1918 divided by its total population in that year. Variable constructed with data from the 1918 census (Uribe-Castro, 2019).
- Male Share 1928: Municipality's male population in 1928 divided by its total population in that year. Variable constructed with data from the 1928 census (Uribe-Castro, 2019).
- Marital: Binary variable that takes the value of one if the participant reports being married or having a permanent partner for two years or more. Collected in all rounds of fieldwork.
- Marital Status (Census): Binary variable that takes the value of one if the census respondent reports being married or having a permanent partner (i.e., not legally married but living with a partner for two years or more). Taken from the 1973 and 2005 Colombian censuses. The 1973 census does not have information about partners being permanent or not, so marital status in 1973 is only equal to one if the census respondent reports being married.
- Name Common: Indicator variable that equals 1 if the participant has a name identified as common by Colombia's National Civil Registry. The list includes 24 names: Maria, Jose, Luis, Luz, Ana, Carlos, Juan, Antonio, Jesus, Alberto, Enrique, Marina, Carmen, Eduardo, Sandra Milena, Santiago, Valentina, Sebastian, Daniela, Mariana, Natalia, Alejandro, Nicolas, and Samuel. List of names

taken from Colombia's National Civil Registry. Variable constructed with information collected in round 3 of fieldwork.

- Name Religious: Indicator variable that equals 1 if the participant has a name of a religious figure mentioned in the Bible. This list includes the following names: Miryan, Ana, José, Uriel, Miguel, Elizabeth, María, Claudia, Pedro, Juan, María, Ligia, John, Miriam, Daniel, Juan, José, Isabel, Susana, Héctor, Jordan, Pablo, Isabel, Sara, Ruth, Lidia, Cristina, Matías, Esteban, Adán, Mariana, Mary, Martha, Marta, Maribel, Andrés, Jonathan, Joseph, or Jospheh. Variable constructed with information collected in round 3 of fieldwork.
- Negative Sentiment: Indicator variable that equals one if, when asked about the historical role of physicians in the leper colony, the participant's answer was characterized by negative sentiment. The sentiment of the participant's response was computed using Rinker (2021)'s "sentiment R", as described in greater detail in Appendix Section A.33. Variable constructed with information collected in round 3 of fieldwork.
- Nomad (Motive for out-migration): Indicator variable that equals 1 if the census respondent out-migrated from their former municipality of residence, and the reported reason for emigrating is due to a nomadic lifestyle. Variable constructed with data from the Colombian population census of 2005
- Obedience: Variable that captures the intensity with which obedience was inculcated in children of a given ethnicity, on a 1-10 scale. Variable constructed with data from the Standard Cross-Cultural Sample.
- Oblique or Horizontal: Indicator variable that equals one if, when asked from who or where did they learn the historical information they reported, participants mentioned neighbors, friends, family friends, close contacts, or relatives other than direct ancestors. Variable constructed with information collected in round 3 of fieldwork.
- Out-migrant: Indicator variable that equals 1 if the census respondent resides in a municipality in 2005 that is different than their municipality of residence five years before. Variable constructed with data from the Colombian population census of 2005
- Own House: Indicator variable that equals one if the people residing in the housing unit were its owners (regardless of whether such a unit had been fully paid or not). Taken from the 1973 Colombian housing census.

- Population Density 1918: Municipality's number of people in 1918 divided by its total area in squared kilometers. Variable constructed with data from the 1918 census (Uribe-Castro, 2019) and municipalities' areas (from Colombia's national statistical department, DANE).
- Population Density 1928: Municipality's number of people in 1928 divided by its total area in squared kilometers. Variable constructed with data from the 1928 census (Uribe-Castro, 2019) and municipalities' areas (from Colombia's national statistical department, DANE).
- Primary: Binary variable that takes the value of one if the participant reports having a complete primary education. Collected in all rounds of fieldwork.
- Primary (Census): Binary variable that takes the value of one if the census respondent reports having a complete primary education or higher. Taken from the 1973 and 2005 Colombian censuses.
- Relevance of Neighbors' Opinions: Response to the question "On a scale from 1 to 10, with the former being not at all and the latter being very much, how much do you care about the opinions that your neighbors hold of you?". Variable collected in round 3 of fieldwork.
- Relevance of Superiors' Opinions: Response to the question "On a scale from 1 to 10, with the former being not at all and the latter being very much, how much do you care about the opinions that superiors (bosses, local authorities, people in position of power) hold of you?". Variable collected in round 3 of fieldwork.
- Religiosity: Response to the question "On a scale from 1 to 10, with the former being not relevant at all and the latter being the most relevant, how central is God in your life? Variable collected in round 1 of fieldwork.
- Reports (Historical) Knowledge: Indicator variable that equals one if the participant answered something when asked to share in their own words what they knew about the historical role of physicians in the leper colony. Variable collected in round 3 of fieldwork.
- Responsibility: Variable that captures the intensity with which responsibility was inculcated in children of a given ethnicity, on a 1-10 scale. Variable constructed with data from the Standard Cross-Cultural Sample.
- Role of Religious Orders: Indicator variable that equals one if, when asked about the historical role of physicians in the leper colony, the participant mentioned the role of religious officials or the church

more broadly in the former colony. Specifically, the indicator equals one when the response has one of the following root words in Spanish: "padre", "cura", "iglesia", "relig", or "salesia". Variable constructed with information collected in round 3 of fieldwork.

- Roof with Concrete Slabs: Indicator variable that equals one if the roof of the housing unit has a concrete slab. Taken from the 1973 Colombian housing census.
- Secondary: Binary variable that takes the value of 1 if the participant reports having a complete secondary education. Collected in all rounds of fieldwork.
- Secondary (Census): Binary variable that takes the value of one if the census respondent reports having a complete secondary education or higher. Taken from the 1973 and 2005 Colombian censuses.
- Security Threat (Motive for out-migration): Indicator variable that equals 1 if the census respondent out-migrated from their former municipality of residence, and the reported reason for emigrating is security threats. Variable constructed with data from the Colombian population census of 2005
- Selfless Altruism: Response to the question "On a scale from 1 to 10, with the former being 'nothing like me' and the latter being 'identical to me', how similar do you consider yourself to be to someone who selflessly helps and assists others?" Collected in round 1 of fieldwork.
- Sleeping Sickness: Indicator variable that equals 1 if sleeping sickness was widespread and prevalent in the ethnicity. Variable constructed with data from the Standard Cross-Cultural Sample.
- Spirochetes: Indicator variable that equals 1 if spirochetes was widespread and prevalent in the ethnicity. Variable constructed with data from the Standard Cross-Cultural Sample.
- Trust in Colombians: Response to the question "On a scale from 1 to 10, how much do you trust Colombians more generally?" Collected in round 3 of fieldwork.
- Trust in Colombian Government: Response to the question "On a scale from 1 to 10, how much do you trust the Colombian government?" Collected in round 3 of fieldwork.
- Trust in COVID Vaccines: Response to the question "On a scale from 1 to 10, with the former being 'not at all safe or effective' and the latter being 'completely safe and effective', how much do you trust the safety and efficacy of the vaccines against COVID-19?" Collected in round 3 of fieldwork.

- Trust in Engineers: Response to the question “On a scale from 1 to 10, with the former being no trust at all and the latter being complete trust, how much do you trust engineers?” Collected in round 2 of fieldwork.
- Trust in HPV Vaccine: Outcome that captures participant’s trust in the safety of the vaccine against the Human Papillomavirus, on a 1 to 10 scale (higher values corresponding to higher trust in the vaccine). The original question that participants answered was the following: “On a scale from 1 to 10, with the former being ‘not dangerous at all’ and the latter being ‘extremely dangerous,’ how dangerous do you believe the vaccine against the Human Papillomavirus is?” The outcome was then re-scaled so that higher values corresponded to higher trust in the safety of the vaccine (the outcome examined in the paper is thus defined as 11 minus the original numerical value collected).
- Trust in Physicians: Response to the question “On a scale from 1 to 10, with the former being no trust at all and the latter being complete trust, how much do you trust physicians?” Collected in rounds 1 and 2 of fieldwork.
- Trust in Researchers: Response to the question “On a scale from 1 to 10, with the former being no trust at all and the latter being complete trust, how much do you trust university researchers?” Collected in round 1 of fieldwork.
- Trust in Lawyers: Response to the question “On a scale from 1 to 10, with the former being no trust at all and the latter being complete trust, how much do you trust lawyers?” Collected in round 2 of fieldwork.
- Trust in Locals (Fieldwork): Response to the question “On a scale from 1 to 10, how much do you trust people who live in these surroundings?” Collected in round 3 of fieldwork.
- Vaccination Rate of Newborns: Rate of infant vaccination in a given year and given municipality. Data is comprised of infant vaccination rates for (BCG, inactivated polio vaccine, the vaccine against hepatitis B, Hib vaccine against type b influenza, DTaP vaccine against diphtheria-tetanus-pertussis, MMR vaccine against measles-mumps-rubella, and the vaccine against yellow fever) between 1998 and 2010. Data comes from Colombia’s Health Ministry.
- Walls Made Out of Bricks: Indicator variable that equals one if the walls of the housing unit are made out of bricks. Taken from the 1973 Colombian housing census.

Appendix A10 - Summary Statistics of Variables Collected in the 3 Rounds of Fieldwork

	(1)	(2)	(3)	(4)	(5)
Variable	Obs	Mean	Std	Min	Max
Panel A - Demographic Characteristics					
Ancestry	900	0.439	0.497	0	1
Female	900	0.547	0.498	0	1
Age	900	45.438	17.420	17	93
Children	900	1.607	1.607	0	11
Marital	900	0.498	0.500	0	1
Primary	900	0.853	0.354	0	1
Secondary	900	0.579	0.494	0	1
Panel B - Incentivized Measures					
Alloc In-group (1st Fieldwork)	265	4.189	3.354	0	16
Alloc Out-group (1st Fieldwork)	265	4.649	3.908	0	16
Alloc Total (1st Fieldwork)	265	8.838	6.666	0	32
Alloc Difference (1st Fieldwork)	265	0.460	2.936	-12	16
Alloc In-group (2nd Fieldwork)	360	70.847	89.823	0	800
Alloc Out-group (2nd Fieldwork)	360	171.992	140.557	0	800
Alloc Total (2nd Fieldwork)	360	242.839	192.259	0	1,600
Alloc Difference (2nd Fieldwork)	360	101.144	136.694	-200	800
Pill (Antiparasitic Take-up)	360	0.314	0.465	0	1
Panel C - Interventions (Second Round of Fieldwork)					
Exclusion-Hist	360	0.256	0.437	0	1
Medicine-Hist	360	0.261	0.440	0	1
Flora-Hist	360	0.225	0.418	0	1
Panel D - Survey Measures					
Altruism Venezuelan Refugees	265	7.491	2.379	1	10
Altruism Large Businessowners	275	2.898	2.522	1	10
Concern for Neighbors' Opinions	275	2.687	2.515	1	10
Conocern for Superiors' Opinions	274	7.201	2.692	1	10
Direct Ancestor	275	0.345	0.476	0	1
Doctors Uninformed	275	0.127	0.334	0	1
Family Separation	275	0.109	0.312	0	1
Forced Displacement	275	0.033	0.178	0	1
Reports (Historical) Knowledge	275	0.880	0.326	0	1
Institutional	275	0.156	0.364	0	1
Isolation of the Ill	275	0.244	0.430	0	1
Name Religious	275	0.324	0.469	0	1
Name Common	275	0.291	0.455	0	1
Oblique or Horizontal	275	0.236	0.426	0	1
Role of Religious Orders	275	0.102	0.303	0	1
Negative Sentiment	242	0.566	0.497	0	1
Stigma	275	0.211	0.409	0	1
Trust Colombians	275	5.153	2.428	1	10
Trust COVID Vaccine	273	6.777	2.257	1	10
Trust Colombian Government	275	2.764	2.218	1	10
Trust HPV Vaccine	220	6.645	3.261	1	10
Trust Physicians	625	7.021	2.410	1	10
Trust Engineers	360	5.908	2.570	1	10
Trust Lawyers	360	4.544	2.552	1	10
Trust Locals	275	6.269	2.686	1	10
Trust Researchers	265	8.008	1.958	1	10

Appendix A.11 - Characteristics of Respondents in Round 2 vs Rounds1-3

	(1)	(2)	(3)	(4)	(5)
	Second Round = 1 Mean	Second Round = 0 Mean	OLS Coefficient on <i>Second Round</i>	SE	Obs.
Panel A - Participants who reported not having an excluded ancestor					
Female	0.497	0.561	-0.064	(0.047)	505
Age	46.407	44.652	1.754	(1.577)	505
Primary	0.808	0.841	-0.034	(0.036)	505
Secondary	0.537	0.512	0.025	(0.047)	505
Marital Status	0.486	0.503	-0.017	(0.047)	505
Children	1.994	1.390	0.604	(0.148)	505
Panel B - Participants who reported having an excluded ancestor					
Female	0.519	0.590	-0.070	(0.050)	395
Age	46.962	44.528	2.433	(1.798)	395
Primary	0.858	0.906	-0.048	(0.033)	395
Secondary	0.650	0.656	-0.005	(0.048)	395
Marital Status	0.514	0.486	0.028	(0.051)	395
Children	1.978	1.297	0.681	(0.167)	395
Panel C - Participants who reported having an excluded ancestor - 2nd vs 1st Round					
Female	0.519	0.585	-0.066	(0.059)	301
Age	46.962	43.085	3.877	(2.048)	301
Primary	0.858	0.856	0.002	(0.042)	301
Secondary	0.650	0.627	0.023	(0.057)	301
Marital Status	0.514	0.492	0.022	(0.059)	301
Children	1.978	1.720	0.258	(0.218)	301
Panel D - Participants who reported having an excluded ancestor - 2nd vs 3rd Round					
Female	0.519	0.596	-0.077	(0.063)	277
Age	46.962	46.340	0.621	(2.273)	277
Primary	0.858	0.968	-0.110	(0.032)	277
Secondary	0.650	0.691	-0.041	(0.060)	277
Marital Status	0.514	0.479	0.035	(0.064)	277
Children	1.978	0.766	1.212	(0.139)	277

Note: The table shows differences in demographic characteristics between study participants in the second round of fieldwork, when reported ancestry was incentivized, and study participants in other rounds of fieldwork. Each panel examines these differences for a subset of participants, per the details in the title of each panel. Columns 1 and 2 show the mean value of each variable in the row headings for those who participated and did not participate in the second round of fieldwork, respectively. Column 3 shows the coefficients from regressions of each variable in the row heading on an indicator that equals 1 for individuals who participated in the second round of fieldwork, while Column 4 presents robust standard errors of such coefficient. Column 5 shows the number of observations for each regression.

Appendix A.12 - Validation of Experimental Measures

	Altruism				Trust in Medicine	
	(1)	(2)	(3)	(4)	(5)	(6)
	In-group Allocation	Out-group Allocation	Total Allocation	Altruism Venezuelan Refugees	Trust in HPV Vaccine	Antiparasitic Take-Up
Self-Reported Altruism	0.217 (0.119)	0.192 (0.096)	0.409 (0.198)	0.253 (0.084)		
Trust in Physicians					0.420 (0.110)	0.004 (0.011)
Covariates						
Female	-0.673 (0.581)	0.180 (0.489)	-0.494 (0.977)	0.657 (0.325)	-0.367 (0.425)	0.050 (0.050)
Age	-0.023 (0.094)	0.001 (0.077)	-0.022 (0.146)	0.009 (0.051)	-0.150 (0.067)	-0.001 (0.009)
Marital	-0.266 (0.506)	0.115 (0.438)	-0.151 (0.863)	-0.265 (0.306)	0.494 (0.460)	-0.086 (0.052)
Primary	-1.108 (0.759)	-0.757 (0.541)	-1.866 (1.113)	0.565 (0.470)	0.606 (0.734)	0.082 (0.082)
Secondary	1.843 (0.633)	0.926 (0.522)	2.769 (1.047)	-0.512 (0.373)	0.081 (0.588)	-0.107 (0.063)
Age square	0.001 (0.001)	0.000 (0.001)	0.001 (0.002)	-0.000 (0.001)	0.001 (0.001)	0.000 (0.000)
Children	-0.012 (0.163)	-0.207 (0.103)	-0.219 (0.216)	-0.019 (0.100)	0.122 (0.167)	0.002 (0.019)
Observations	265	265	265	265	220	360
R-squared	0.048	0.042	0.042	0.083	0.122	0.029
Round 1	✓	✓	✓	✓	✓	
Round 2						✓
Round 3						
Mean Dep. Var	4.649	4.189	8.838	7.491	6.645	0.314

Note: The table examines the correlations between the experimental outcomes and views and attitudes as reported by participants in survey responses. Each column reports results from a single OLS regression of the dependent variable described in the column heading on the variables described in the row headings. Self-reported altruism is the reported level of altruism that each participant gave on a scale from 1 to 10. Trust in physicians is the level of trust (on a 1-10 scale) that each study participant reported toward such professionals. In-group Allocation is the total amount given by each participant in a Dictator Game to an anonymous receiver from the same municipality. Out-group Allocation is the amount given to an anonymous receiver from Apulo, a municipality outside the region where the Dictator Games were conducted. Total Allocation is the total sum of what each participant gave in both rounds of Dictator Games. Allocations are in thousand Colombian Pesos of 2018 (1 USD was approximately 2,900 COP at the time). Participants allocated 16,000 at most in each round of Dictator Games (i.e., they allocated COP 32,000 at most in total). Altruism Venezuela Refugees is the reported support for policies to assist Venezuelan refugees, on a scale from 1 to 10. Trust in HPV Vaccine is the trust reported by study participants in the safety and efficacy of the vaccine against the Human Papilloma Virus (on a 1-10 scale). Antiparasitic take-up is an indicator that equals 1 for those participants who redeemed a dose of antiparasitics offered for free to every study participant in the second round of fieldwork. Round denotes the round of fieldwork in which the outcome in the column heading was collected. Heteroskedastic-robust standard errors in parentheses.

Appendix A.13 - Infant Vaccinations

Considering the documented aversion toward medicine and physicians, are there observable legacies in health choices? To answer this question, I turn to vaccination records of newborns in Agua de Dios, its adjacent municipalities, and municipalities that are distant from the region but are nonetheless comparable in geo-climatic conditions. These vaccines are provided for free by the state, which implies that differences in rates are mostly demand rather than supply-driven. Although data at the ancestry level are not available (i.e., it is unfeasible to compare the vaccination rates of descendants of the excluded versus descendants of the non-excluded), municipality-level vaccination rates provide suggestive evidence of an adverse legacy on the take-up of vaccines, considering that more than three-fourths of descendants of the excluded live in Agua de Dios. If wariness toward doctors mapped on these health choices, one would expect a significantly worse performance of Agua de Dios in the rates of various vaccines administered for free to newborns (children less than one year of age).

I thus examine the municipality-year vaccination rate for seven types of vaccines provided for free to newborns (BCG, inactivated polio vaccine, the vaccine against hepatitis B, Hib vaccine against type b influenza, DTaP vaccine against diphtheria-tetanus-pertussis, MMR vaccine against measles-mumps-rubella, and the vaccine against yellow fever) for the years 1998–2010 in the municipalities of the region of study (Tocaima, Ricaurte, and Agua de Dios). I use that information to estimate the following expression:

$$Y_{m,t,p} = \alpha + \beta \text{AguaDios}_m + \gamma_t + \gamma_p + \epsilon_{m,t,p}$$

Where $Y_{m,t,p}$ is the rate of infant vaccination in municipality m for the type of vaccine p and year t , and γ_t and γ_p are year and type-of-vaccine fixed effects, respectively. The coefficient of interest, β , captures the average difference in vaccination rates between the former leper colony and its contiguous municipalities, accounting for vaccine-type and year-specific shocks.

The results, displayed in the table below, show that vaccination rates of newborns in the former leper colony are significantly worse than those of climatically similar (but distant) municipalities by as much as 6 percentage points (see Appendix A.8 for details on how these climatically similar municipalities were selected). More importantly, the rate of infant vaccinations in Agua de Dios is 22 percentage points below the rates of its nearest neighbors. The fact that these vaccines are provided for free by the state suggests that the results are demand driven and can likely be explained

by attitudes toward medicine. Hence, the survey evidence presented in Table 4 is indeed capturing an overall mistrust that is conducive to worse health choices and outcomes. All in all, these results reinforce existing findings in the literature (Alsan et al., 2018; Lowes and Montero, 2021b; Martinez-Bravo and Stegmann, 2022), demonstrating the relevance of particular historical processes when understanding the differential success/failure of health policies and health outcomes today.

	Vaccination Rate of Newborns			
	(1)	(2)	(3)	(4)
Agua de Dios	-0.225 (0.028)	-0.225 (0.025)	-0.059 (0.033)	-0.059 (0.033)
Mean Dep. Var	0.859	0.859	0.748	0.748
Climatically Similar			✓	✓
Contiguous Municipalities	✓	✓		
Type FE		✓		✓
Year FE	✓	✓	✓	✓
Observations	240	240	240	240
R-squared	0.650	0.739	0.351	0.362

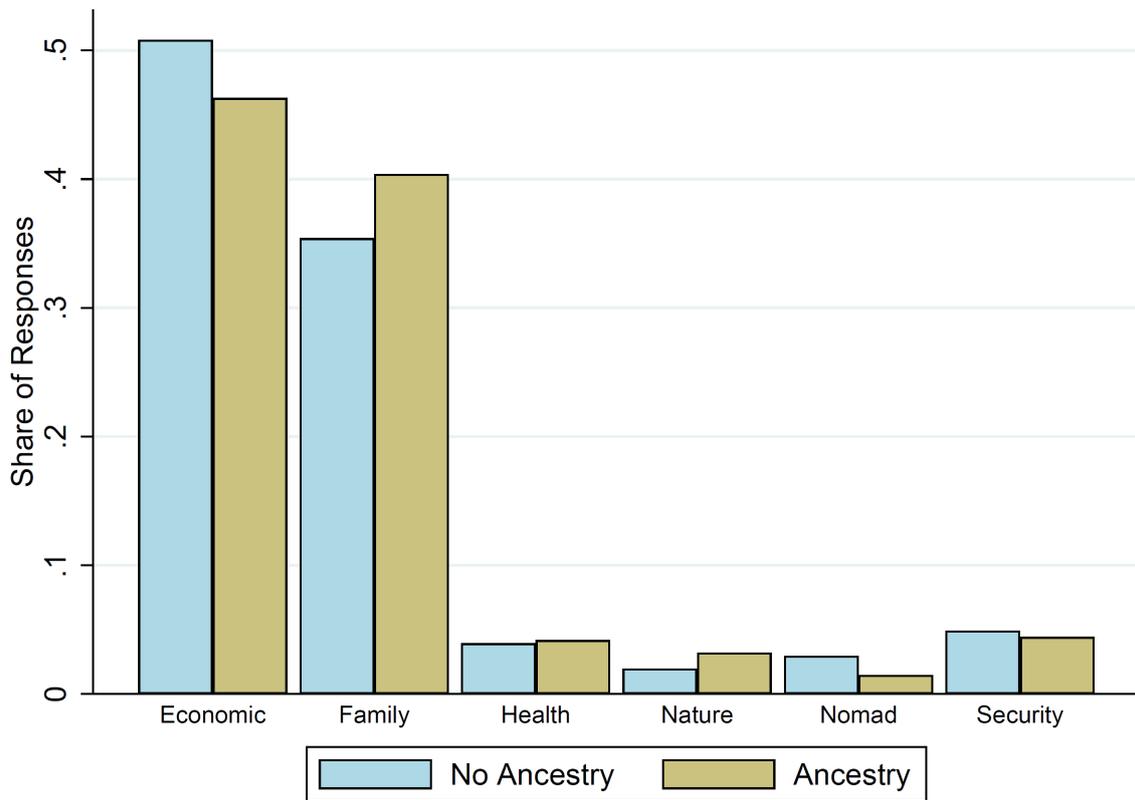
Note: The table documents the negative legacy of the medically-induced social exclusion of Agua de Dios for current health choices, as demonstrated by the lower municipality-level vaccinations rates of newborns (less than 1 year) for several vaccines (BCG, inactivated polio vaccine, the vaccine against hepatitis B, Hib vaccine against type b influenza, DTaP vaccine against diphtheria tetanus-pertussis, MMR vaccine against measles-mumps-rubella, and the vaccine against yellow fever) between 1998 and 2010. Data comes from Colombia’s Health Ministry, and the unit of observation is a municipality’s rate of vaccination for a given vaccine in a given year. Agua de Dios is a dummy variable that takes the value of 1 for the vaccination rates in Agua de Dios and 0 otherwise. Columns 1-2 report the coefficients from a regression where the control group is provided by the vaccination rates of the contiguous municipalities of Tocaima and Ricaurte. Columns 3-4 report the coefficients from regressions in which the control group is given by the vaccination rates in Mariquita and Puerto Salgar, which are the municipalities closest to Agua de Dios along climatic characteristics in all of Colombia and in the sub-national unit where the region of study is located, respectively (see Appendix Sections A.4 and A.8). Type FE correspond to fixed effects for the type of vaccine (BCG, inactivated polio vaccine, the vaccine against hepatitis B, Hib vaccine against type b influenza, DTaP vaccine against diphtheria-tetanus-pertussis, MMR vaccine against measles-mumps-rubella, and the vaccine against yellow fever). Year FE are fixed effects for the year of recorded municipality-level vaccination rates. Heteroskedastic-robust standard errors are in parentheses.

Appendix A.14 - Reported Health Outcomes in Census - Respondents Originally From the Colony vs Those Originally From Contiguous Region

	(1)	(2)	(3)	(4)	(5)	(6)
	Healthy Legs	Healthy Arms	Healthy Eyes	Healthy Ears	Healthy Voice	Health Independence
Ancestry	0.018 (0.006)	0.004 (0.006)	0.063 (0.009)	0.019 (0.003)	0.006 (0.001)	0.002 (0.001)
Mean Dep. Var	0.961	0.983	0.937	0.977	0.986	0.993
Observations	10914	10914	10914	10914	10914	10914
R-squared	0.002	0.000	0.012	0.003	0.001	0.000

Note: The table examines differences in self-reported health outcomes between those who report being originally from Agua de Dios and those who report being originally from the contiguous municipalities of Tocaima and Ricaurte. Data comes from the Colombian census of 2005, the unit of observation is an individual, and the sample is restricted to census respondents who report that their original place of residence (five years before the census) was the region of the study (Agua de Dios, Ricaurte or Tocaima). Each column reports results from a single OLS regression of the dependent variable described in the column heading on the variables described in the row heading. Ancestry is an indicator variable that equals 1 if the respondent manifests that they used to live in the municipality that was formerly a leper colony (Agua de Dios) regardless of their place of residence in 2005. Outcome variables in columns 1-5 are indicators that equal 1 if the respondent claims that she enjoys good health with respect to her legs, arms, eyes, ears, and voice, respectively. The outcome in Column 6 is an indicator that equals 1 if the respondent claims that her health status allows her to enjoy physical and material independence from others. Standard errors (in parentheses) are clustered at participants' municipalities of residence in 2005.

Appendix A.15 - Distributions of Reasons for Out-Migrating from the Study's Region, Split by Ancestral Status



Note: The figure shows the share of out-migrants from Agua de Dios (Ancestry) and those from Ricaurte and Tocaima (No Ancestry) that reported each reason for out-migrating from their respective place of origin. Data comes from the Colombian Census of 2005, the unit of observation is an individual, and the sample is restricted to Census respondents who report being originally from the region of the study (Tocaima, Agua de Dios, or Ricaurte) but who out-migrated from their municipalities of origin. Economic is an indicator that equals one if the respondent report out-migrating due to employment or educational reasons. Family is an indicator that equals one if the respondent simply reports that they out-migrated due to reasons related to their family. Nomad is an indicator that equals one if the respondent mentions that they out-migrated due to a nomad lifestyle. Finally, Health, Natural Hazard, and Security are indicators that equal one if the respondent mentions that they out-migrated due to health reasons, climatic risks, or security threats, respectively.

Appendix A.16 - Out-Migration Reasons Regression

	Main reason for out-migrating is ... (1=Yes)					
	(1)	(2)	(3)	(4)	(5)	(6)
	Economic	Climate	Security Threat	Health	Family Reason	Nomad Lifestyle
Panel A - No Covariates						
Ancestry	-0.045 (0.038)	0.012 (0.012)	-0.005 (0.016)	0.003 (0.015)	0.050 (0.037)	-0.015 (0.011)
Observations	711	711	711	711	711	711
R-squared	0.002	0.001	0.000	0.000	0.003	0.003
Mean Dep. Var	0.482	0.027	0.046	0.041	0.383	0.021
Panel B - With Covariates						
Ancestry	-0.060 (0.044)	0.018 (0.018)	-0.025 (0.020)	0.018 (0.018)	0.046 (0.041)	0.002 (0.011)
Covariates						
Female	-0.206 (0.044)	-0.010 (0.016)	0.013 (0.020)	0.033 (0.020)	0.191 (0.041)	-0.021 (0.013)
Age	-0.006 (0.002)	0.001 (0.001)	0.001 (0.001)	0.003 (0.001)	0.001 (0.002)	-0.000 (0.000)
Marital	0.077 (0.048)	-0.002 (0.018)	-0.032 (0.023)	-0.038 (0.022)	-0.031 (0.044)	0.026 (0.010)
Primary	0.077 (0.077)	-0.010 (0.037)	-0.008 (0.039)	-0.048 (0.050)	-0.014 (0.073)	0.002 (0.006)
Secondary	-0.044 (0.051)	0.020 (0.020)	0.020 (0.021)	-0.012 (0.018)	-0.012 (0.047)	0.027 (0.018)
Observations	504	504	504	504	504	504
R-squared	0.078	0.015	0.020	0.061	0.046	0.022
Mean Dep. Var	0.567	0.036	0.048	0.046	0.288	0.016

Note: The table examines differences in reported reasons for migration between those who out-migrated from Agua de Dios and those who out-migrated from Tocaima and Ricaurte. Data comes from the Colombian census of 2005, the unit of observation is an individual, and the sample is restricted to census respondents who at the time of the census lived in a municipality different than the one where they lived five years before the census. Each panel-column reports results from a single OLS regression of the dependent variable described in the column heading on the variable(s) described in the row heading(s). Outcome variables in columns 1-6 are indicators that equal 1 if the respondent reports the reason stated in the column header. Ancestry is an indicator variable that equals 1 for those who report that Agua de Dios was the municipality where they used to live five years before the census. Observations vary across panels because there are some census respondents for whom some demographic characteristics are missing. Heteroskedastic-robust standard errors are in parentheses.

Appendix A.17 Correlates of Altruism in Colombia and the Sub-National Region of the Study

	Altruism (1-6 Scale)					Altruism (Binary)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A - Colombia										
Employment	-0.077 (0.056)				-0.028 (0.056)	-0.023 (0.023)				-0.005 (0.023)
Family Important		0.053 (0.073)			0.034 (0.074)		0.015 (0.030)			0.006 (0.030)
Primary Education			-0.001 (0.074)		0.041 (0.096)			0.015 (0.030)		0.035 (0.038)
Secondary Education				-0.012 (0.056)	0.057 (0.076)				-0.001 (0.022)	0.013 (0.029)
Children					-0.000 (0.018)					-0.005 (0.007)
Marital					-0.017 (0.063)					0.002 (0.024)
Female					0.125 (0.053)					0.050 (0.022)
Age					0.007 (0.002)					0.003 (0.001)
Observations	1506	1506	1506	1506	1506	1506	1506	1506	1506	1506
R-squared	0.001	0.000	0.000	0.000	0.015	0.001	0.000	0.000	0.000	0.013
Mean Dep. Var	5.156	5.156	5.156	5.156	5.156	0.790	0.790	0.790	0.790	0.790

	Altruism (1-6 Scale)					Altruism (Binary)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel B - Subnational Region Where Leper Colony is Located										
Employment	-0.190 (0.136)				-0.084 (0.135)	-0.051 (0.049)				-0.021 (0.053)
Family Important		0.147 (0.177)			0.141 (0.181)		0.079 (0.067)			0.083 (0.069)
Primary Education			0.866 (0.382)		0.883 (0.432)			0.173 (0.110)		0.179 (0.134)
Secondary Education				0.194 (0.180)	-0.046 (0.165)				0.028 (0.060)	-0.017 (0.067)
Children					-0.005 (0.048)					0.004 (0.016)
Marital					0.189 (0.143)					0.048 (0.052)
Female					0.285 (0.121)					0.090 (0.047)
Age					0.007 (0.005)					0.002 (0.002)
Observations	310	310	310	310	310	310	310	310	310	310
R-squared	0.007	0.003	0.043	0.005	0.088	0.004	0.005	0.012	0.001	0.044
Mean Dep. Var	5.123	5.123	5.123	5.123	5.123	0.797	0.797	0.797	0.797	0.797

Note: The table documents the correlations between reported altruism and various demographic characteristics in Colombia and in the sub-national region where the leper colony was located. Data comes from the 6th round of the World Values Survey, and the unit of observation is an individual. Each panel-column reports results from a single OLS regression of the dependent variable described in the column heading on the variable(s) described in the row heading(s). The outcome in Columns 1-5 is given by individual responses to the question "How much like you is someone who claims that it is important to help people living nearby; to care for their need?", which range from "Very much like me" (=1) to "Not at all like me" (=6). Columns 6-10 have as an outcome an indicator that equals 1 for respondents who answered 5 or 6. Family Important is an indicator that equals one if, when asked how important family is in life, the respondent reports that it is very important. Employment is an indicator that equals one if the respondent reports being employed full-time or part-time. Panel A uses data from all Colombian respondents in the WVS who answered such a question, while Panel B uses data from respondents in the sub-national region where the leper colony is located (Cundinamarca). Heteroskedastic-robust standard errors are in parentheses.

Appendix A.18 - Are Outmigrants From the Former Leper Colony Different Than Outmigrants From Its Contiguous Municipalities?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Female	Age	Marital	Primary	Secondary	Employed	Children	Deceased Children
Out-Migrant*Ancestry	0.023 (0.040)	-4.303 (1.499)	0.033 (0.044)	0.011 (0.027)	-0.059 (0.036)	-0.024 (0.024)	0.109 (0.287)	-0.040 (0.046)
Out-Migrant	-0.013 (0.029)	-5.295 (1.141)	0.082 (0.033)	0.029 (0.020)	0.030 (0.026)	0.016 (0.013)	-0.853 (0.206)	-0.135 (0.030)
Mean Dep. Var	0.469	36.381	0.581	0.838	0.266	0.960	3.225	0.210
Observations	10914	10914	7891	10914	10914	3138	3947	3033
R-squared	0.000	0.008	0.005	0.001	0.007	0.003	0.011	0.015

Note: The table compares the observable attributes of out-migrants from the municipality that was formerly a leper colony with those of out-migrants from the contiguous municipalities of Ricaurte and Tocaima. Data comes from the 2005 Census in Colombia, and the unit of observation is an individual. Data is restricted to respondents who report that they used to live in the region of study (Agua de Dios, Ricaurte, or Tocaima) five years before the census (which includes both those who continued to live in those places and those who out-migrated). Each column reports results from a single OLS regression of the dependent variable described in the column heading on the variables described in the row headings and municipality-of-origin fixed effects. Out-migrant is an indicator variable that equals 1 for those who report out-migrating from the municipality where they used to live five years before the census. Ancestry is an indicator variable that equals 1 for those who report that Agua de Dios was the municipality where they used to live five years before the census. Heteroskedastic-robust standard errors are in parentheses.

Appendix A.19 - Study Participants Compared to Respondents in Distant Municipalities

	Self-Reported Altruism			Trust in HPV Vaccine		
	(1)	(2)	(3)	(4)	(5)	(6)
	Excluding Descendants of the Excluded	Excluding Descendants of the Non- Excluded in Region of Study	Full Baseline Sample and Distant Municipalities	Excluding Descendants of the Excluded	Excluding Descendants of the Non- Excluded in Region of Study	Full Baseline Sample and Distant Municipalities
Panel A - No Covariates						
Region of Study	-0.019 (0.222)			0.549 (0.343)		
Ancestry		0.570 (0.206)	0.578 (0.189)		-1.328 (0.402)	-1.598 (0.379)
Observations	341	312	459	254	224	349
R-squared	0.000	0.023	0.017	0.010	0.052	0.057
Panel B - With Baseline Covariates						
Region of Study	0.024 (0.288)			0.405 (0.425)		
Ancestry		0.433 (0.244)	0.616 (0.194)		-0.806 (0.475)	-1.452 (0.396)
Covariates						
Female	0.643 (0.231)	0.250 (0.207)	0.508 (0.186)	-0.286 (0.350)	-0.054 (0.418)	-0.095 (0.321)
Age	0.037 (0.039)	0.037 (0.036)	0.044 (0.029)	-0.148 (0.061)	-0.189 (0.062)	-0.158 (0.051)
Marital	0.164 (0.227)	-0.192 (0.207)	0.049 (0.186)	0.434 (0.359)	-0.017 (0.393)	0.352 (0.325)
Primary	0.262 (0.328)	-0.517 (0.330)	0.124 (0.259)	-0.065 (0.672)	1.422 (0.820)	0.460 (0.571)
Secondary	0.038 (0.200)	-0.077 (0.197)	-0.020 (0.153)	-0.027 (0.374)	0.388 (0.421)	-0.060 (0.305)
Age Squared	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.001 (0.001)	0.002 (0.001)	0.001 (0.001)
Children	0.077 (0.070)	0.095 (0.066)	0.089 (0.057)	0.251 (0.107)	0.129 (0.138)	0.135 (0.103)
Observations	341	312	459	254	224	349
R-squared	0.045	0.072	0.058	0.061	0.107	0.090
Mean Dep Var	7.718	7.942	7.867	7.177	6.344	6.742

Note: The table shows that those who have an excluded ancestor are the ones who stand out as having significantly higher altruism and significantly lower trust in medicine (whereas descendants of non-excluded locals exhibit behaviors that are not statistically or economically different from those of individuals in distant yet similar places). The table examines data from the first round of fieldwork combined with survey data collected in (1) Colombian municipalities that are as similar as possible to the former leper colony in terms of average climatic characteristics, and (2) an economically-depressed neighborhood in Bogota (Barrio Kennedy). The climatically-similar municipalities were selected based on municipalities with the lowest Climatic Dissimilarity Indices (CDI) in Colombia and in the sub-national unit of the study's region (see Online Appendix A.8), and CDI is defined as the sum of the squared percentage differences in the geo-climatic attributes of each municipality c and those of Agua de Dios, $CDI(c) = \sum_{i=1}^N \left(\frac{x_i^c - x_i^{AD}}{x_i^{AD}} \right)^2$, where N is the number of climatic attributes, $c \in C$ are municipalities in Colombia, x_i^c is the average value of each attribute i in municipality c , and x_i^{AD} is the average value of each attribute i in Agua de Dios. The geo-climatic attributes included in the calculation of the $CDI(c)$ are ruggedness, temperature, precipitation, precipitation seasonality, and temperature seasonality. The climatically-similar municipalities were the ones with the lowest CDI in (a) the entire country and (b) in the sub-national unit where the study's region is located (besides the contiguous municipalities). Columns 1 and 4 combine data from the first round of fieldwork with data collected in these distant places, dropping those who have an excluded ancestor, to compare local non-excluded participants in the region of the study with respondents in distant places. Columns 2 and 5 do the same but drop local non-excluded respondents in the region of study to compare participants who are descendants of the excluded with respondents in distant places. Columns 3 and 6 pool the data collected in the first round of fieldwork with that collected in distant places without dropping any participants. Heteroskedastic-robust standard errors in parenthesis.

Appendix A.20 - Are There Differences in Social Conformity or Desirability Between Study Participants With and Without an Excluded Ancestor?

	Susceptibility to Others' Opinions			Baseline Results Accounting for Trust in Researchers			
	(1) Concern for Neighbors' Opinions	(2) Concern for Superiors' Opinions	(3) Trust in Researchers	(4) Total Allocation	(5) In- vs Out- Allocation Difference	(6) Trust in Physicians	(7) Trust in HPV Vaccine
Ancestry	0.307 (0.328)	-0.917 (0.373)	-0.391 (0.242)	3.865 (0.873)	1.287 (0.399)	-1.475 (0.275)	-1.651 (0.462)
Covariates							
Trust in Researchers				-0.036 (0.178)	0.001 (0.092)	0.287 (0.092)	0.277 (0.137)
Female	-0.498 (0.317)	-0.091 (0.331)	-0.165 (0.264)	-0.125 (0.887)	-0.802 (0.406)	0.515 (0.261)	-0.025 (0.423)
Age	0.028 (0.054)	0.012 (0.056)	0.001 (0.038)	-0.021 (0.141)	-0.030 (0.085)	-0.009 (0.043)	-0.126 (0.067)
Marital	0.203 (0.370)	-0.019 (0.361)	-0.100 (0.252)	0.212 (0.836)	-0.274 (0.379)	0.349 (0.270)	0.588 (0.452)
Primary	-0.269 (0.543)	-0.604 (0.558)	0.778 (0.455)	-1.471 (1.110)	-0.299 (0.695)	-0.102 (0.437)	0.295 (0.733)
Secondary	0.051 (0.344)	0.268 (0.397)	0.121 (0.327)	2.017 (0.988)	0.705 (0.472)	0.154 (0.355)	0.284 (0.579)
Age Squared	-0.000 (0.001)	-0.001 (0.001)	0.000 (0.000)	0.000 (0.001)	0.001 (0.001)	0.000 (0.001)	0.001 (0.001)
Children	0.472 (0.490)	0.740 (0.458)	-0.166 (0.095)	-0.057 (0.218)	0.243 (0.163)	-0.026 (0.105)	0.071 (0.173)
Observations	275	274	265	265	265	265	220
R-squared	0.029	0.083	0.058	0.109	0.107	0.197	0.135
Mean Dep. Var	2.687	7.201	8.008	8.838	0.460	7.502	6.645
Round 1			✓	✓	✓	✓	✓
Round 2							
Round 3	✓	✓					

Note: The table examines differences in social desirability and social conformity across ancestry status. Each column reports results from a single OLS regression of the dependent variable described in the column heading on the variables described in the row headings. Column 1 has as an outcome variable participants' reported concern for the opinions and views of neighbors on a 1-10 scale. Column 2 has as an outcome variable participants' reported concern, on a 1-10 scale, for the opinions and views of superiors or those who could be in positions of power. Column 3 has as an outcome variable participants' reported trust in researchers on a 1-10 scale. Outcomes in Columns 4-7 are the same ones as those examined and discussed in Tables 3 and 4 (see Appendix A.9 for definitions of variables). Round denotes the round of fieldwork in which the outcome in the column heading was collected. Heteroskedastic-robust standard errors are in parentheses. Overall, the table suggests that descendants of those who were excluded tend to display a similar concern for the opinions of neighbors as locals who are not descendants of the excluded, similar trust in researchers, and significantly lower concern for the opinions of superiors. It also indicates that the baseline results are immune to controlling for trust in researchers.

Appendix A.21 - Income Distribution

	Participant Belongs to ... (1=Yes)		
	(1)	(2)	(3)
	<i>Estrato 1</i>	<i>Estrato 2</i>	<i>Estrato 3</i>
	Low Income	Medium-Low Income	Medium Income
Panel A - No Covariates			
Ancestry	-0.081 (0.053)	0.061 (0.061)	0.020 (0.040)
R-squared	0.009	0.004	0.001
Panel B - With Baseline Covariates			
Ancestry	-0.056 (0.055)	0.057 (0.064)	-0.001 (0.040)
Covariates			
Female	-0.099 (0.057)	0.052 (0.063)	0.047 (0.037)
Age	0.002 (0.010)	-0.006 (0.011)	0.005 (0.006)
Marital	-0.075 (0.062)	0.086 (0.070)	-0.012 (0.043)
Primary	-0.084 (0.108)	0.032 (0.114)	0.053 (0.051)
Secondary	-0.053 (0.064)	-0.033 (0.071)	0.086 (0.043)
Age square	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)
Children	-0.091 (0.094)	0.043 (0.098)	0.048 (0.048)
Observations	248	248	248
R-squared	0.058	0.019	0.052
Mean Dep Var	0.218	0.690	0.093
Round 1			
Round 2			
Round 3	✓	✓	✓

Note: The table examines the income group of study participants in the third round of fieldwork. It compares the share of those belonging to each estrato, which is a categorization of the housing units that serves as a proxy for the income group of each participant. Each panel-column reports results from a single OLS regression of the dependent variable described in the column heading on the variable(s) described in the row heading(s). Ancestry is an indicator that equals 1 for participants who reported that they have an ancestor who lived in the colony during the epoch of isolation. Outcomes in the column headings are indicators that equal 1 if the respondent claims that she belongs to estrato 1 (low income), estrato 2 (medium-low income), or estrato 3 (medium income). Round denotes the round of fieldwork in which the outcome in the column heading was collected. Heteroskedastic-robust standard errors in parentheses.

Appendix A.22 - Heterogeneity by Income

	Trust in locals	Altruism Large Business Owners	Doctors Uninformed	Doctors' Role In Stigmatizing
	(1)	(2)	(3)	(4)
Panel A - Restricted to Medium Income Respondents				
Ancestry	0.492 (0.387)	0.767 (0.386)	0.135 (0.059)	0.135 (0.063)
R-squared	0.077	0.102	0.115	0.090
Observations	194	194	194	194
Mean Dep var	6.593	3.031	0.134	0.191
Panel B - Restricted to Low Income Respondents				
Ancestry	1.116 (0.907)	0.601 (0.766)	0.019 (0.085)	-0.057 (0.132)
R-squared	0.098	0.069	0.106	0.173
Observations	81	81	81	81
Mean Dep var	5.494	2.580	0.111	0.259
Panel C - Interaction of Ancestry and Low Income Indicator				
Ancestry	0.436 (0.378)	0.752 (0.383)	0.131 (0.058)	0.143 (0.064)
Ancestry × Low Income	1.512 (1.098)	0.088 (0.948)	-0.082 (0.125)	-0.127 (0.159)
Low Income	-1.728 (0.601)	-0.949 (0.416)	-0.008 (0.050)	0.112 (0.078)
R-squared	0.094	0.083	0.089	0.089
Observations	248	248	248	248
Mean Dep var	6.298	2.843	0.125	0.210
Demographic Controls	✓	✓	✓	✓
Round 1				
Round 2				
Round 3	✓	✓	✓	✓

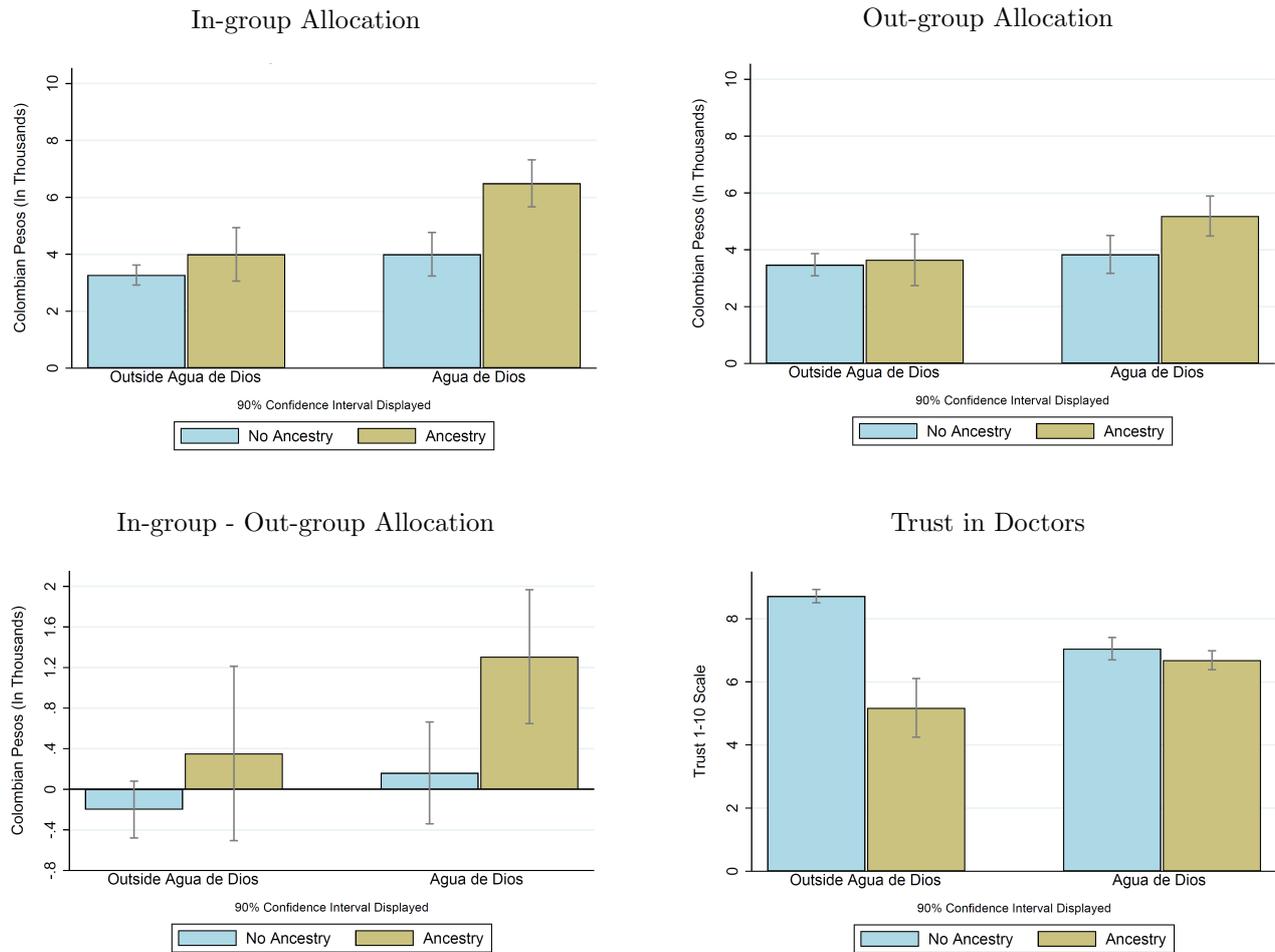
Note: The table examines the heterogeneity of the baseline patterns based on the income of respondents in the third round of fieldwork, which is the only one in which a proxy for participants' income groups was collected. Each panel-column reports results from a single OLS regression of the dependent variable described in the column heading on the variable(s) described in the row heading(s). Panel A restricts the sample to participants who belong to estratos 2 and 3, which correspond to medium-low and medium-income groups. Panel B restricts the sample to estrato 1, which corresponds to the lowest income group. Panel C examines the interaction between having an excluded ancestor and being from a low-income group. Ancestry is an indicator that equals 1 for participants who reported that they have an ancestor who lived in the colony during the epoch of isolation. Low income is an indicator that equals 1 if the participant's reported estrato is 1 (low income). Trust in Locals is participants' reported trust in local people, on a 1-10 scale. Altruism Large Business Owners is the reported support for policies to assist large business owners who struggled during the pandemic. Doctors' Role In Stigmatizing is an indicator variable that equals one if the participant shared a historical narrative that mentions that people in the colony were stigmatized/mistreated by doctors. Doctors Uninformed is an indicator variable that equals one if such a historical narrative mentions that doctors were misinformed about leprosy in the past. Round denotes the round of fieldwork in which the outcome in the column heading was collected. Heteroskedastic-robust standard errors in parentheses.

Appendix A.23 - Heterogeneity by Education

	In-group Allocation	Out-group Allocation	Total Allocation	In-group- Out-group Allocation	Trust in Doctors
	(1)	(2)	(3)	(4)	(5)
Panel A - Restricted to Respondents without Secondary Education					
Ancestry	2.068 (0.727)	0.584 (0.615)	2.651 (1.181)	1.484 (0.648)	-1.317 (0.402)
R-squared	0.148	0.079	0.093	0.196	0.105
Observations	118	118	118	118	194
Mean Dep var	4.169	3.695	7.864	0.475	7.155
Panel B - Restricted to Respondents with Secondary Education					
Ancestry	2.803 (0.706)	1.864 (0.661)	4.667 (1.283)	0.939 (0.473)	-1.397 (0.280)
R-squared	0.168	0.088	0.138	0.115	0.118
Observations	147	147	147	147	245
Mean Dep var	5.034	4.585	9.619	0.449	7.306
Panel C - Interaction of Ancestry and Secondary Education					
Ancestry	1.949 (0.712)	0.571 (0.595)	2.520 (1.145)	1.379 (0.640)	-1.326 (0.397)
Ancestry × Secondary Educ.	1.116 (0.969)	1.278 (0.831)	2.394 (1.627)	-0.161 (0.782)	-0.039 (0.478)
Secondary Educ.	0.875 (0.594)	0.100 (0.542)	0.975 (1.010)	0.775 (0.522)	0.182 (0.315)
R-squared	0.145	0.075	0.117	0.107	0.099
Observations	265	265	265	265	439
Mean Dep var	4.649	4.189	8.838	0.460	7.239
Demographic Controls	✓	✓	✓	✓	✓
Round 1	✓	✓	✓	✓	✓
Round 2					✓
Round 3					

Note: The table examines the heterogeneity of the baseline patterns based on participants' educational attainments. Each panel-column reports results from a single OLS regression of the dependent variable described in the column heading on the variable(s) described in the row heading(s). Panels A and B restrict observations to those without and with secondary education, respectively. Panel C examines the heterogeneity of the results as captured by the interaction between ancestry and an indicator for whether the participant completed secondary education. Ancestry is an indicator that equals one for participants who reported that they have an ancestor who lived in the colony during the epoch of isolation. In-group Allocation is the total amount given by each participant in a Dictator Game to an anonymous receiver from the same municipality. Out-group Allocation is the amount given to an anonymous receiver from Apulo, a municipality outside the region where the Dictator Games were conducted. Total Allocation is the total sum of what each participant gave in both rounds of Dictator Games. In-group-Out-group Allocation is the difference between what a participant assigned to the in-group and to the outgroup recipient. Allocations are in thousand Colombian Pesos of 2018 (1 USD was approximately 2,900 Colombian Pesos), and participants allocated 16,000 at most to each type of recipient (i.e., they allocated COP 32,000 at most). Round denotes the round of fieldwork in which the outcome in the column heading was collected. Observations from the second round of fieldwork correspond to those collected without an intervention (i.e., observations from the second round of fieldwork in this table are restricted to the control group - see Appendix A.34 for further details about the protocol in the second round). Heteroskedastic-robust standard errors in parentheses.

Appendix A.24 - Place-Based Patterns and Heterogeneity



Note: The figure displays the unconditional averages of different outcomes for different subgroups of study participants. For each outcome, the figures show the average for those with and without an excluded ancestor, both inside and outside Agua de Dios (the former leper colony). Also displayed are the 90 percent confidence intervals of such unconditional averages. In-group Allocation is the amount chosen by each participant in a Dictator Game in which she got to allocate up to COP 16,000 to someone from the same municipality (1 USD was approximately 2900 COPs at the time, and COP 16,000 amounted to half of a daily minimum wage). Likewise, Out-group Allocation is the amount chosen by each participant in a Dictator Game in which she got to allocate up to COP 16,000 to someone from a distant municipality called Apulo. In-group - Out-group Allocation is the difference between what each participant gave to someone from their municipality and someone from the distant municipality of Apulo. Trust in Physicians is the level of trust reported by each study participant toward medical doctors (on a 1-10 scale). Ancestry refers to those who are descendants of people who lived in the colony during the epoch of isolation, whereas No Ancestry refers to participants who are descendants of people who did not live in the colony during the epoch of isolation.

Appendix A.25 - Place-Based Heterogeneity Regressions

	In-group Allocation	Out-group Allocation	Total Allocation	In-group- Out-group Allocation	Trust in Doctors
	(1)	(2)	(3)	(4)	(5)
Panel A - Accounting for Municipality Fixed-Effects					
Ancestry	1.837 (0.524)	0.725 (0.454)	2.562 (0.883)	1.112 (0.426)	-1.099 (0.438)
R-squared	0.155	0.078	0.125	0.109	0.162
Panel B - Interaction of Ancestry and 'Agua de Dios' indicator					
Ancestry	0.718 (0.626)	0.136 (0.667)	0.855 (1.145)	0.582 (0.603)	-3.634 (0.605)
Ancestry x Agua de Dios	1.750 (0.943)	0.934 (0.886)	2.684 (1.635)	0.816 (0.822)	3.190 (0.664)
AguaDios	0.615 (0.535)	0.600 (0.513)	1.215 (0.964)	0.015 (0.412)	-1.636 (0.253)
R-squared	0.163	0.080	0.131	0.111	0.173
Ancestry + (Ancestry*Agua Dios)	2.468 (0.718)	1.070 (0.600)	3.538 (1.195)	1.398 (0.569)	-0.444 (0.285)
Observations	265	265	265	265	439
Mean Dep var	4.649	4.189	8.838	0.460	7.239
Demographic Controls	✓	✓	✓	✓	✓
Round 1	✓	✓	✓	✓	✓
Round 2					✓
Round 3					

Note: The table examines the heterogeneity of the baseline patterns based on the location of responses in the first round of fieldwork, when data was collected both inside and outside Agua de Dios (the former leper colony). Each panel-column reports results from a single OLS regression of the dependent variable described in the column heading on the variable(s) described in the row heading(s). Ancestry is an indicator that equals one for participants who reported that they have an ancestor who lived in the colony during the epoch of isolation. In-group Allocation is the total amount given by each participant in a Dictator Game to an anonymous receiver from the same municipality. Out-group Allocation is the amount given to an anonymous receiver from Apulo, a municipality outside the region where the Dictator Games were conducted. Total Allocation is the total sum of what each participant gave in both rounds of Dictator Games. In-group-Out-group Allocation is the difference between what a participant assigned to the in-group and the outgroup anonymous recipients. Allocations are in thousand Colombian Pesos of 2018 (1 USD was approximately 2,900 COPs at the time), and participants allocated COP 16,000 at most to each type of recipient (i.e., their total allocation was COP 32,000 at most). Panel A accounts for municipality-of-residence fixed effects. Panel B examines the heterogeneity of the results as captured by the interaction between Ancestry and an indicator of whether the participant resides in Agua de Dios. Round denotes the round of fieldwork in which the outcome in the column heading was collected. Observations from the second round of fieldwork correspond to those collected without an intervention (i.e., observations from the second round of fieldwork in this table are restricted to the control group - see Appendix A.34 for further details about the protocol in the second round). Heteroskedastic-robust standard errors in parentheses.

Appendix A.26 - Behavioral Legacy of Leprosy in Other Contexts - Ethnic-Level Evidence from Standard Cross-Cultural Sample

	<u>Local Trust</u>	<u>Industriousness</u>	<u>Honesty</u>	<u>Responsibility</u>	<u>Obedience</u>
	(1)	(2)	(3)	(4)	(5)
Panel A - Leprosy					
Leprosy	1.163 (0.589)	-0.467 (0.263)	0.215 (0.687)	-0.397 (0.213)	-0.281 (0.446)
R-squared	0.088	0.088	0.031	0.078	0.114
Panel B - Malaria					
Malaria	-0.119 (0.664)	-0.290 (0.332)	0.523 (0.697)	-0.308 (0.259)	-0.738 (0.510)
R-squared	0.060	0.076	0.037	0.066	0.126
Panel C - Spirochetes					
Spirochetes	-0.481 (0.558)	-0.226 (0.288)	0.093 (0.649)	-0.302 (0.224)	-0.143 (0.431)
R-squared	0.065	0.074	0.030	0.068	0.113
Panel D - Leishmaniasis					
Leishmaniasis	-0.441 (0.545)	0.113 (0.298)	-0.177 (0.578)	-0.241 (0.232)	0.913 (0.505)
R-squared	0.064	0.071	0.031	0.064	0.138
Panel E - Sleeping Sickness					
Sleeping Sickness	-0.115 (0.955)	-0.204 (0.390)	-0.469 (1.389)	-0.046 (0.285)	-0.784 (0.818)
R-squared	0.060	0.072	0.032	0.056	0.121
Observations	135	155	107	144	154
Controls	✓	✓	✓	✓	✓

Note: The table shows that ethnicities in which leprosy was severe and widespread are also those in which elders tend to inculcate local trust among children. Conversely, these ethnicities are not more likely to inculcate other behaviors, and other tropical diseases do not predict these patterns. Data is at the ethnicity level, and it comes from the Standard Cross-Cultural Sample (SCCS). The SCCS contains data on whether different pathogens (including leprosy, leishmaniasis, sleeping sickness, malaria, schistosomiasis, and spirochetes) were present and widespread in the ethnicity. The variables in the row headings are constructed with this information, each of these being indicators that equal one if the corresponding disease was severe and widespread in the ethnicity. For a subset of these 186 ethnicities, the SCCS also contains information on whether elders inculcated different values in children. These variables are on a 1-10 scale, with the former corresponding to a value that is not inculcated in children and the latter corresponding to a value for which there is an extremely strong inculcation. Outcomes in columns 2, 4, and 5 are the average of the degree to which the corresponding value is inculcated in early boys, early girls, late boys, and late girls. In contrast, outcomes in columns 1 and 3 are the degree of the value's inculcation directly provided by the SCCS. Each column-panel reports results from a single OLS regression of the dependent variable described in the column heading on the variable described in the row heading and absolute latitude, pathogen stress, altitude (in meters), and year in which the information about the ethnicity was published. Heteroskedastic-robust standard errors in parenthesis.

Appendix A.27 - Behavioral Legacy of Social Exclusion in Other Contexts - Evidence from Discriminated Migrants in European Social Survey

	Care for Others' Wellbeing		Loyal to close Contacts		Have a Good Time		Follow Rules		Be Rich and Have Money	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Discriminated Against	0.042 (0.014)	0.041 (0.014)	0.034 (0.010)	0.027 (0.010)	0.011 (0.009)	-0.005 (0.008)	0.006 (0.007)	0.010 (0.006)	0.008 (0.004)	-0.000 (0.004)
R-squared	0.042	0.080	0.057	0.094	0.037	0.077	0.036	0.066	0.024	0.059
Observations	20928	20928	20928	20928	20928	20928	20928	20928	20928	20928
Mean Dep. Var	0.285	0.285	0.387	0.387	0.179	0.179	0.106	0.106	0.047	0.047
Country FE	✓		✓		✓		✓		✓	
ESS Round FE		✓		✓		✓		✓		✓
Region FE		✓		✓		✓		✓		✓
Individual Cont.		✓		✓		✓		✓		✓

Note: This table shows that, among second-generation migrants in Europe, those who report being members of a group that is discriminated against in their country of birth tend to be more concerned with the well-being of others and also more loyal to close people. Conversely, they exhibit beliefs in other dimensions that are not significantly different from those of second-generation migrants who are not discriminated against. Data comes from the first 10 rounds of the European Social Survey, and the observations are at the level of the individual. The analysis is restricted to second-generation migrants (i.e., respondents who report that their parents were born in a different country than theirs), and it leverages a variable collected in rounds 5-10 of the ESS, which contains information on whether the respondent reports being from a group discriminated against in their country of birth. Discriminated against is thus an indicator that equals 1 if the respondent reports belonging to a group discriminated against. The ESS also contains information on the extent to which the participant believes it important to behave in different ways. Outcomes are defined with this information, with each outcome being an indicator variable that equals 1 if the respondent believes it is very important to behave in the way described in each column heading. Each column reports results from a single OLS regression of the dependent variable described in the column heading on the variables described in the row headings. Country FE and Region FE correspond to country-of-response and (sub-national) region-of-response fixed effects, respectively. Individual Cont is a host of controls of respondent's characteristics, including gender, marital status, years of education, age, and the number of children ever fathered/ given birth. Robust standard errors clustered at the (sub-national) region level are in parentheses.

Appendix A.28 - Check of Randomization Implemented in Second Round of Fieldwork

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Ancestry	Age	Female	Marital	Children	Primary	Secondary
ExclusionHist	0.125 (0.074)	-2.977 (2.571)	0.027 (0.074)	0.005 (0.074)	-0.218 (0.255)	0.063 (0.053)	0.093 (0.073)
MedicalHist	-0.026 (0.074)	-4.225 (2.543)	-0.005 (0.074)	-0.090 (0.073)	-0.085 (0.264)	0.034 (0.055)	0.079 (0.073)
FloraHist	0.093 (0.077)	-2.561 (2.588)	0.036 (0.077)	0.123 (0.076)	0.332 (0.273)	-0.039 (0.062)	0.055 (0.076)
Observations	360	360	360	360	360	360	360
R-squared	0.016	0.008	0.001	0.022	0.013	0.010	0.005
Mean Dep. Var	0.508	46.689	0.508	0.500	1.986	0.833	0.594
Round 1							
Round 2	✓	✓	✓	✓	✓	✓	✓
Round 3							

Note: The table presents balance checks on the main demographic variables of participants in the second round of fieldwork, showing evidence that suggests that the randomization of participants across different treatments was successful. The unit of observation is the individual, and the data comes from the information collected in the second round of fieldwork, when the randomization exercise was implemented. Each column reports results from a single OLS regression of the dependent variable described in the column heading on the variables described in the row headings. ExclusionHist, MedicalHist, and FloraHist are indicators that equal 1 if the participant was exposed to a narrative about the exclusion endured by those with leprosy, the misinformation that characterized the actions of doctors, or the history of a local tree, respectively. The omitted category corresponds to the set of participants who were exposed to no narrative (pure control). See Appendix A.9 for a detailed definition of treatments and outcome variables, as well as Appendix A.34 for details on how the variables were collected. Heteroskedastic-robust standard errors in parentheses.

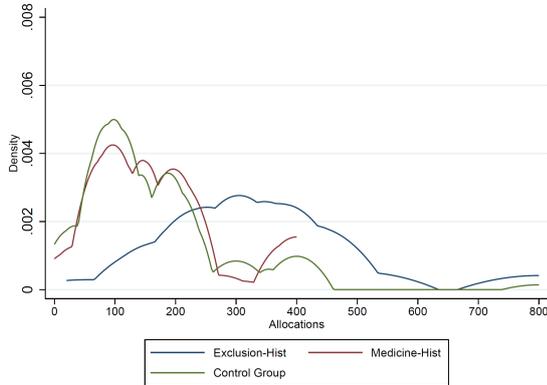
Appendix A.29 - Irrelevance of Falsification Treatment in Second Round of Fieldwork

	In-group Allocation	Out-group Allocation	Total Allocation	In-group- Out-group Allocation	Trust in Physicians	Anti- Parasitic Take-Up
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A - Pure Control and Placebo Narrative - Full Sample						
FloraHist	-4.520 (16.730)	1.764 (12.832)	-2.756 (26.576)	-6.284 (13.523)	-0.276 (0.392)	-0.031 (0.074)
Mean Dep. Var	138.218	68.563	206.782	69.655	6.839	0.362
Observations	174	174	174	174	174	174
R-squared	0.000	0.000	0.000	0.001	0.003	0.001
Panel B - Pure Control and Placebo Narrative - Only Those With Ancestry						
FloraHist	-20.264 (27.131)	-1.468 (19.875)	-21.731 (44.448)	-18.796 (16.927)	-0.317 (0.591)	-0.036 (0.098)
Mean Dep Var	160.568	62.273	222.841	98.295	6.466	0.284
Observations	88	88	88	88	88	88
R-squared	0.007	0.000	0.003	0.015	0.003	0.002
Panel C - Pure Control and Placebo Narrative - Only Those With No Ancestry						
FloraHist	3.222 (19.473)	7.644 (16.340)	10.867 (30.087)	-4.422 (19.678)	-0.093 (0.519)	0.004 (0.111)
Mean Dep Var	115.349	75.000	190.349	40.349	7.221	0.442
Observations	86	86	86	86	86	86
R-squared	0.000	0.002	0.001	0.001	0.000	0.000
Round 1						
Round 2	✓	✓	✓	✓	✓	✓
Round 3						

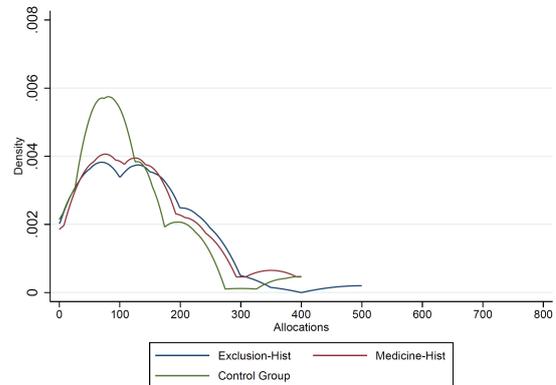
Note: The table examines data from the second round of fieldwork to show that the falsification treatment did not affect participants' behaviors. The data used to build this table is thus restricted to participants in the second round of fieldwork who were exposed to either no intervention or who were exposed to the falsification treatment. Each panel-column reports results from a single OLS regression of the dependent variable described in the column heading on the variables described in the row heading, restricting the observations per the description of the panel title. The placebo treatment, FloraHist, is an indicator variable that equals 1 for those who were exposed to a narrative about a tree that is native to the region of study. Panel A uses data from all participants who were assigned to the pure control or the placebo treatment group. Panel B focuses on participants assigned to these groups who reported having an ancestor who was excluded in the leper colony. Panel C focuses on participants assigned to these groups who reported not having an ancestor excluded in the leper colony. See Appendix A.9 for a detailed description of the outcome variables, as well as Appendix A.34 for further details on how the variables were collected. Heteroskedastic-robust standard errors are in parentheses.

Appendix A.30 - Distribution of Outcomes by Treatment and Ancestry Status

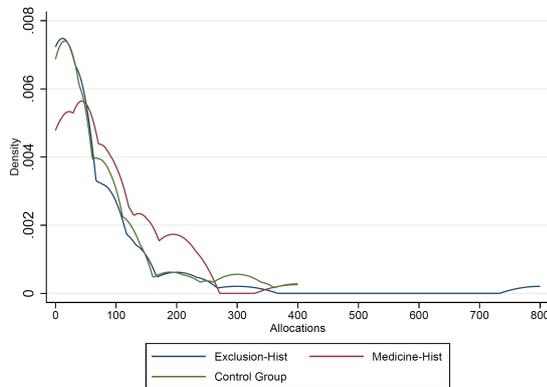
Ancestry - In-group allocation



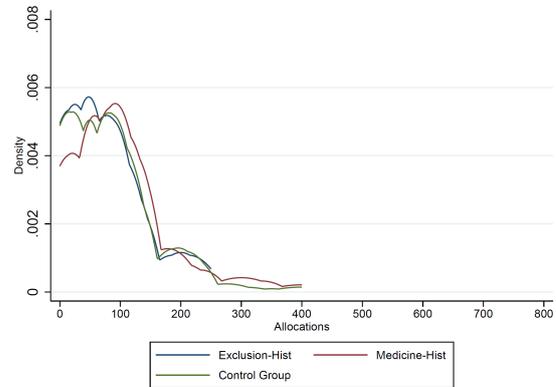
No Ancestry - In-group allocation



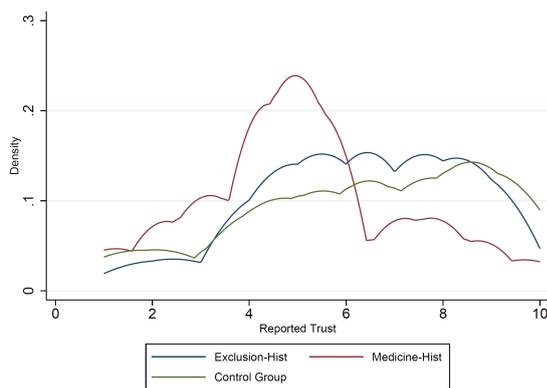
Ancestry - Out-group allocation



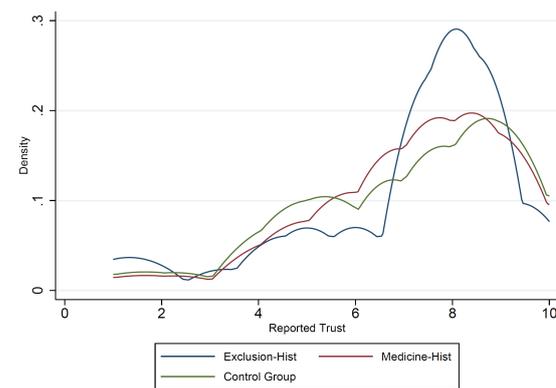
No Ancestry - Out-group allocation



Ancestry - Trust in Doctors



No Ancestry - Trust in Doctors



Note: The figures display the distributions of the allocations given by participants in the second round of fieldwork to in-group and out-group recipients, as well as the reported trust in physicians, split by ancestral status and treatment status. Ancestry refers to those who are descendants of people who lived in the colony during the epoch of isolation, whereas No Ancestry refers to participants who are descendants of people who did not live in the colony during the epoch of isolation. In-group refers to anonymous receivers from the same municipality as the respondent, whereas Out-group refers to an anonymous receiver in Pandi, a municipality outside the region of study (see Appendix Table A.5 for more information on this outgroup). Each participant got to allocate up to COP 800,000 between themselves, an anonymous in-group recipient, and an anonymous outgroup recipient (i.e., approximately \$250 at the time). Trust in physicians is on a 1-10 scale. Exclusion-Hist and Medical-Hist correspond to participants who were exposed to a narrative about the mistreatment of those with leprosy or about the medical misconceptions related to leprosy, respectively. Control Group corresponds to participants who were exposed to either no treatment or to a placebo narrative about the history of a local tree. See Appendix Sections A.9 and A.10 for descriptions of the variables, as well as Appendix Section A.34 for further details about how they were collected.

Appendix A.31 - Effect of Interventions of Second Round of Fieldwork on Attitudes Towards Lawyers and Engineers

	Trust in Lawyers		Trust in Engineers	
	(1)	(2)	(3)	(4)
Ancestry	0.145 (0.517)	-0.176 (0.523)	-0.566 (0.563)	-0.877 (0.549)
ExclusionHist*Ancestry	-0.270 (0.743)	0.030 (0.731)	0.319 (0.762)	0.634 (0.738)
MedicalHist*Ancestry	-0.063 (0.781)	0.253 (0.757)	-0.964 (0.740)	-0.720 (0.712)
ExclusionHist	0.601 (0.564)	0.288 (0.551)	0.155 (0.605)	-0.166 (0.572)
MedicalHist	0.378 (0.523)	0.032 (0.520)	0.534 (0.547)	0.267 (0.532)
Observations	360	360	360	360
R-squared	0.015	0.092	0.054	0.145
Mean Dep. Var	4.544	4.544	5.908	5.908
Baseline Covariates		✓		✓
Round 1				
Round 2	✓	✓	✓	✓
Round 3				

Note: The table examines the effects of reminding participants of different types of narratives on perceptions about lawyers and engineers. Data comes from the second round of fieldwork, and the unit of observation is an individual. Each column reports results from a single OLS regression of the dependent variable described in the column heading on the variables described in the row headings. Odd (even) columns don't control (do control) for the baseline covariates, which include gender, age, age squared, marital status, educational attainment, and the number of children of each study participant in the second round of fieldwork. All columns account for the placebo treatment (FloraHist), as well as its interaction with Ancestry. ExclusionHist is an indicator that equals 1 if the participant was exposed to a narrative about the stigma suffered by those with leprosy. MedicalHist is an indicator that equals 1 if the participant was exposed to a narrative about the historical medical misinformation about leprosy. Ancestry is an indicator that equals 1 for participants who reported that they have an ancestor who lived in the colony during the epoch of isolation. Outcomes in columns 1 and 2 correspond to reported trust in lawyers, whereas outcomes in columns 3 and 4 correspond to reported trust in engineers, both on a 1-10 scale. See Appendix A.9 and A.10 for definitions and summary statistics of these outcomes and Appendix A.34 for further information on how these variables were collected. Heteroskedastic-robust standard errors in parentheses.

Appendix A.32 - Other Historical Contents: Isolation, Displacement, Family Separation, and Religion

	Reported Knowledge of Historical Content (1=Reports/Exhibits)				Names (1=Has a name that is ...)	
	(1)	(2)	(3)	(4)	(5)	(6)
	Isolation of The III	Family Separation	Forced Displacement	Role of Religious Orders	Religious	Common
Panel A - No Covariates						
Ancestry	-0.031 (0.054)	-0.020 (0.039)	0.031 (0.026)	0.039 (0.041)	0.009 (0.060)	-0.022 (0.058)
Observations	275	275	275	275	275	275
R-squared	0.001	0.001	0.007	0.004	0.000	0.001
Panel B - With Baseline Covariates						
Ancestry	-0.031 (0.057)	-0.032 (0.040)	0.037 (0.026)	0.036 (0.042)	0.030 (0.062)	0.002 (0.061)
Covariates						
Female	-0.051 (0.055)	0.118 (0.036)	0.010 (0.019)	0.021 (0.039)	-0.011 (0.059)	-0.051 (0.057)
Age	-0.006 (0.009)	0.000 (0.007)	-0.001 (0.003)	0.003 (0.007)	0.005 (0.010)	0.011 (0.009)
Marital	-0.019 (0.061)	0.030 (0.043)	-0.009 (0.026)	-0.001 (0.044)	-0.070 (0.065)	0.005 (0.063)
Primary	-0.046 (0.100)	0.093 (0.072)	-0.041 (0.050)	-0.047 (0.071)	-0.183 (0.113)	-0.108 (0.114)
Secondary	0.053 (0.060)	-0.043 (0.047)	-0.015 (0.024)	0.041 (0.040)	-0.001 (0.065)	-0.063 (0.065)
Age Squared	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Children	0.089 (0.076)	0.015 (0.050)	0.028 (0.021)	-0.007 (0.056)	-0.019 (0.083)	-0.090 (0.080)
Observations	275	275	275	275	275	275
R-squared	0.015	0.043	0.020	0.011	0.052	0.055
Mean Dep Var	0.244	0.109	0.033	0.102	0.324	0.291
Round 1						
Round 2						
Round 3	✓	✓	✓	✓	✓	✓

Note: The table examines data from the third round of fieldwork to assess the correlations between having an excluded ancestor and different views and attitudes as captured by participants' names as well as the historical narratives they shared. Each panel-column reports results from a single OLS regression of the dependent variable described in the column heading on the variable(s) described in the row heading(s). Ancestry is an indicator that equals 1 for participants who reported that they have an ancestor who lived in the colony during the epoch of isolation. Outcomes in Columns 1-4 are indicator variables that equal 1 if, when asked about doctors' role in the historical isolation of patients, the participant mentioned the theme described in the respective column heading. Outcomes in Columns 5-6 are indicator variables if the participant has a name that is common in Colombia or a name that is religious, respectively. See Appendix A.9 and A.10 for definitions and summary statistics of these outcomes and Appendix A.34 for details about the third round of fieldwork. Heteroskedastic-robust standard errors in parentheses.

Appendix A.33 - Robustness of Sentiment Analysis

In the third round of fieldwork, participants were asked to report what they knew about the historical role of physicians in the leper colony (see Appendix Section A.34 for more details of the third round of fieldwork). The responses were open-ended, allowing several pieces of information to be extracted from such answers (see Table 6 for specific outcomes built upon these responses, and Appendix Sections A.9 and A.10 for a detailed description of these variables as well as their summary statistics). Crucially, these open-ended responses allowed for examining whether participants tended to provide responses characterized by a negative sentiment or not. This outcome was achieved using tools of sentiment analysis provided by the command R package “sentimentr,” a package developed by Tyler Rinker (Rinker, 2021) that has been used in settings not unlike that of this project, such as in studies of the sentiment of critically ill patients (Weissman et al., 2019).

The package is based on a dictionary of 11,710 words, each of which has a corresponding sentiment score in the $[-2,1]$ interval (words that convey a negative sentiment have negative scores, while words that convey a positive sentiment have positive scores). The overall score, which is the sum of these word-level scores, captures whether the sentence conveys an overall negative sentiment. The package also has 140 valence shifters, which amplify or attenuate the score of each word, thereby allowing an alternative sentiment score at the sentence level that seeks to capture the sentiment as originally intended. The construction of the “negative sentiment” outcome assessed in the baseline results (column 4, Table 6) is derived from the following steps:

1. Translating responses from Spanish to English: The original responses were first recorded by enumerators and later transcribed. These responses were in Spanish, and they use idioms and expressions that convey a meaning in Colombia but may not necessarily mean the same in other contexts. The first step was thus translating these responses to English, ensuring that the translation captured, to the extent possible, the meaning as intended by each participant. The original data contain both the literal responses provided by each participant (in Spanish) as well as the English translation, which is the main input for the sentiment analysis.
2. Word filtering in R: Responses had to be filtered in some simple (intuitive) ways to make sure that the sentiment analysis actually represented the sentiment displayed by participants. Hence, the following elements and words contained in participants’ responses were filtered (dropped) before conducting the sentiment analysis:

- (a) Characters that do not represent any concrete meaning, such as punctuation marks, numbers, alphanumeric symbols, and digits, which were transcription errors by enumerators.
 - (b) Words that are frequent but provide little information, such as pronouns and connectors (also known as “stop words”).
 - (c) Spuriously negative or positive words: given the context and the question asked (regarding leprosy, a disease that tends to have a negative connotation), participants naturally included in their responses words that R tends to interpret as heavily negative even if that was not the intention of respondents. To avoid this innate bias in each participant’s sentiment score, I dropped the following (medical, leprosy-related) words before conducting the sentiment analysis: injury, injuries, surgery, sanatorium, amputate, die, dies, died, deadly, virus, death, hansen, hansens, hansenn, hanssen, leprosy, leper, lepers, sick, sickness, sickest, healthy, health, ill, disease, doctor, doctors, shelter, shelters, patient, patients, medicine, medicines, medication, contagious, disinfected, disinfect, lazaretto, lazaret, lazarettos, infection, gave, cure, cured, cures, hospital, hospitals, clinic, clinics, hospitalize, medical, drug, drugs, contagion, sighs, nurse, nurses, painkillers.
3. Computing a sentiment score per response: Based on the dictionary, and once a set of scores was established at the word level, the next step was to collapse these values to obtain opinion/observation scores. The procedure was simply to add the score of words per observation, where each word has the same weight in the estimation. The main outcome was thus computed as an indicator that equals one when the overall score of response is negative and zero otherwise.

Robustness Exercise

The table below provides several robustness checks that suggest the results are not driven by the specific dictionary chosen or by the fact that the sentiment was computed at the word level instead of at the sentence level. In the table, Panels A (without baseline covariates) and B (with baseline covariates) compute the scores after filtering out medical words with an innate negative connotation, while Panels C (without baseline covariates) and D (with covariates) compute the scores without filtering out these medical words. Each column of the table examines the association between having an excluded ancestor and whether the historical narrative shared by the participant conveyed a negative sentiment according to different R packages designed for those purposes:

1. Column 1 presents the original sentiment analysis examined in the main text and described above, which is based on the package “sentiment” (Rinker, 2021).
2. Column 2 presents the results when the sentiment is computed with the same package (sentimentr). However, instead of computing and averaging “word sentiments,” this score computes a sentence-level sentiment, which uses valence shifters to amplify or attenuate the score of different words.
3. Column 3 presents the same word-level exercise as column 1 but using the “syuzhet” package in R instead of the original “sentimentr” package. This alternative package is based on a dictionary developed by the Nebraska Literary Lab, which contains 10,748 words, each with a score in the interval $[-1,1]$. The variable is constructed as an indicator that equals 1 if the sum across the score of words used by a participant is below 0.
4. Column 4 presents the sentence-level sentiment using the syuzhet package and lexicon. Like the outcome in column 2, this sentiment score retrieves the sentiment computed by the package for an entire sentence based on valence shifters (instead of averaging across words).
5. Column 5 presents the same word-level exercise as column 1, using the “Afinn” R package instead of sentimentr. This package leverages a dictionary of words developed by Finn Arup Nielsen, which includes internet slang and obscene words. It contains 2,477 words, each with a score in the interval $[-5,5]$. The variable is constructed as an indicator that equals 1 if the sum across the score of words used by a participant is below 0.
6. Column 6 presents the sentence-level sentiment using the “Afinn” package. Like the outcomes in columns 2 and 4, this sentiment score retrieves the sentiment computed by the package for an entire sentence based on valence shifters (instead of averaging across words).

	(1) Negative Sentiment (Used in Table 6)	(2) SentimentR Negative Sentence Sentiment	(3) Syuzhet Negative Word Sentiment	(4) Syuzhet Negative Sentence Sentiment	(5) Afinn Negative Word Sentiment	(6) Afinn Negative Sentence Sentiment
--	--	--	--	---	--	---

Panel A - Dropping Biased Words - Without Baseline Covariates

Ancestry	0.164 (0.065)	0.169 (0.065)	0.170 (0.065)	0.170 (0.065)	0.114 (0.066)	0.120 (0.066)
Mean Dep Var	0.566	0.574	0.562	0.562	0.541	0.537
R-squared	0.025	0.027	0.027	0.027	0.012	0.013

Panel B - Dropping Biased Words Words – With Baseline Covariates

Ancestry	0.150 (0.067)	0.156 (0.066)	0.152 (0.067)	0.146 (0.067)	0.074 (0.067)	0.083 (0.067)
Mean Dep Var	0.566	0.574	0.562	0.562	0.541	0.537
R-squared	0.074	0.085	0.080	0.096	0.093	0.095

Panel C - Including Biased Words – Without Baseline Covariates

Ancestry	0.107 (0.057)	0.107 (0.057)	0.101 (0.057)	0.096 (0.058)	0.065 (0.061)	0.050 (0.063)
Mean Dep Var	0.727	0.727	0.731	0.723	0.686	0.661
R-squared	0.013	0.013	0.012	0.011	0.005	0.003

Panel D - Including Biased Words - With Baseline Covariates

Ancestry	0.104 (0.060)	0.106 (0.060)	0.096 (0.060)	0.093 (0.059)	0.044 (0.064)	0.036 (0.065)
Mean Dep Var	0.727	0.727	0.731	0.723	0.686	0.661
R-squared	0.061	0.061	0.059	0.094	0.040	0.046
Round 1						
Round 2						
Round 3	✓	✓	✓	✓	✓	✓
Observations	242	242	242	242	242	242

Note: The table presents robustness checks of the sentiment analysis presented in Table 6. The unit of observation is the individual, and the data comes from the information collected in the third round of fieldwork. Each panel-column reports results from a single OLS regression of the dependent variable described in the column heading on the variables described in the row headings, subject to the controls and details described in the panel heading. Baseline covariates (Panels B and D) include gender, age, age squared, number of children, marital status, and indicators for primary and secondary education. Ancestry is an indicator that equals one for participants who reported that they have an ancestor who lived in the colony during the epoch of isolation. Details of the construction of the variables in the column heading are discussed in the text of Appendix Section A.33. Heteroskedastic-robust standard errors are in parentheses.

Appendix A.34 - Details of Recruitment and Protocols in Each Round of Fieldwork

This Appendix Section provides further details of the protocols that I implemented in each of the three rounds of fieldwork.

First Round of Fieldwork

This protocol was registered in EGAP under the ID # 20180724AB; the details that follow are consistent with what was registered. The experimental sessions were conducted on Saturdays in June, July, and August of 2018. Invitations were distributed Wednesdays and Thursdays before each session. A total of 12 sessions were implemented, with sessions 1–6 occurring in Agua de Dios, sessions 7–9 in Tocaima, and sessions 10–12 in Ricaurte. Recruitment for each session was conducted following a random walk method: with the help of local leaders, the neighborhoods within each municipality were identified along with the central locations of those neighborhoods. Starting from such nodes, the enumerators arbitrarily selected one path and invited no more than one member per household they approached. In the hope of achieving 25 subjects per session, 33 participants were invited for each of the sessions in Agua de Dios (a total of 200 invitations), while 32 were invited per session in the contiguous regions (a total of 192 invitations). The total number of participants was 265, with 138 choosing to participate in Agua de Dios and 127 in its neighboring municipalities. All sessions were conducted in the school closest to the center of each municipality. The invitations contained the school's address and the informal name with which people associate the school of each municipality. These invitations also provided the specific hour at which each participant had to arrive; enumerators made it clear that late arrivals would not be allowed in. One member of the research team was standing at the school entrance, verifying the authenticity of the invitation as a prerequisite for entering. Once inside, subjects gathered on a patio where they waited before being called, one by one, into classrooms specially arranged for the interview and the experimental decision.

Enumerators instructed participants to enter a room where data were collected. Once the participants entered their assigned room, they would encounter two chairs placed in the middle of the room, each one facing the other at a distance of approximately four meters. The enumerator was in one of these chairs, and the participant was asked to sit in the chair in front. After the participant sat down, enumerators began by collecting different outcomes. The interview began with some basic demographic questions, followed by experimental decisions and, subsequently, some survey questions

(see Appendix Sections A.9 and A.10 for a detailed description of these variables and their summary statistics).

For the experimental decisions, the enumerator went over the details of what the participant was expected to do. The enumerator then handed eight bills of COP 2,000 to the participant, along with an empty white envelope, and emphasized that this money was theirs and they were free to allocate whatever amount they chose for the receiver (including 0). The enumerator then instructed the participant to walk toward an area in the back of the room, behind a wide and large cardboard, which resembled a voting booth. Participants were told that they had to make their choices for the game in this space. Behind this cardboard, the participant would encounter a voting box with a small orifice and a small window showing the closed white envelopes inserted by other participants. Given the window's size and the envelopes' homogeneous color and (large) size, participants could not possibly determine the amounts allocated by other participants, nor could they alter the choices of other participants. Once in this private area, the participant would place the chosen amount for their recipient in the envelope, close the envelope, and deposit the envelope with the allocation into the box. The box (and the cardboard behind which it was located) were behind the interviewer's chair (who would remain seated while the participant made their choice), which afforded another layer of privacy to the decision maker. The participant could then keep in their pockets (or wherever they decided) whatever amount they chose for themselves.

Each participant went through this process twice—one after the other—with a specific area in the room and a specific box corresponding to in-group recipients (anonymous receivers from the same municipality as the participant) and another area of the room (and thus another box) corresponding to out-group recipients (anonymous receivers from Apulo). As discussed in Section 4, Apulo was selected as an out-group given its historical similarity to the region of my study, plus the fact that it is not contiguous to the former leper colony. The allocations selected by each participant (both in- and out-group allocations) were delivered in the respective municipality. Hence, there was no deception, and allocations were honored.

After completing the two experimental decisions, participants returned to their seats and answered the remaining survey questions. The interview ended with the interviewer asking the participant to avoid sharing the session's content and their answers with anyone. The session ended when all participants completed their respective decisions and interviews. The envelopes given to each participant were marked inside with a number (both envelopes used by each participant were marked

with the same number), which coincided with the number of the questionnaire that the enumerator filled out when recording participants' responses. This allowed the research team to match the allocations chosen by each participant with their demographic characteristics and survey responses without needing to collect any personal identifying information. At the end of all sessions, we had collected no personal information (no IDs, first or last names) but simply 265 completed (numbered) questionnaires, 265 closed and numbered envelopes with the in-group allocations, and 265 closed and numbered envelopes with the out-group allocations. Hence, there was no way for the researchers to know what specific participants chose to allocate, and both participants and the research team were aware that recontacting participants would be impossible.

Second Round of Fieldwork

The second protocol was registered in EGAP under the ID # 20190527AA, and the details that follow are consistent with what was registered. This protocol, designed to assess mechanisms, was implemented on Saturday, June 1, 2019, with the goal of collecting data from a different sample than that of the first round of fieldwork. Six enumerators were sent to the central nodes of neighborhoods in the municipality of Agua de Dios. Starting from these nodes, each enumerator started walking along a path they arbitrarily chose. The enumerators were instructed to approach every other house they encountered along such a route and tell whoever answered the door that answering a couple of questions for five to ten minutes would allow them to enter a lottery for COP 800,000 (approximately \$250 at the time) and a free dose of antiparasitics. The enumerators were instructed to interview, in private, one subject in the household before moving on to the following household.

The enumerator recorded participants' responses in a numbered questionnaire, and there was a different questionnaire (with a different number) for each participant. The enumerator began by administering the treatment to the participant, which amounted to reading out loud one of the following texts:

- No Treatment: The enumerator collected responses without priming.
- Exclusion-Hist: The enumerator read a brief excerpt commenting on the social stigma to which leprosy patients have been subjected historically and globally. The exact excerpt read as follows: "Since antiquity and across continents and nations, people suffering from leprosy have been subject to social stigma and exclusion. With the help of state authorities and civilians, those suffering from the disease have historically faced strict physical isolation and marginalization."

- Medical-Hist: The enumerator read a brief excerpt commenting on the historical medical misinformation relative to leprosy. The exact excerpt read as follows: “For centuries, civilians and experts alike believed that leprosy was a highly contagious and deadly disease. Such a belief was ratified by international summits like the Berlin Congress of Leprosy, where it was concluded that the disease was indeed as dangerous as it was commonly believed. Nowadays, it is known that the disease is, in fact, neither deadly nor highly contagious.”
- Flora-Hist: The enumerator read a brief excerpt commenting on the history of the region’s trees and geology. The exact excerpt read as follows: “The Chicala Tree is a floral species that is originally from the southern regions of North America. It was brought into Colombia and specifically to Cundinamarca by indigenous communities due to its resistance to fire and its potential decorative use. The climatic conditions in the Tequendama Region in Cundinamarca have allowed the Chicala to flourish.”

The enumerators collected participants’ information in a form that had a unique number for each participant. This number determined the type of treatment given to each participant. Those interviewed with forms numbered 1–99 received treatment 1 (Exclusion-Hist); those interviewed with forms 100–198 received treatment 2 (Medical-Hist); those with 199–297 received treatment 3 (Flora-Hist); and those interviewed with forms 298–396 received no treatment (No-Treat).

After reading the corresponding excerpt to each participant, the enumerator proceeded to inquire about allocations for the dictator game. In this modified version of the dictator game, enumerators reminded participants that they would earn COP 800,000 if they won the lottery and asked how much they would be willing to donate from this prize to someone chosen arbitrarily and anonymously in Agua de Dios (in-group allocation) and someone chosen arbitrarily and anonymously from Pandi, Cundinamarca (out-group allocation).

As mentioned in Section 4, Pandi was selected with the explicit goal of providing an outgroup that was also similar to the region of study but was more distant than Apulo. There are other distant municipalities that are similar in socioeconomic characteristics to the region of study, but these alternative municipalities belong to distant sub-national regions in Colombia and were thus likely unknown by study participants. To avoid participants’ confusion and prevent potential issues when conducting the second round of fieldwork, I thus decided to choose Pandi as an outgroup for this second round of fieldwork.

When enquiring about allocations in this round, enumerators reminded participants that these allocations would be enforced if they won the lottery. To ensure that participants trusted the lottery's outcome and randomness, they were given a number (between 001 and 399) and were told that the winning number would be the last three digits of the "Lotería de Bogotá" of June 6, 2019, a popular lottery that plays every Thursday and is well known in the region (the winning number turned out to be 466, so no allocations were enforced as no prize was given, per the rules explained to participants).

Participants were also told that if they won the lottery, the research team would contact them (their phone numbers and their addresses were recorded). Moreover, they were provided with a local number they could contact if they wanted to dispute or discuss the lottery's outcome. After making the choices corresponding to the modified dictator game, participants gave their opinion on their degree of trust (on a 1–10 scale) in physicians, lawyers, and engineers (see Appendix Sections A.9 and A.10 for a detailed description of these variables and their summary statistics). The last question that each participant answered pertained to their ancestry (specifically whether any of their parents or grandparents experienced the exclusion in Agua de Dios). Participants were told that further proof of their reported ancestry could be required before giving the prize to the winner.

Next, the enumerator provided a voucher to the participant and told them that they could redeem it (for free) in exchange for a full dose of antiparasitics (a dose of albendazole and secnidazole). The voucher, participants were told, could be redeemed at the local and trusted pharmacy (Drogueria Popular, located in the center of the municipality and run and owned by Saul Morales, a native of Agua de Dios). The voucher had written the final date on which it could be redeemed (June 5, giving participants four days to collect it). After giving the voucher to the participant, the interview ended and the enumerator would move on to the next household/interviewee.

The ethical dimension of the randomization exercise was carefully considered before, during, and after implementation, beyond IRB approval. As mentioned, treatments were designed with the inputs provided by a local oral historian so that they would mention narratives that participants with an excluded ancestry had likely heard.¹ The contribution of the local oral historian to this treatment also reduced the likelihood that the message's tone and language would upset participants. Moreover, the enumerator privately delivered the treatment to the participant (i.e., at the entrance of their home), which prevented any potential social pressure that could produce anxiety or fear of retaliation from third-party observers. The fact that the dose of antiparasitics was provided to all

¹I thank Efraín Oyaga for his invaluable help, which allowed the randomization exercise to happen.

participants further ensured equal access to the experiment's benefits.

Last, one of the agreements with the local oral historian of the leper colony was the socialization and discussion of the projects' findings after its completion, with the explicit goal of debriefing the results and any unexpected consequences. Honoring such an agreement, I shared the results with the oral historian (Efrain Oyaga), explaining the project's findings and providing printed copies of the manuscript. Moreover, a columnist from Colombia's second-largest newspaper, *ElEspectador*, wrote an op-ed discussing the paper's conclusions vis-à-vis the colony's history, highlighting the responsibility of the state and civil society beyond the errors that physicians made in the past.² Crucially, Agua de Dios's local newspaper, *LaTribuna*, reprinted the op-ed on its front page, ensuring the project's main findings and its nuanced discussion would be distributed.

Third Round of Fieldwork

The third protocol was registered in EGAP under the ID # 20210817AA. Registration occurred before I had access to the data, and the details that follow are consistent with what was registered. The third round of fieldwork was designed and implemented with the intention of surveying a new set of participants and examining additional questions that remained unclear or unaddressed in the first two rounds of fieldwork. Specifically, this round was intended to collect survey data about pro-sociality and altruism toward different actors and to dig into the content and sources of what individuals claim they know about the region's history.

This round of fieldwork was planned and implemented during the worst outbreak of the COVID pandemic in Colombia, which coincided with generalized social unrest in the country (between April and mid-June 2021). These constraints limited my ability to collect all the outcomes that I intended to and left few alternatives to recruiting participants other than outdoors in public spaces (given the circumstances, however, this setting mitigated a hostile response from people fearful of strangers knocking on their doors). Given these constraints, I hired and worked alongside a professional enumerator (Rodolfo Méndez Torres, rodolfo.mendtorr@gmail.com) to collect data in Agua de Dios and Tocaima in public spaces. This professional enumerator hired and trained three other enumerators. The enumerators then visited the two municipalities (one at a time) between June 16 and 25, 2021. Following the process of rounds one and two of fieldwork, the enumerators stood in the central nodes of each neighborhood in Agua de Dios and Tocaima. They simply remained in the vicinity of these (outdoors, public) central nodes, walking close to the nodes but without following a specific

²See <https://www.elespectador.com/opinion/columnistas/marc-hofstetter/el-puente-de-los-suspiros-column-900407/>

path. They thus approached and surveyed anyone who would be open to answering the survey. As discussed in Section 4, two of the main outcomes collected in this round of fieldwork were: (a) open-ended responses of what each participant knew about the historical role of doctors in the leper colony and (b) open-ended responses of all the sources that informed the historical views shared by each participant. These responses were recorded by the enumerators, and the recordings were then transcribed by the professional enumerator that I initially hired. The professional enumerator then shared with me these transcriptions (along with all other outcomes collected) in an excel file. The original file contained the first and last names of each respondent, as well as their neighborhood of residence. To ensure that no personal identifying observation was shared with the general public, both the neighborhood and the last names of respondents were dropped before sharing the raw data in the repository.

All enumerators followed strict protocols to avoid contagion (using face masks constantly and properly, maintaining at least two meters between themselves and the surveyed subject, and frequently disinfecting/washing their hands). A total of 275 individuals were surveyed (see Appendix Sections A.9 and A.10 for a description of variables and summary statistics). No experimental data were collected, given the constraints.

Auxiliary Survey Data in Distant Places (Appendix A.19)

In order to address concerns that the baseline results could be driven by the control group (i.e., descendants of non-excluded locals), I collected data in municipalities with climatic conditions akin to those in Agua de Dios as well as in an economically-depressed neighborhood in Bogotá. Specifically, I computed the climatic dissimilarity indices displayed in Appendix A.8, which indicate that the municipality of Mariquita is the Colombian municipality outside the region of study that holds the most similar climatic conditions to Agua de Dios. Moreover, such indices indicate that Puerto Salgar is the most climatically similar to Agua de Dios out of all municipalities in Cundinamarca (the subnational region where the study is located) and the third most similar out of all the municipalities in Colombia (besides the municipalities contiguous to Agua de Dios). I then hired the same enumerator that collected the data in the third round of fieldwork (Rodolfo Méndez Torres, rodolfo.mendtorr@gmail.com) and asked him to collect data in such municipalities and in Kennedy, an economically-depressed neighborhood in Bogotá. After identifying the localities and neighborhoods in each of these municipalities, the professional enumerator hired two additional enumerators and collected survey data in open spaces from each of these places. This data-collection process took

place in August 2019.

Enumerators approached subjects and asked them if they wanted to participate in a brief non-political survey that would only take a couple of minutes. After these subjects agreed to participate, enumerators collected baseline demographic characteristics (age, gender, education level, number of children, and marital status) and then asked the same survey questions about altruism and trust in medicine that were collected in the first round of fieldwork (see Appendices A.9 and A.10 for further details about these variables). Since the original variable capturing trust in the HPV vaccine had to be rescaled so that higher values corresponded to higher trust in the vaccine, on this occasion, the scale of this outcome was between 1-10, with greater values corresponding to higher trust in the vaccine (per the rescaling implemented for the baseline results). In addition to this, participants who reported not knowing about the vaccine were transcribed in the data as a 0 instead of a missing, so the code converts these responses as missing values to ensure that the comparison with the first round of fieldwork is adequate. All in all, the enumerators surveyed 66 subjects in Kennedy (Bogotá), 63 in Mariquita, and 65 in Puerto Salgar.