Migration, Culture, and Long-Run Development: Evidence from Brazil's German Colonies[*]

Gabriela Araujo[†]

Draft 30 June 2025

Abstract

This paper examines the long-term impact of early settlers on economic prosperity. Specifically, this paper investigates the long-term economic impact of 19th-century German immigration to southern Brazil, focusing on how early human capital shocks shaped development in a largely undeveloped region. Exploiting spatial variation in proximity to historical German settlements, I find that municipalities closer to these colonies exhibit significantly higher current literacy rates and income levels. The empirical strategy combines a quasi-natural experiment with robustness checks, placebo tests, and a comparative analysis with Italian settlement areas to support causal interpretation. I show that literacy is a key mediating channel, and that historical proximity to German colonies is associated with structural shifts away from agriculture and toward human capital-intensive industries. These findings highlight the enduring influence of early skilled immigration on local development.

Keywords: Human capital transmission, economic persistence, economic history.

1 Introduction

The long-term effects of historical immigration on economic development have gained increasing attention in recent economic history. In particular, understanding how the arrival of relatively skilled or educated immigrant groups shaped the developmental trajectories of frontier regions offers important insights into the origins of regional inequality. This paper investigates the long-term economic impact of 19th-century German immigration to southern Brazil, a region that, at the time of settlement, was largely undeveloped and sparsely populated.

German immigration to Brazil in the nineteenth century was relatively limited in magnitude. Estimates suggest that around 28,000 Germans migrated to Brazil between 1824 and 1872 (de Geografia e Estatística (2000)). From 1872 to 1910, Germans accounted for less than 3 percent of all immigrants to Brazil, while Italian immigrants outnumbered them by a factor of fourteen. Unfortunately, disaggregated immigration estimates for the state of Rio Grande do Sul, the southernmost state of Brazil and the focus of this study, are not available. Nonetheless, historical accounts and the concentration of settlements suggest that German immigration played an outsized role in shaping this region's early development.

German settlers arrived between 1824 and 1867, during a formative period in which the institutional and economic foundations of Brazil were still taking shape. These immigrants brought with them relatively high levels of education, strong associational norms, and agricultural and artisanal skills. Unlike other parts of Brazil, which were shaped by plantation economies and slave labor, German settlements in Rio Grande do Sul were organized as smallholder agricultural colonies. This unique historical episode provides a valuable quasi-natural experiment for studying how an early injection of human capital into an otherwise underdeveloped frontier affects long-run educational and economic outcomes.

This paper exploits spatial variation in proximity to these early German colonies to estimate their persistent effects on modern outcomes. Using municipal-level data from the 1980 Brazilian Census, I examine how distance to historical German settlements correlates with contemporary literacy rates and income per capita. The identification strategy relies on the assumption that the original settlement locations were determined primarily by geopolitical and land availability considerations, rather than by pre-existing differences in development. To support a causal interpretation, I implement a battery

of robustness checks, placebo tests, and covariate balance assessments.

The results show that municipalities closer to historical German colonies exhibit significantly higher current literacy and income levels. These patterns are robust to the inclusion of a wide range of geographic and socioeconomic controls. The evidence is consistent with long-lasting externalities from early human capital accumulation and localized institutional development, which appear to have compounded over time and contributed to persistent regional disparities. By focusing on a setting where initial conditions were close to a *tabula rasa*, this paper contributes to our understanding of how early human capital shocks can leave deep and durable imprints on economic development.

To shed light on the mechanisms underlying these effects, I explore how German settlement influenced structural change. In particular, I examine shifts in sectoral employment shares, showing that municipalities closer to German colonies experienced earlier transitions from agriculture toward manufacturing and services. Furthermore, to isolate the role of culture and religion—particularly the Protestant ethic—I compare the long-run effects of German versus Italian settlements in the same region. This comparative strategy helps disentangle general effects of European migration from those specific to Protestantism and its associated values.

This research contributes to multiple literatures. First, it adds to the literature on the long-run economic effects of migration, building on seminal studies such as Hornung (2014) and Murard and Sakalli (2020). Second, it engages with the literature on historical persistence, which examines how deep-rooted determinants continue to shape modern outcomes through channels like human capital and cultural transmission (Galor, 2005; Michalopoulos and Papaioannou, 2020; Nunn, 2009; Spolaore and Wacziarg, 2013). Third, it contributes to the literature on religion and economic performance by extending the insights of Weber (1934) and Becker and Woessmann (2009) to a New World context. This study also relates to work on the persistence of cultural traits (Grosjean, 2011; Guiso et al., 2016; Tabellini, 2008; Voigtländer and Voth, 2012).

This paper offers several novel contributions. First, it leverages a rare quasi-natural experiment to examine the long-run consequences of human capital shocks driven by migration. Second, it provides empirical evidence on the enduring developmental impact of German immigration in Brazil, a setting that has received limited attention in the economics literature. Third, it analyzes these dynamics during an early phase of state formation and economic organization.

3

Empirically, the relationship between proximity to German colonies and income appears to be mediated largely through educational gains. Consistent with structural transformation models, these municipalities also experienced a shift toward more human capital-intensive industries over time.

The remainder of the paper is structured as follows: Section ?? provides historical background on German immigration to Brazil. Section ?? describes the data. Section ?? presents the empirical model. Section ?? reports the main results. Section ?? conducts placebo tests, and Section ?? implements robustness checks. Sections ?? and ?? explore the mechanisms through which the effects of German settlement persist. Section ?? examines Italian migration to Rio Grande do Sul. Section ?? concludes.

2 Historical Background

In 1824, over 120 Germans arrived in the first contingent of settlers, which is considered the starting point of immigration and colonisation in Brazil (Luebke (1987)). At that time, the Brazilian government was encouraging Europeans to migrate to Brazil to supplying the country with an alternative workforce in the face of the declining, though still pervasive, slave labour.[1] In particular, Dom Pedro I, the founder and first ruler of the Empire of Brazil, was determined to increase the population of Brazil's southern frontier (Luebke (1987)).

To encourage migration to Brazil, the Crown granted various advantages to immigrants. These included the receipt of portions of land, the payment of tickets from Europe to Brazil as well as immediate naturalisation and religion freedom (Lando and Barros (1976)). To this end, free passage, and free land in the state of Rio Grande do Sul was promised to German immigrants (Oberacker Jr (1975)). In decision number 80 of the imperial government, dated March 31st, 1824, it becomes clear and the government's intention to create new colonies in the province of Rio Grande do Sul

> A Colony of Germans is expected shortly in this Court, which cannot cease
> to be of recognised usefulness to this Empire due to the superior advantage
> of if you employ white people, free and industrious, both in the Arts and
> in Agriculture: And with the S.M. the Emperor that the Land on which

---

[1]Moreover, there were some Brazilians with racist sentiments who welcomed the immigration as a means to increase the white proportion of the population and thereby counterbalance the huge, newly freed black population (Luebke (1987)).

4

the Estabelecimento do Linho Canhamo (later named São Leopoldo) lies in the Province of S. Pedro (later named Rio Grande do Sul), is the most appropriate to settle the Germans: Sends the same August Lord, for the Secretary of State for Foreign Affairs, that the President of the Government of that Province, proceed: 1st, to have the same Land measured, to be divided in dates of 400 fathoms; 2nd, that he immediately give part of the amount of Lands and Couples, that they could be arranged in them, since the arrival of the Colonists was very close ... (Iotti (2001)).

From the 1830s to the mid-1840s, foreign immigration to Brazil was stagnated due to the serious political disturbances that threatened the integrity of the empire. In the state of Rio Grande do Sul, the Ragamuffin War (1835-1845) also contributed to this interruption. Furthermore, from 1830 onwards, financial aid to immigrants was suppressed (Saviani Filho (2008)).

After this period of stagnation, changes in the legislation dedicated to immigration and colonisation were made in 1844 and lasted until the end of the empire in 1889 (CUNHA (2006)). In 1848 the government budget law general ordered the provinces to grant 36 square leagues of vacant lands directed exclusively to colonisation and prevented from being routed by slave arm (Klug et al. (2012)). The main objective of colonisation at this time became related to agricultural production. From 1845 official German colonies began to be founded from São Leopoldo towards the west, in regions which, once populated, enabled the opening of communication between the southern and the northern part of the province (CUNHA (2006)).

Notably, this land had been ignored by the Portuguese and was unsuited for grazing cattle and was not rich in precious minerals or gems (Luebke (1987)). German immigrants settled almost exclusively in remote pastoral and forested lands, isolated from other settlements, where they had to cultivate the land and construct their own houses. They were able to organise themselves independently, build their own churches, schools and municipal authorities (Schütz (1974)). The majority of the Germans in Rio Grande do Sul were small farmers who cultivated the soil with the help of their typically large families (Luebke (1987)).

In contrast to the religion composition of the Brazilian population, which was mostly

Catholic, more than half of German immigrants were Protestant (Luebke (1987)).[2,3] With the arrival of German immigrants to Brazil, Protestants constituted a significant share of the population in German-Brazilian communities. According to Luebke (1987) one estimate made in the early 1920s reports that 54 percent of the Germans in Rio Grande do Sul were Protestant, the remainder Catholic.

Although German Protestant immigrants did not have the primary objective of evangelising Brazilians, they helped to create the conditions that facilitated the introduction of missionary Protestantism in Brazil. Protestant missionary activity began in Brazil later in the mid-nineteenth century. Missionaries founded schools, hospitals, radio stations, printing presses, and clinics, yet made few converts until the second half of the 20th century ((Ferreira, 1959); (BELLOTTI, 2000); (Chesnut, 1997)). Between 1824 and 1870, forty or more German evangelical churches were organised (Luebke (1987)).

Like other European immigrants, German immigrants were more educated than Brazilian natives and placed a high value on literacy (Stolz et al. (2013)).[4] Given the lack of public schools near the forested areas where German immigrants settled in Rio Grande do Sul, families took turns teaching the children in the colony. Later, the educational function was shifted to the church (Luebke (1987)). Furthermore, German schools were also of better quality compared to public schools. In 1879, the director of education in Rio Grande do Sul reported that the German schools were organised, well equipped, and well attended. In contrast, government schools were of much lower quality and lacking in equipment (Neiva and Diégues Júnior (1955)). Luebke (1987) also notes that in 1920 only 24.5 per cent of the Brazilian population was classified as literate while in Rio Grande do Sul, which had a large number of German and Italian immigrants, registered the highest literacy rate of any state in the country (38.8 per cent) and the figure soared to 62 percent in the German colony of São Leopoldo.

---

[2]Non-Catholics were tolerated by the Brazilian government so long as their houses of worship did not bear the standard insignia of churches – e.g. crosses and spires (Luebke (1987)).

[3]Emilio Willems has observed that religion drew a line of cultural demarcation between Catholics and Protestants among the Germans of Brazil. Differences in their philosophical conceptions and roles of behaviour may be traced to the churches' differing hierarchies of values regarding law, government, and economic pursuits (Luebke (1987)).

[4]Using data from the National Archive of Rio de Janeiro on passenger lists of ships that arrived in Brazil from the beginning of the 19th century, Stolz et al. (2013) show that German immigrant groups were more numerate than native Brazilians. This paper uses data were recorded at the ports in Rio de Janeiro and Santos, which were the most important ports. According to Ferenczi (1929), around 85 percent of all immigrants entered Brazil via those ports.

3   Data and Descriptive Statistics

This study utilizes data from multiple sources. Primary data sources include Brazilian national censuses and geospatial datasets.

First, I employ data from the national censuses of Brazil for the years 1872, 1940, 1950, 1960, 1970, and 1980. These censuses provide demographic and economic information at the municipal level. Notably, municipal-level data for the key variables of interest are not available in censuses conducted between 1872 and 1940. Furthermore, detailed occupational data, crucial for examining the occupational structure, are only available from the 1980 census.

Second, geographical data are obtained from the NASA JPL Shuttle Radar Topography Mission Global 1 Arc Second (SRTM1) for elevation and terrain characteristics. Additionally, I utilize OpenStreetMap (OSM) data, processed in QGIS, to derive distances from municipalities to rivers and the coast. The OSM data was downloaded in 2025 from Geofabrik.

Finally, these administrative and geographical datasets are spatially matched to shapefile data on municipalities in Rio Grande do Sul, obtained from the Brazilian Institute of Geography and Statistics (IBGE).[5] The spatial data are used to construct the distance from municipality capitals to the city of São Leopoldo, a key variable in this analysis (see Figure 1).

Descriptive statistics for key variables in 1980 are presented in Table 1. Distances to the 1868 German colonies and to the state capital, Porto Alegre, are calculated from the centroid of each municipality. Notably, Table 1 reveals a negative correlation between distance to the 1868 German colonies and literacy rates, suggesting that municipalities in closer proximity to these settlements exhibit higher levels of literacy.

---

[5]These data are available at https://geodata.mit.edu/.

Figure 1: Location of German Colonies in Rio Grande do Sul (Brazil); 1980

Table 1: Descriptive Statistics

Variable	All	Control	Treated
ln income	9.017	9.000	9.428
	[0.954]	[0.964]	[0.524]
Illiteracy rate	0.188	0.189	0.167
	[0.061]	[0.062]	[0.058]
Distance to Porto Alegre (10 km)	23.239	23.565	15.158
	[12.799]	[12.822]	[9.590]
Share Protestant	0.114	0.108	0.250
	[0.116]	[0.108]	[0.208]
Share German-born	6.427	6.156	11.555
	[32.887]	[33.620]	[12.515]
Share Slaves in 1872	0.164	0.168	0.105
	[0.053]	[0.049]	[0.077]
Farm Size in 1872	0.078	0.077	0.088
	[0.103]	[0.105]	[0.099]

Notes: Mean values, standard deviations in brackets.

4   Empirical Strategy

4.1   Estimating equations

In order to estimate the effect of the German colonies on contemporary outcomes, I use
the following econometric model:

$$Y_{1980,j} = \alpha + \beta d\left(G_j\right) + \delta_{k(j)} + \gamma GEO_j + \epsilon_j \tag{1}$$

where Y_{jk} denotes the current human capital or log of per capita GDP in municipality
j in region k. G measures the presence of a German colony at the municipality level
and the d function is the distance to the nearest German colony in 10 kilometers. β is
the coefficient of interest, which in the case of a beneficial effect would be negative in
the distance. **GEO** is a vector that contains a set of geographic variables that includes
latitude, longitude, elevation, terrain, distance to Porto Alegre (10 km), distance to the
nearest coast (10 km) and distance to the nearest river (10 km) with a corresponding
vector of coefficients γ. δ captures a region-fixed effect, α is a generic constant and ϵ is
an idiosyncratic error term.

To actually estimate the equations above, I use OLS with fixed effects. I cluster the
standard errors at the region-level throughout, except when the number of observations
is too small and I jointly report bootstrap standard errors. I also conduct robustness
tests using spatially adjusted Conley standard errors (results not shown). As an exception,
I use OLS and probit specifications for the individual level data in the cultural section.

4.2   Identification

The causal identification of the German settlement effect hinges on several assumptions.
First, the historical record suggests that the foundation of the missions proceeded
in a relatively haphazard manner. This issue will be examined further in the section
examining covariate balance test for conditional independence of German colony assignment.
I also control directly for geographic conditions such as lower altitude and proximity
to rivers that might have influenced the initial settlement choices. Since the Brazilian
government might assigned Germans to places with more favorable geographic conditions,
it is important to include these variables directly in the estimation. Finally, I use
standard econometric techniques such as placebo tests and comparison to the Italian
settlement to confirm the causal effect of the German settlement.

5 Main Results

Table 2 presents the main empirical results on the long-term effects of German colonial settlement on contemporary literacy and income levels across municipalities. The estimated coefficients indicate that municipalities located closer to former German colonies exhibit significantly better outcomes in both education and income. Specifically, a 10 km reduction in distance to ex-German colonies is associated with a decline in illiteracy rates of approximately 0.33 to 0.37 percentage points and an increase in log income per capita of about 3.9 per cent to 3.8 per cent, depending on the specification. These effects remain statistically significant and of similar magnitude even after including a comprehensive set of geographical controls. The consistency of the estimates across specifications strengthens the interpretation that proximity to areas of German settlement has had a persistent, positive influence on human capital accumulation and economic development, plausibly through historical spillovers in education, skills, and local institutions.

Table 2: Current illiteracy and income and colonial German settlement

	Illiteracy rate		ln income	
	(1)	(2)	(3)	(4)
Distance to ex-German	0.0033*	0.0037*	-0.0393*	-0.0376*
colonies (10 km)	(0.0014)	(0.0017)	(0.0114)	(0.0151)
Geographical controls	No	Yes	No	Yes
Observations	232	231	232	231
R-squared	0.1889	0.2448	0.0673	0.1525

Notes: All regressions include meso-region fixed effects. Geographic controls: distance to Porto Alegre (10 km), municipality area, latitude, longitude, elevation, terrain, distance to the closest river (10 km), and distance to the closest coast (10 km). Standard errors in parentheses and clustered at the meso-region level. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Source: Brazil Census, 1980.

5.1 Covariate Balance Test for Conditional Independence of German Colony Assignment

Although the main regressions include geographic controls, concerns may remain that the placement of German colonies was not random. For the identifying assumption to hold (that German settlement was quasi-random) it must be conditionally uncorrelated

with pre-existing geographic characteristics that could have influenced either settlement location or long-term development. To assess this, I conduct a covariate balance test by regressing several geographic features on distance to ex-German colonies. These features serve as proxies for land quality and accessibility, and may have influenced both settler and government decisions in land allocation.

Table 3: Geographical covariates balance test

| | Dependent Variable | | | | | | | |
	Terrain	Distance to Porto Alegre	Distance to Closest river	Distance to Closest coast	Municipality Area	Latitude	Longitude	Elevation
Distance to Ex-German Colonies	-0.113*	1.364***	0.040	8.061	119.553	0.008	-0.111*	4.610
	(0.061)	(0.350)	(0.089)	(4.165)	(73.439)	(0.054)	(0.057)	(9.892)
Observations	231	232	232	232	232	232	232	232
R-squared	0.048	0.401	0.004	0.107	0.233	0.002	0.227	0.011

Notes: All regressions include meso-region fixed effects. Robust standard errors clustered by meso-region are shown in parentheses. $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The results indicate that distance to ex-German colonies is largely uncorrelated with most pre-determined geographic characteristics. Specifically, there is no statistically significant relationship between distance and proximity to the nearest river (Column 4) or elevation (Column 9), both of which could serve as proxies for trade accessibility and topographical suitability. While the coefficient on distance to Porto Alegre is strongly significant (Column 2), this reflects the geographical concentration of settlements in the state rather than endogenous placement. The weak correlation with terrain and longitude (both significant at the 10 per cent level) suggests some minor geographic sorting, but the small magnitudes and low explanatory power (R-squared values below 0.05) mitigate concerns of systematic bias. Overall, the balance test provides suggestive evidence that the identifying variation in distance to former German colonies is plausibly exogenous with respect to major geographic determinants of development.

Additional analyses below strengthen the validity of the identification strategy. First, placebo tests using artificially constructed treatments—based on the Mahalanobis distance to a set of randomly selected municipalities—do not produce statistically significant associations with current literacy or income outcomes. This suggests that the observed effects are unlikely to be spurious artifacts of spatial correlation. Second, when the same analysis is conducted using distance to Italian settlements instead of German ones, the results are not statistically significant and show stronger correlations with pre-determined geographical covariates. Lastly, excluding São Leopoldo, the

earliest and most prominent German colony, from the sample does not materially change the results. Together, these findings mitigate concerns about non-random assignment of German colonies and provide further support for the interpretation that proximity to historical German settlements has had a persistent, independent impact on development outcomes.

6 Placebo Tests

As previously discussed, German colonies may have been systematically established in locations that were inherently more favorable, beyond what is captured by observable geographic characteristics. To address the potential endogeneity of settlement location, I implement a series of placebo tests. Specifically, I identify municipalities that resemble the German colonies based on observable geographic features.

The inclusion of geographic controls in the main regression model already mitigates the concern that treatment effects are driven purely by observable locational advantages. However, this strategy cannot rule out the possibility that German colonies were systematically located in areas with advantageous unobserved characteristics such as soil quality or proximity to informal trade routes. The placebo tests therefore serve a complementary role by asking whether municipalities that are similar to treated ones in observable geography, but not actually treated, also exhibit treatment-like effects. If the estimated effects are absent or significantly attenuated in the placebo groups, this suggests that the results are not driven by unobserved geographic selection. Moreover, I also conduct a more conservative placebo test, which excludes municipalities overlapping with original 1872 German areas. This helps rule out concerns about spatial spillovers. Taken together, these placebo strategies reinforce the interpretation that the observed effects are attributable to the historical presence of German settlers rather than pre-existing geographic advantages, whether observed or unobserved.

6.1 Placebo Assignment Using Mahalanobis Distance

To investigate whether the observed effects are driven by the inherent characteristics of the locations where German colonies were established—rather than by the presence of the settlers themselves—I construct a placebo treatment group using a matching strategy based on observable geographic characteristics. Specifically, I identify municipalities that closely resemble the actual German-settled municipalities in terms of four covariates: distance to Porto Alegre, distance to the nearest riverline, distance to the coast, and

elevation. These variables are also included as controls in the main regression analysis and are selected to capture key geographic features relevant to early settlement patterns and economic development.

To match municipalities based on these observables, I use the Mahalanobis distance metric. The Mahalanobis distance between a vector of characteristics X_i for municipality i and the mean vector μ of the treated group is given by:

$$D_M(X_i, \mu) = \sqrt{(X_i - \mu)'\Sigma^{-1}(X_i - \mu)},$$

where Σ is the sample covariance matrix of the covariates. Unlike the Euclidean distance, the Mahalanobis distance accounts for the correlation structure among variables and scales them appropriately, making it particularly suitable for matching on multiple correlated continuous variables. Each control municipality is ranked by its distance to the mean profile of treated municipalities, and those with the smallest distances are selected as placebo-treated units.

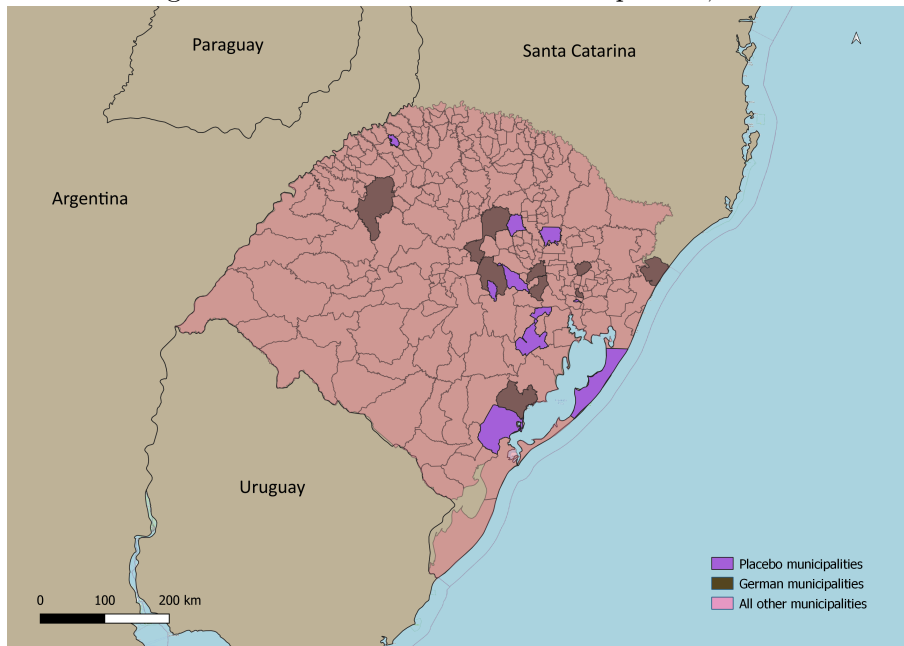Figure 2: Location of Placebo Municipalities, 1980

Table [X] reports the results of a placebo test in which "treatment" is reassigned to control municipalities that resemble historically German-settled areas based on observable geographic characteristics. This exercise aims to evaluate whether the effects observed in the main analysis are likely to be driven by favorable geography rather than by the

13

historical presence of German colonies.

The placebo results show no statistically significant association between distance to placebo-treated municipalities and illiteracy rates. The coefficients are close to zero both with and without geographic controls (0.001 and 0.000, respectively), suggesting no systematic educational advantage associated with proximity to the placebo group.

For income, the placebo test yields a statistically significant negative coefficient when no geographic controls are included: a 10 km decrease in distance to a placebo-treated municipality is associated with a 2.4 per cent increase in income (significant at the 10 per cent level). However, this effect disappears entirely once geographic covariates are added, with the coefficient dropping to -0.006 and losing statistical significance.

These patterns are informative. The disappearance of the income effect upon inclusion of geographic controls indicates that the initial placebo income association was likely capturing geographic sorting rather than a treatment-like effect. In contrast, the main treatment results (from the actual German colonies) remain statistically significant and sizable even after controlling for the same geographic features.

Together, these findings underscore that geography alone cannot account for the socioeconomic advantages observed near former German colonies. The placebo exercise acts as a falsification test, helping to rule out that observed treatment effects are simply artifacts of locational advantage. The fact that the placebo regressions yield no robust effects—while the main treatment remains significant—bolsters a causal interpretation of the long-run influence of German settlement.

Table 4: Placebo Test: Effect of Proximity to Placebo-Treated Municipalities

	Illiteracy rate	Illiteracy rate	ln income	ln income
Distance to ex-German	0.001	0.000	-0.024*	-0.006
colonies (10 km)	(0.001)	(0.000)	(0.010)	(0.009)
Geographical controls	No	Yes	No	Yes
Observations	223	223	223	223
R-squared	0.1858	0.2262	0.0460	0.1482

Notes: All regressions include meso-region fixed effects. Geographic controls: distance to Porto Alegre (10 km), municipality area, latitude, longitude, elevation, terrain, distance to the closest river (10 km) and distance to the closest coast (10 km). Standard errors are shown in parentheses and are clustered at the meso-region level. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Source: Brazil Census, 1980.

In a further robustness check, I impose an additional restriction: only municipalities that, in 1980, did not geographically overlap with any 1872 German-settled municipality are eligible for placebo treatment. This restriction is motivated by historical considerations. German colonies were established prior to 1872, which is also the first year for which I have census and municipal-level data. If a 1980 municipality overlaps with a German settlement from 1872, then assigning it to the placebo treatment group could lead to contamination through spillover effects, as part of its territory may have historically benefited from the presence of Germans. By excluding such overlapping municipalities, this stricter placebo test addresses concerns about endogenous spatial persistence and potential treatment diffusion.

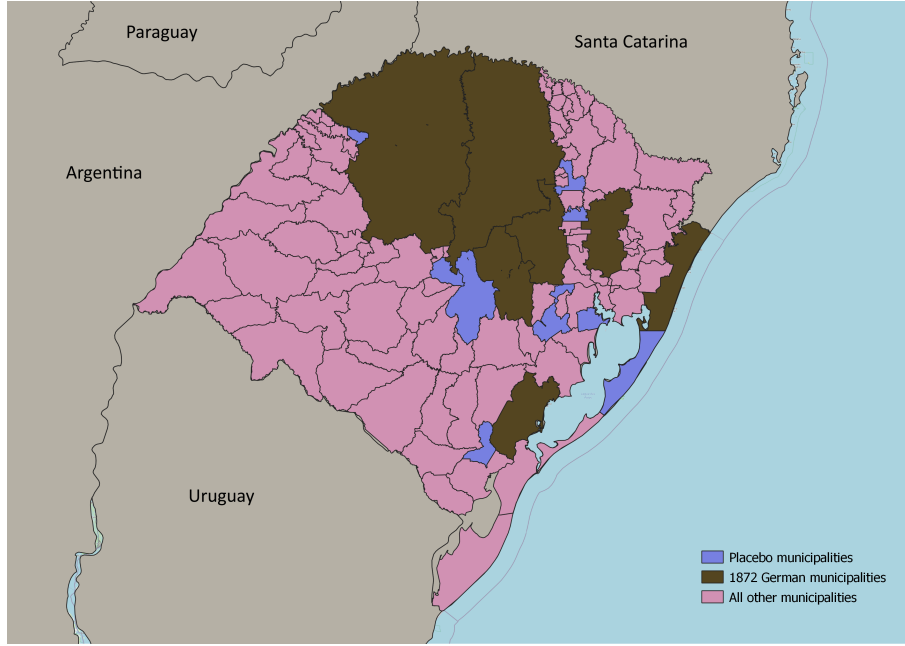Figure 3: Location of Strict Placebo Municipalities, 1980

Table 5: Strict Placebo Test: Effect of Proximity to Placebo-Treated Municipalities

	Illiteracy rate	Illiteracy rate	ln income	ln income
Distance to ex-German	0.001	-0.001	-0.025	0.006
colonies (10 km)	(0.002)	(0.002)	(0.015)	(0.043)
Geographical controls	No	Yes	No	Yes
Observations	101	101	101	101
R-squared	0.2475	0.3212	0.0756	0.2567

Notes: All regressions include meso-region fixed effects. Geographic controls: distance to Porto Alegre (10 km), municipality area, latitude, longitude, elevation, terrain, distance to the closest river (10 km) and distance to the closest coast (10 km). Standard errors are shown in parentheses and are clustered at the meso-region level. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Source: Brazil Census, 1980.

Table [X] presents results from a stricter placebo test where municipalities that historically belonged to German-settled areas in 1872 are explicitly excluded from the placebo treatment group. This test addresses potential spillover effects and contamination concerns by focusing only on 1980 municipalities outside the original German colonial boundaries.

The estimated coefficients for the effect of proximity to these stricter placebo-treated municipalities are small, statistically insignificant, and vary in sign across specifications. For illiteracy, the coefficients are 0.001 both without and with geographic controls, indicating no meaningful relationship. For log income, the coefficients are -0.016 (without controls) and 0.019 (with controls), neither of which are statistically distinguishable from zero.

Together, the placebo analyses provide robust falsification evidence supporting the causal interpretation of the main results. The first placebo test, which assigns treatment to geographically similar but untreated municipalities, finds no consistent socioeconomic advantage after controlling for geography. The stricter placebo test, which excludes any municipality overlapping with historical German settlements to avoid spillovers, similarly finds no significant treatment-like effects.

The consistent failure of placebo treatments to reproduce the significant effects seen in the actual treatment groups strengthens the argument that the socioeconomic advantages proximate to German colonies are not merely artifacts of geographic location or spillover but reflect a genuine historical causal impact.

7 Robustness

Table 6 investigates the sensitivity of the results to the exclusion of São Leopoldo, the first German colony and regional German hub. São Leopoldo may have played a disproportionately large role in driving the observed long-term effects on income and literacy due to its historical status as the initial settlement point, which likely attracted additional migrants, investments, and infrastructure development. This "hub" effect could generate differential growth advantages not representative of other German colonies, potentially biasing the overall estimates. By excluding São Leopoldo, I test whether the results are robust to removing this influential municipality, ensuring that the estimated effects reflect a broader settlement pattern rather than being driven by a single dominant location. The findings show that the estimates remain very similar, supporting the robustness of the main results.

17

Table 6: Robustness Check: Excluding São Leopoldo (German Hub)
Effect of Distance to ex-German Colonies on Illiteracy and Income

	Illiteracy rate		ln income	
Distance to ex-German	0.0030**	0.0030*	-0.0363**	-0.0331*
colonies (10 km)	(0.0012)	(0.0013)	(0.0126)	(0.0139)
Geographic controls	No	Yes	No	Yes
Observations	232	231	232	231
R-squared	0.1811	0.2388	0.0633	0.1511

Notes: Both regressions include meso-region fixed effects. Geographic controls include distance to Porto Alegre (10 km), municipality area, latitude, longitude, elevation, terrain, distance to the closest river (10 km), and distance to the closest coast (10 km). Standard errors in parentheses, clustered at the meso-region level. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Source: Brazil Census, 1980.

8 Channels of Persistence

This section investigates potential mechanisms through which proximity to 19th-century German settlements may have shaped divergent development paths. I test whether the relationship between proximity to these colonies and modern income operates through observable historical and cultural characteristics. Specifically, I focus on structural transformation, occupation specialisation, urban agglomeration and health as potential mechanisms that can help explain the persistent differences in human capital across municipalities.

Table 7 shows that the strong negative association between distance to German colonies and current income per capita remains robust to the inclusion of various controls. Column (1) reports the baseline estimate, where each 10-kilometer increase in distance is associated with a 3.9 per cent decline in log income. In Column (2), I control for current illiteracy rates, allowing me to assess whether education levels mediate the relationship. The coefficient on distance attenuates modestly to -0.029, consistent with the view that early human capital accumulation was one channel through which German presence shaped long-run outcomes.

Column (3) includes the contemporary Protestant share to examine whether the results are driven by religious affiliation. The estimated effect remains similar in magnitude and statistically significant, suggesting that Protestantism may have contributed to development but does not fully explain the observed relationship.

Column (4) adds the estimated literacy rate in 1872, capturing pre-existing differences

in educational attainment linked to early settlement. The coefficient on distance remains stable, indicating that long-run effects are not purely the result of initial educational advantages.

In Columns (5) and (6), I include the estimated Protestant share. Both variables are positively and significantly associated with modern income. These results are consistent with a persistent cultural influence (e.g., Weberian work ethic).

Column (7) introduces the share of the population enslaved in 1872, under the hypothesis that municipalities with lower slave dependence may have followed more inclusive growth trajectories. While the point estimate on distance remains large and negative, the sample size drops substantiall, and standard errors increase, limiting statistical precision.

Finally, Column (8) includes all key controls jointly. Although the coefficient on distance declines slightly to −0.023, it remains statistically significant at the 10 per cent level, suggesting that the relationship is robust to the inclusion of multiple historical factors. Notably, the coefficients on initial illiteracy and Protestant share remain highly significant, reinforcing the importance of early human capital and cultural traits.

Overall, these controls are important for testing the validity of the identification assumption, that distance to German colonies is uncorrelated with other unobserved determinants of economic development. Reassuringly, the estimated effects remain statistically significant and relatively stable in magnitude across specifications. The fact that the distance coefficient remains significant even in the fully saturated model supports the interpretation that German settlements generated persistent regional spillovers beyond the characteristics captured in available historical data.

Table 7: Effect of Proximity to German Colonies on Current Income, Controls

Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Log Income (1980)				
Distance to German Colonies (per 10km)	-0.039**	-0.029*	-0.029**	-0.040**	-0.040**	-0.040***	-0.022	-0.023*
	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	[6.340]	(0.012)
Illiteracy rate		-3.150***						-2.804**
		(0.750)						(0.860)
Protestant share			1.350***					1.047***
			(0.160)					(0.261)
Initial literacy				0.011***				0.121
				(0.001)				(0.103)
Initial Protestant share					0.059***			-0.474
					(0.005)			(0.463)
Initial manufacturing share						-0.040**		-0.504
						(0.011)		(0.424)
Slave share (1872)							-5.870	
							[6.340]	
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	230	230	230	230	230	230	28	230

Notes: Robust standard errors in parentheses; bootstrapped standard errors in square brackets. All regressions include meso-region fixed effects. Distance to German colonies is measured from the municipality centroid to the nearest historical German colony. Initial condition variables (literacy, Protestant share, manufacturing share) are estimated using a distance-weighted shift-share approach. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

8.1 Manufacturing and Agriculture Employment Shares

8.1.1 Identification Strategy

This study aims to examine the long-term impact of 19th-century German migration on the industrial structure of Rio Grande do Sul, Brazil. Specifically, I analyze the evolution of agricultural and manufacturing employment shares across municipalities from 1872 to 1980, focusing on the role of distance to São Leopoldo, the primary German settlement hub.

The identification strategy leverages the historical context of German migration to Brazil, which began in the mid-19th century due to Brazilian government incentives. This early-stage migration provides a unique setting to investigate the persistence of human capital and cultural influences on occupational structural transformation.

The identification of the German migration effect rests on several key assumptions. First, historical accounts suggest that the initial selection of São Leopoldo as a settlement

location was effectively exogenous. Second, I control for a comprehensive set of geographic factors, including proximity to rivers and the coast, which could have influenced the government's settlement decisions. Moreover, I conduct a comparative analysis with Italian migration to isolate the causal effect of Protestantism on structural transformation. This within-migrant comparison aims to address potential confounding factors related to general migration patterns.

To further bolster the identification strategy, I provide empirical evidence supporting specific persistence mechanisms of transmission. Additionally, I perform robustness checks to rule out alternative transmission channels that could potentially confound the results.

8.1.2 Results

Occupational structural transformation serves as a critical transmission channel for economic persistence. As documented in Doepke and Zilibotti (2008), human capital investment decisions are closely linked to occupational choices. This section examines the impact of distance to São Leopoldo on manufacturing and agricultural employment shares across municipalities, utilizing census data from 1872, 1940, 1950, 1960, 1970, and 1980. The analysis focuses on distance to São Leopoldo, the first and most significant German settlement in the region, characterized by a notably high concentration of German immigrants (Table ?? and Figure ??).
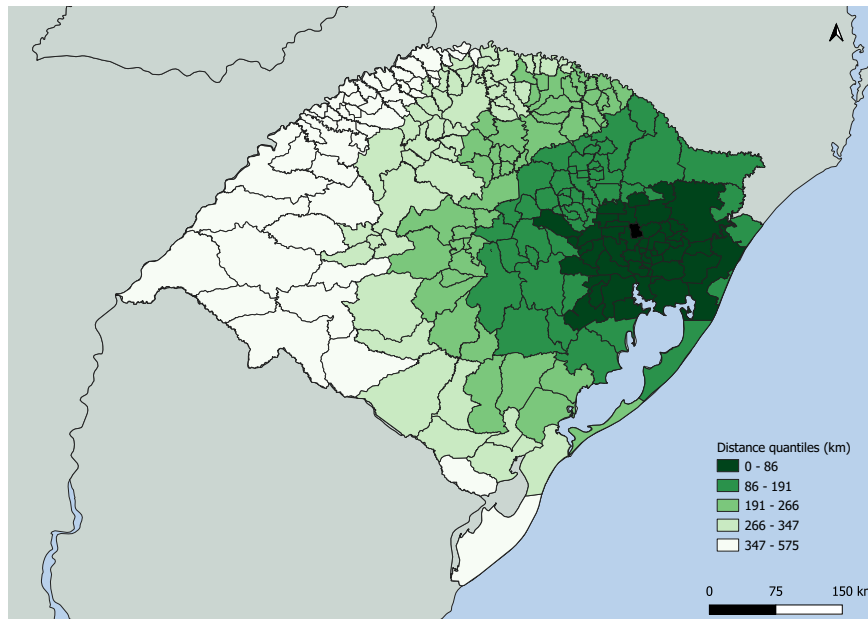
The hypothesis posits that areas with higher human capital, proxied by proximity to São Leopoldo, experienced a shift towards manufacturing and away from agriculture (Table ??). Indeed, the empirical results demonstrate that municipalities closer to São Leopoldo exhibit a statistically significant increase in manufacturing employment shares and a corresponding decline in agricultural employment shares. The magnitude of the estimated coefficients varies with the inclusion of geographic controls, but statistical significance remains consistent. Specifically, a 10-kilometer reduction in distance to São Leopoldo is associated with a 3.6 per cent increase in manufacturing employment share without geographic controls, rising to 35.3 per cent with controls. Conversely, the same reduction in distance leads to a 1.5 percent decrease in agricultural employment share without controls and an 8.8 percent decrease with controls. Overall, I find that municipalities located closer to this hub exhibit significantly higher manufacturing employment shares and correspondingly lower agricultural employment shares. This result is consistent with the presence of localized industrial development and more

Table 8 - Summary Statistics in 1872

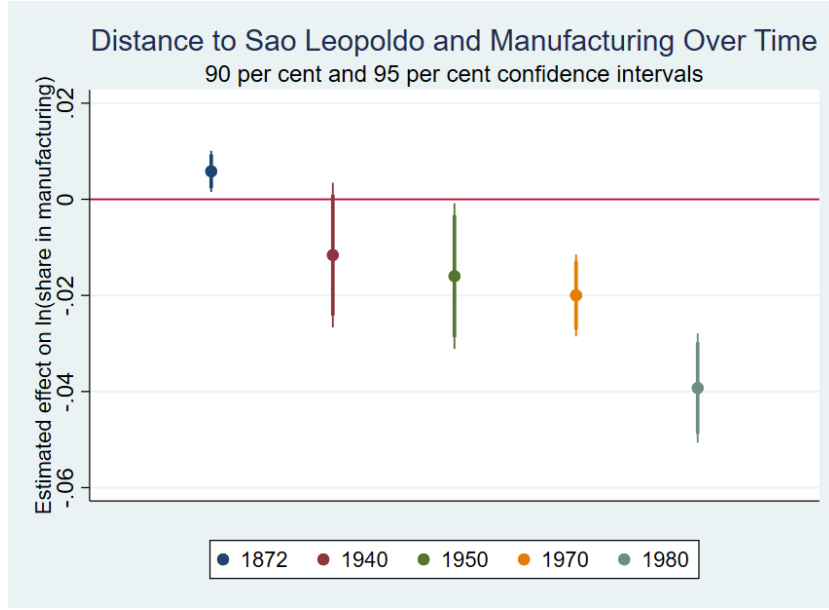| | Whole sample | | | | São Leopoldo | Garibaldi | Triunfo |
	Mean	Min	Max	Std dev			
Distance to São Leopoldo (10 km)	23.95	0	57.53	15.70	0	6.60	5.88
Latitude	-30.31	-32.57	-28.26	1.07	-29.76	-29.29	-29.94
Longitude	-53.09	-57.10	-50.27	1.78	-51.15	-51.56	-51.72
Elevation	147.22	3.18	678.86	179.15	4.78	697.83	18.35
Terrain	3.04	0.00	22.00	4.82	0.00	2	7
Distance to the nearest river (10 km)	0.50	0.02	1.88	0.51	0.02	0.30	0.43
Distance to the coast (10 km)	21.80	0.93	59.53	16.90	12.05	18.65	14.59
Share of foreigners	8.89	1.81	21.39	6.27	17.22		16.27
Share of Germans	3.37	0.15	16.22	5.19	16.22		13.46
Distance to Porto Alegre (10 km)	22.82	0	56.75	15.53	3.07	8.81	4.82
Population	13,143	4,753	35,843	7,839	29,308		10,111

Source: Brazil Census 1872; NASA Jpl Shuttle Radar Topography Mission Global 1 Arc; OpenStreetMap contributors (2025); Author's calculations. The unit of observation of the variables is the municipality.

Figure 4: Distance to São Leopoldo

(a) The map shows the municipal level boundaries for the state of Rio Grande do Sul (Brazil) in 1980 and the shades of green denote the quantiles of the distance to São Leopoldo, along with the state boundary and national level boundaries.

Figure 5

(a) Year-specific regression results with confidence intervals. The dependent variable is distance to São Leopoldo (10 km) and the regressions control for region fixed effects, distance to Porto Alegre (10 km), elevation, terrain, altitude, longitude, distance to the nearest coastline (10 km) and distance to the nearest river (10 km).

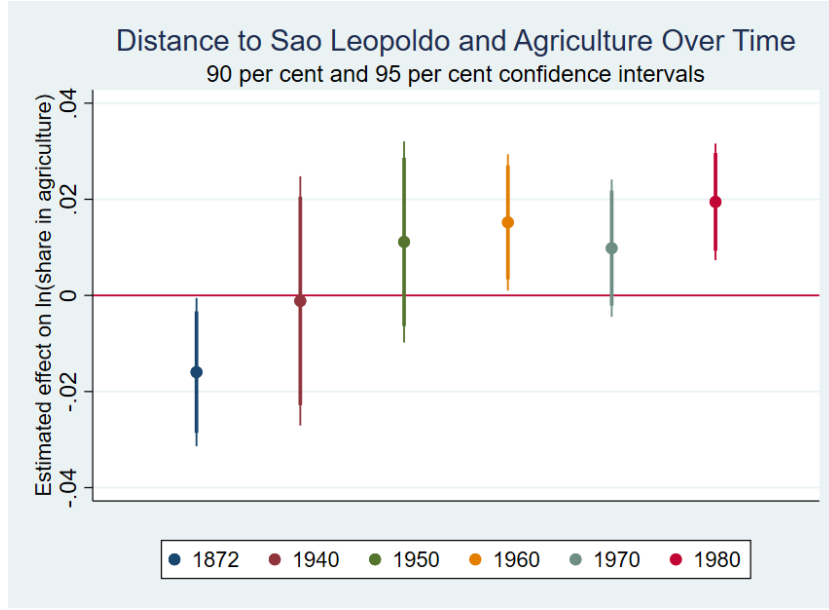advanced occupational structures in areas surrounding São Leopoldo.

To further explore the temporal dynamics of this relationship, I investigate the evolution of the impact of distance to São Leopoldo on agricultural and manufacturing employment shares from 1872 to 1980. This is accomplished by estimating separate cross-sectional regressions for each census year and examining the time-varying coefficient on distance to São Leopoldo (Figures 5 and ??). The results indicate a strengthening negative effect of German migration on manufacturing employment shares over time, particularly in the postwar period. In 1872, municipalities closer to São Leopoldo exhibited a statistically significant decrease in manufacturing employment shares. By 1950, this effect reversed, with closer proximity leading to a statistically significant increase in manufacturing employment shares. This positive effect intensified in subsequent decades. Conversely, the effect on agricultural employment shares transitioned from negative to positive over time. The postwar patterns align with the broader industrialization process in Brazil (Szirmai (2012)). Overall, these findings suggest that structural transformation serves as a significant persistence mechanism for the Protestantism effect.

Table 9: The Impact of the Distance to São Leopoldo on Manufacturing and Agricultural Employment Shares

					log share in:					
			Manufacturing					Agriculture		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Distance to São Leopoldo (10 km)	−0.036***	−0.056***	−0.062***	−0.366***	−0.353***	0.015***	0.026***	0.026***	0.12***	0.088**
	(0.003)	(0.007)	(0.010)	(0.037)	(0.038)	(0.003)	(0.004)	(0.006)	(0.038)	(0.038)
	{0.012}	{0.020}	{0.030}	{0.065}	{0.049}	{0.015}	{0.016}	{0.025}	{0.046}	{0.024}
	[0.005]	[0.011]	[0.015]	[0.055]	[0.054]	[0.006]	[0.008]	[0.010]	[0.071]	[0.064]
Distance to Porto Alegre (10 km)				0.328***	0.310***				−0.104***	−0.074**
Additional geo controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Region fixed effects	No	No	Yes	No	Yes	No	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	668	668	668	668	668	805	805	805	805	805

Notes: White standard errors in parentheses, standard errors in curly brackets are corrected for clustering at meso-region and standard errors in square brackets are corrected for spatial correlation (i.e. Conley standard errors). Asterisks denote significance at the 10(*), 5(**) and 1(***) percent level. The unit of observation of the variables is the municipality. Geographical controls: latitude, longitude, elevation, terrain and distance to the nearest coastline, distance to the nearest river. Sources: See Data section.

Figure 6

Distance to Sao Leopoldo and Agriculture Over Time
90 per cent and 95 per cent confidence intervals

Legend: 1872, 1940, 1950, 1960, 1970, 1980

(a) Year-specific regression results with confidence intervals. The dependent variable is distance to São Leopoldo (10 km) and the regressions control for region fixed effects, distance to Porto Alegre (10 km), elevation, terrain, altitude, longitude, distance to the nearest coastline (10 km) and distance to the nearest river (10 km).

However, when I broaden the analysis to include the distance to the nearest German colony, rather than just São Leopoldo, the results are no longer statistically significant. This divergence points to important heterogeneity across colonies in terms of their economic orientation and long-term influence. While São Leopoldo and its neighboring settlements appear to have been focal points of early industrial activity and skill transfer, many other German colonies remained predominantly rural and agricultural. In these areas, German settlers may still have contributed to long-run development through other channels, such as literacy gains, religious norms, or the establishment of educational institutions, but not necessarily through structural transformation in employment.

These findings suggest that the transmission mechanisms linking historical German immigration to long-run development outcomes are not uniform across space. Instead, the evidence is consistent with multiple, context-dependent channels of persistence. For some municipalities—particularly those near São Leopoldo—economic transformation through industrialization and non-agricultural employment appears to have played a central role. For others, the enduring legacy may be rooted more in cultural or

educational factors introduced by the settlers.

This spatial heterogeneity in mechanisms helps reconcile why proximity to German settlements is strongly predictive of modern income and literacy overall, even though the specific pathways through which these outcomes emerge vary across colonies. In other words, the broad effect of historical German immigration on economic development in southern Brazil appears to be mediated through different channels in different places, shaped by local settlement characteristics, economic potential, and patterns of diffusion.

8.2   Occupational Specialisation

I expand the occupational analysis by looking at skill-intensive manufacturing industries following the categorisation of Ciccone and Papaioannou (2009). I test whether distance to 1868 German colonies led to specialisation in industries that depend more intensely on human capital accumulation. Namely, the dependent variable is an indicator variable that equals 1 if an individual reports working in one of industries that according to Ciccone and Papaioannou (2009) is human capital intensive. These are, ordered from most to least human capital skill intensive (HCINT, in years): industrial chemicals, electric machinery, transport equipment, beverages, plastic products and fabricated metal products. I find a negative and significant effect for 6 out of the 7 industries in the top skill intensity ranking (Table 9). The findings within industries complement at the intensive margin the broader structural transformation results reported in Table 8.

8.3   Alternative Mechanisms

In this section, I explore alternative transmission channels: agglomeration and health.

One potential alternative channel of transmission of the Protestantism effect is population density. German colonies may have created the seeds for future urban agglomerations ((Becker et al., 2010); (Bai and Kung, 2015)). Table 10 Columns (1) and (2) display the results of this empirical test. I find that places closer to German colonies are more dense today but the results are not statistically significant.

Another possible mechanism is health. I study health as a human capital investment complementary to education ((Calvi et al., 2019); (Cagé and Rueda, 2018)). I find a negative and significant relationship between distance from German colonies in the nineteenth century and health as measured by the Brazilian health index. Moving

26

Table 9 - Distance to German colonies Effect on Skill-intensive Industries

	Industrial chemicals	Electric machinery	Transport equipment	Beverages industries	Plastic products	Fabricated metal products
HCINT	13.031	12.357	12.346	11.967	11.678	11.577
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to German Colonies (100 km)	−0.209 (0.016)***	−0.430 (0.036)***	−0.720 (0.031)***	−0.227 (0.027)***	−0.760 (0.087)***	1.410 (0.163)***
Mesoregion FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1 925 700	1 925 700	1 925 700	1 925 700	1 925 700	1 780 228
Pseudo R^2	0.0424	0.0984	0.1777	0.0853	0.1067	0.3812

Source: Brazil Census 1980 individual-level data from the IPUMS project; Quadro Statistico e Demografico de S. Pedro do Rio Grande do Sul, 1868. The dependent variable is an indicator that equals to 1 if an individual reports working in the industry in 1980 ordered from most to least by their HCINT according to Ciccone and Papaioannou (2009). Controls: foreign-born, women, ln(population), distance to Porto Alegre (100 km), municipality area, latitude and longitude. The estimation is for a probit model with robust standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. See Appendix 2 for more details on the industry classification.

100km closer to a German colony increases the level of health development by 0.0321 points. This is approximately equivalent to moving from the 50th to the 70th percentile of the distribution of the index for Rio Grande do Sul. Overall this result is consistent with the broader structural transformation results described previously.

Table 10 - Protestantism Effect on Alternative Transmission Mechanisms				
	Population density	Population density	Health index	Health index
	(1)	(2)	(3)	(4)
Distance to 1868 German Colonies (100 km)	-15.5495 (17.637)	-20.824 (28.863)	-0.032 (0.016)**	-0.042 (0.023)*
Mesoregion FE	No	Yes	No	Yes
Observations	232	232	232	232
R^2	0.1916	0.2672	0.1700	0.2014
Source: Brazil Census 1980; Quadro Statistico e Demografico de S. Pedro do Rio Grande do Sul, 1868. Population density equals total population in a given municipality divided by municipality area. The Brazilian IFDM Health Index in 2005 (earliest period available) is the dependent variable in columns (3) and (4). Controls: % foreign-born, % women, distance to Porto Alegre (100 km), latitude and longitude. The estimation is by OLS with robust standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.				

9 Empirical extension

In this section I present the main empirical extension of the main results described so far. I compare German colonies to the Italian colonies.

9.1 Italian Migration

The analysis of Italian migration provides a valuable comparative perspective, allowing for the examination of two distinct European migration experiences within a similar geographical context. This comparison is pertinent as both German and Italian migrants were subject to similar settlement policies by the Brazilian government, characterized by shared economic and geographic considerations. While both groups exhibited higher literacy rates than the native population, one key distinction lies in their religious affiliation: Italians were predominantly Catholic, whereas a significant share of Germans were Protestant.

Italian migration to Rio Grande do Sul commenced in 1875 (Rückert (2013)). The government's land allocation strategy for Italians differed from that employed for the Colony of São Leopoldo in two primary aspects: the sale of public land parcels and the implementation of a dedicated legislative and administrative framework to facilitate the establishment and operation of imperial colonies. Specifically, land parcels were offered at relatively low prices, with government incentives and subsidies further easing the purchase process (Giron (2007)). Initial settlements were established in the imperial colonies of Conde D'Eu (Garibaldi) and Dona Isabel (Bento Gonçalves) (Rückert (2013)).

To capture the effect of Italian migration, I use the distance to these Italian colonies. Column 7 in Table ?? presents descriptive statistics for the Italian colonies and for the municipality of Triunfo, from which these colonies originated, revealing a significant presence of German settlers. This overlap represents a limitation to the analysis, as the underlying cultural landscape may have been influenced by German migrants. Consequently, the location of the government's creation of the colonies Conde D'Eu (Garibaldi) and Dona Isabel (Bento Gonçalves) as an Italian settlement site may not have been entirely random, potentially reflecting a preference for locations near existing European population centers.

Table 8: Current illiteracy and income and colonial Italian settlement

| | Illiteracy rate | | ln income | | Illiteracy rate | ln income |
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to ex-Italian	0.011	0.099	0.017	-0.185	0.007	0.082
colonies (10 km)	(0.012)	(0.015)	(0.112)	(0.229)	(0.009)	(0.082)
Distance to ex-German					0.003*	-0.043***
colonies (10 km)					(0.001)	(0.012)
Geographical controls	No	Yes	No	Yes	No	No
Observations	232	231	232	231	231	231
R-squared	0.1591	0.2843	0.0407	0.1513	0.1943	0.0705

Notes: All regressions include meso-region fixed effects. Geographic controls: distance to Porto Alegre (10 km), municipality area, latitude, longitude, elevation, terrain, distance to the closest river (10 km), and distance to the closest coast (10 km). Standard errors in parentheses and clustered at the meso-region level. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Source: Brazil Census, 1980.

I test directly whether the Italian migration in the nineteenth century had the same effects as the German migration in the long term, by reestimating the human capital and income equations, using instead distance to the nearest Italian colony. The results, or lack thereof, can be seen in Table 8. I find no effect for either modern literacy or income. In a horserace between the German and Italian colonies I find that the beneficial effect on education and income is preserved for the German colonies and it is not statistically significant for Italian colonies.

There a few factors that could be driving this disparity. First, the observed differences in results between the two migration groups can be attributed to their distinct cultural characteristics. Specifically, the predominantly Protestant German migrants and the predominantly Catholic Italian migrants likely exhibited different cultural traits. This religious distinction represents a key differentiating factor between the two settlements. Second, the skill set of the settlers. In Table **??** in Appendix D I examine this possibility. Areas closer to Italian colonies have moved towards agriculture and did not industrialise as much as German colonies.

Table 9: Geographical covariates balance test

	Dependent Variable							
	Terrain	Distance to Porto Alegre	Distance to Closest river	Distance to Closest coast	Municipality Area	Latitude	Longitude	Elevation
Distance to Ex-Italian Colonies	-0.137	8.970***	0.048***	7.974**	230.68	0.087	-0.898***	-120.61**
	(0.356)	(0.427)	(0.012)	(2.702)	(222.93)	(0.160)	(0.195)	(32.970)
Observations	231	232	232	232	232	232	232	232
R-squared	0.079	0.974	0.159	0.861	0.604	0.839	0.851	0.674

Notes: All regressions include meso-region fixed effects. Robust standard errors clustered by meso-region are shown in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

10 Conclusion

This paper has investigated the long-term effects and transmission mechanisms of a historical human capital shock by exploiting the 19th-century German migration to southern Brazil as a quasi-natural experiment. Employing a unique combination of archival records and municipal-level data, I find evidence that areas closer to historical German colonies, established between 1824 and 1867, experienced more rapid economic growth as measured by GDP per capita. Furthermore, when examining the long-term impact of German migration on structural transformation, specifically changes in manufacturing and agriculture employment shares, a comparative analysis with Italian

migration was employed. This within-migrant comparison allowed for the isolation of the causal effect of German settlement on structural transformation, addressing potential confounding factors related to general migration patterns. The observed enduring effects are consistent with transmission mechanisms involving structural transformation and occupational specialization, as municipalities closer to German colonies exhibit a shift away from agriculture and a greater specialization in human capital-intensive industries.

This study contributes to the literature on cultural economics and the persistence of human capital by providing empirical evidence of the long-term impact of historical migration patterns on contemporary economic outcomes. Unlike previous research, this paper explicitly identifies and validates specific transmission mechanisms, offering a nuanced understanding of how historical cultural shocks can shape regional economic disparities. The findings underscore the importance of considering historical context in contemporary economic analysis and policy design. For policymakers, this research highlights the potential for long-term benefits from policies that promote human capital development and facilitate the integration of skilled migrants. Furthermore, it suggests that cultural factors, such as religious affiliation, can play a significant role in shaping economic trajectories over extended periods.

Future research could expand upon this study by exploring the specific channels through which Protestantism influenced human capital formation, such as the role of educational institutions or social networks. Additionally, investigating the potential spillover effects of German migration on neighboring regions or examining the impact of other cultural groups on economic development in Brazil would further enhance our understanding of the long-term dynamics of cultural and human capital legacies. Moreover, using more modern data, and seeing if these results still hold, would be a valuable expansion of this research.

References

Bai, Y. and Kung, J. K.-s. (2015). Diffusing knowledge while spreading god's message: Protestantism and economic prosperity in china, 1840–1920. *Journal of the European Economic Association*, 13(4):669–698. [26]

Becker, S. O., Cinnirella, F., and Woessmann, L. (2010). The trade-off between fertility and education: evidence from before the demographic transition. *Journal of Economic Growth*, 15:177–204. [26]

Becker, S. O. and Woessmann, L. (2009). Was weber wrong? a human capital theory of protestant economic history. *The quarterly journal of economics*, 124(2):531–596. [3]

BELLOTTI, K. K. (2000). Uma igreja invisível? protestantes históricos e meios de comunicação no brasil (anos 50 a 80). *Monografia de conclusão de graduação em História. Campinas: IFCH/UNICAMP*. [6]

Cagé, J. and Rueda, V. (2018). Sex and the mission: The conflicting effects of early christian investments on sub-saharan africa's hiv epidemic. *Manuscript, Sciences Po*. [26]

Calvi, R., Mantovanelli, F., and Hoehn-Velasco, L. (2019). The protestant legacy: Missions and human capital in india. *Available at SSRN 3354891*. [26]

Chesnut, R. A. (1997). *Born again in Brazil: The Pentecostal boom and the pathogens of poverty*. Rutgers University Press. [6]

Ciccone, A. and Papaioannou, E. (2009). Human capital, the structure of production, and growth. *The review of economics and statistics*, 91(1):66–82. [26]

CUNHA, J. L. d. (2006). Imigração e colonização alemã. *História geral do Rio Grande do Sul*, 2. [5]

de Geografia e Estatística, I. B. (2000). Brasil: 500 anos de povoamento. [2]

Doepke, M. and Zilibotti, F. (2008). Occupational choice and the spirit of capitalism. *The Quarterly Journal of Economics*, 123(2):747–793. [21]

Ferenczi, I. (1929). International migration statistics. In *International Migrations, Volume I: Statistics*, pages 47–76. NBER. [6]

Ferreira, J. A. (1959). *História da Igreja Presbiteriana do Brasil en comemoração ao seu primeiro centenário*, volume 1. Casa Editora Presbiteriana. [6]

Galor, O. (2005). From stagnation to growth: unified growth theory. *Handbook of economic growth*, 1:171–293. [3]

Giron, L. S. (2007). *História da imigração italiana no Rio Grande do Sul*. Edições EST. [29]

Grosjean, P. (2011). The weight of history on european cultural integration: a gravity approach. *American Economic Review*, 101(3):504–508. [3]

Guiso, L., Sapienza, P., and Zingales, L. (2016). Long-term persistence. *Journal of the European Economic Association*, 14(6):1401–1436. [3]

Hornung, E. (2014). Immigration and the diffusion of technology: The huguenot diaspora in prussia. *American Economic Review*, 104(1):84–122. [3]

Iotti, L. H. (2001). *Imigração e colonização: legislação de 1747 a 1915*. Educs. [5]

Klug, A. Q., Dal Molin, A., Tesmer, D. T., TESSMANN, J. M. d. C., VARGAS, J. G., and RÖHNELT, P. B. C. (2012). A cultura pomerana no rio grande do sul. *Universidade Federal de Pelotas*. [5]

Lando, A. M. and Barros, E. C. (1976). *A colonização alemã no Rio Grande do Sul: uma interpretação sociológica*. Number 11. Editora Movimento. [4]

Luebke, F. C. (1987). Germans in brazil: a comparative history of cultural conflict during world war i. [4, 5, 6]

Michalopoulos, S. and Papaioannou, E. (2020). Historical legacies and african development. *Journal of Economic Literature*, 58(1):53–128. [3]

Murard, E. and Sakalli, S. O. (2020). Mass refugee inflow and long-run prosperity: lessons from the greek population resettlement. Technical report, CReAM Discussion Paper Series. [3]

Neiva, A. H. and Diégues Júnior, M. (1955). Cultural assimilation of immigrants in brazil. [6]

Nunn, N. (2009). The importance of history for economic development. *Annu. Rev. Econ.*, 1(1):65–92. [3]

Oberacker Jr, O. J. (1975). *Jorge Antônio von Schaeffer, criador da primeira corrente emigratoria alema para o Brasil*, volume 6. Editora Metrópole. [4]

Piassini, C. E. (2017). Imigração alemã e política. *Os deputados provinciais Koseritz, Kahlden, Haensel, Brüggen e Bartholomay. Porto Alegre: Assembleia Legislativa do Estado do Rio Grande do Sul*. [34]

Rückert, F. Q. (2013). A colonização alemã e italiana no rio grande do sul: uma abordagem na perspectiva da história comparada. *Revista Brasileira de História & Ciências Sociais*, 5(10). [29]

Saviani Filho, H. (2008). O processo de colonização no rio grande do sul: o caso de são leopoldo no século xix. [5]

Schütz, L. M. M. (1974). Imigração alemã: processo, costumes e influências. *Anais do I Simpósio de História da Imigração e Coloninzação Alemã no Rio Grande do Sul*, pages 271–318. [5]
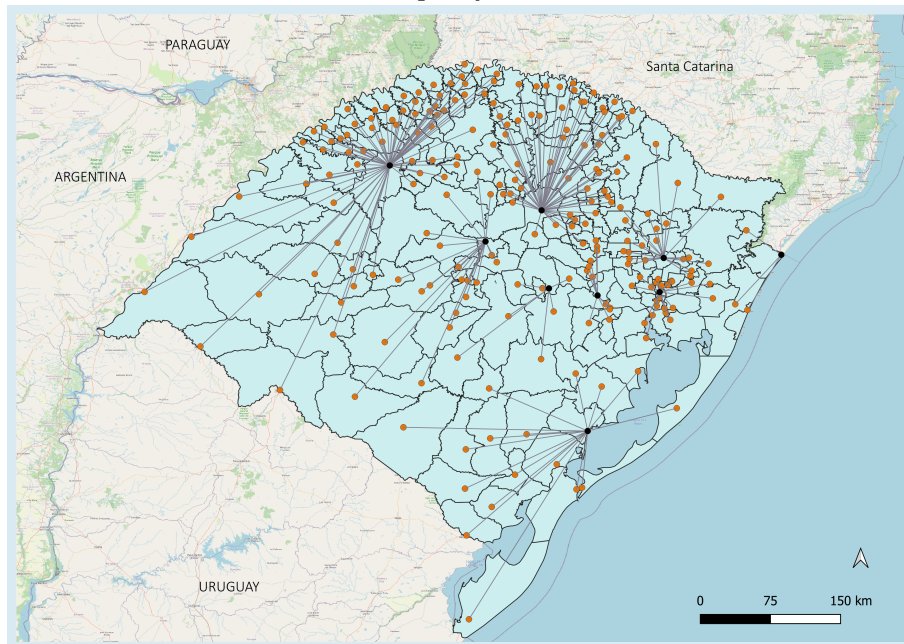
Spolaore, E. and Wacziarg, R. (2013). How deep are the roots of economic development? *Journal of economic literature*, 51(2):325–369. [3]

Stolz, Y., Baten, J., and Botelho, T. (2013). Growth effects of nineteenth-century mass migrations:"fome zero" for brazil? *European Review of Economic History*, 17(1):95–121. [6]

Szirmai, A. (2012). Industrialisation as an engine of growth in developing countries, 1950–2005. *Structural change and economic dynamics*, 23(4):406–420. [23]

Tabellini, G. (2008). Institutions and culture. *Journal of the European Economic association*, 6(2-3):255–294. [3]

Voigtländer, N. and Voth, H.-J. (2012). Persecution perpetuated: the medieval origins of anti-semitic violence in nazi germany. *The Quarterly Journal of Economics*, 127(3):1339–1392. [3]

Weber, M. (1934). Die protestantische ethik und der geist des kapitalismus. [3]

A   Appendix A

The variation that I picking up in the main analysis comes from the distance displayed in the figure below.

German Colonies Founded before 1868	
Municipality name	Year of foundation
São Leopoldo	1824
Torres	1826
Taquari	1849
Santa Cruz do Sul	1849
Estrela	1853
Soledade	1857
Santo angelo	1857
São Lourenco do Sul	1858
Nova Petropolis	1858
Camaquam	1864
Source: Piassini (2017).	

Figure 7: Distance to Nearest Municipality where German Colonies were Located

(a) The map shows the municipal level boundaries for the state of Rio Grande do Sul (Brazil) in 1980, the capital of each municipality, municipality capitals where 1868 German colonies were located (marked with a black circle), the nearest distance between municipality capitals and municipalities where German colonies in 1868 were located, along with the state boundary and national level boundaries.

A Appendix 2

Skill-intensive Industry classification		
Skill-intensive industry	IPUMS Industry code	IPUMS Industry code name
Industrial Chemicals	200	chemical industry
	418	chemical and pharmac. prod. trade
Electric machinery	130	electric material industry
Transport equipment	140	transport material industry
Beverages industries	270	beverage industry
Plastic products	230	plastic material industry
Fabricated metal products	55	firing/prosp metallic minerals
	58	extraction of metallic minerals

A Appendix C: Balance Tests for Placebo Treatment Assignments

To validate the quality of the placebo treatment assignments, I conduct balance tests comparing the means of key geographic covariates used in the matching procedure. Specifically, I assess balance for four variables: distance to Porto Alegre, distance to the nearest river, distance to the nearest coast, and elevation. The standardized mean differences (SMDs) are reported for each variable.

The balance tests examine whether the placebo-treated municipalities differ significantly from control municipalities on these baseline characteristics. Good balance—commonly defined as an SMD below 0.1 indicates that the placebo groups are well matched and thus suitable as a counterfactual for the placebo analyses.

Table 10 presents the results of the balance tests for both placebo exercises described in the main text: (1) the placebo treatment based on Mahalanobis distance matching, and (2) the stricter placebo excluding municipalities historically belonging to German colonies in 1872 based on Mahalanobis distance matching.

Table 10: Balance Test for Placebo-Treated vs. Control Municipalities

Variable	Placebo Test 1	Placebo Test 2 (Strict)
Distance to Porto Alegre (km)	0.063	0.095
Distance to nearest river (km)	0.132	0.142
Distance to nearest coast (km)	0.065	0.067
Elevation (meters)	0.039	0.003

Notes: Standardized mean differences (SMDs) are calculated as the difference in means between placebo-treated and control groups divided by the pooled standard deviation. An SMD below 0.1 is generally considered indicative of good balance. All four geographic variables show excellent balance in both placebo samples.

A   Appendix D

I assess the persistent effects of distance to Italian colonies on current income by re-estimating the specifications used for German colonies, substituting distance to Italian colonies as the key independent variable (Table ??). I observe a statistically significant negative effect on the manufacturing share, albeit with a smaller magnitude compared to the effect of distance to São Leopoldo (see discussion below). Specifically, without geographic controls, the effect of distance to Italian colonies on the manufacturing share is 3.2 per cent, compared to 3.6 per cent for São Leopoldo. With geographic controls, this increases to 8.9 percent, compared to 35.3 percent for São Leopoldo.

The effect on agricultural employment shares is not statistically significant in certain specifications (Table ??, columns 7 and 8). Furthermore, in contrast to the effect of São Leopoldo, I find that with geographic controls, closer proximity to Italian colonies leads to a 6.1 percent increase in the agricultural employment share. Similarly, the year-by-year regression analysis indicates that the estimated effect of distance to Italian colonies on the agricultural share is not consistently statistically significant across all years.
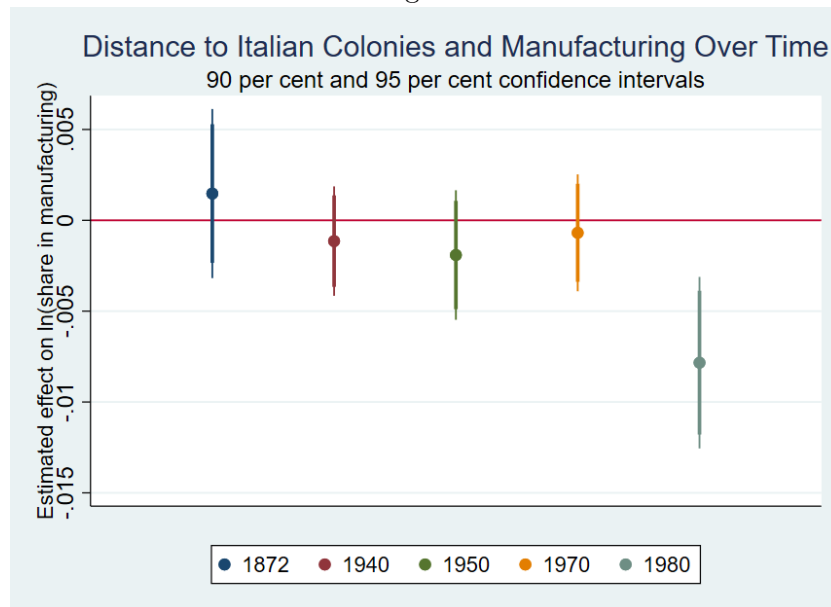
Table 10: The Impact of the Distance to Italian Colonies on Structural Transformation

	log share in:									
	Manufacturing					Agriculture				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Distance to Italian Colonies (10 km)	−0.032*	−0.035**	−0.039***	−0.006	−0.090***	0.006	−0.004	0.003	−0.118**	−0.093
	{0.014}	{0.010}	{0.014 }	{0.019}	{0.020}	{0.014}	{0.013}	{0.010}	{0.034}	{0.061}
	[0.003]	[0.006]	[0.007]	[0.014]	[0.018]	[0.003]	[0.005]	[0.006]	[0.011]	[0.015]
Geographic controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Region fixed effects	No	No	Yes	No	Yes	No	No	Yes	No	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	668	668	668	668	668	805	805	805	805	805

Notes: White standard errors in parentheses, standard errors in curly brackets are corrected for spatial correlation (i.e. Conley standard errors). Asterisks denote significance at the 10(*), 5(**) and 1(***) percent level. The unit of observation of the variables is the municipality. Geographical controls: latitude, longitude, elevation, terrain and distance to the nearest coastline, distance to the nearest river.
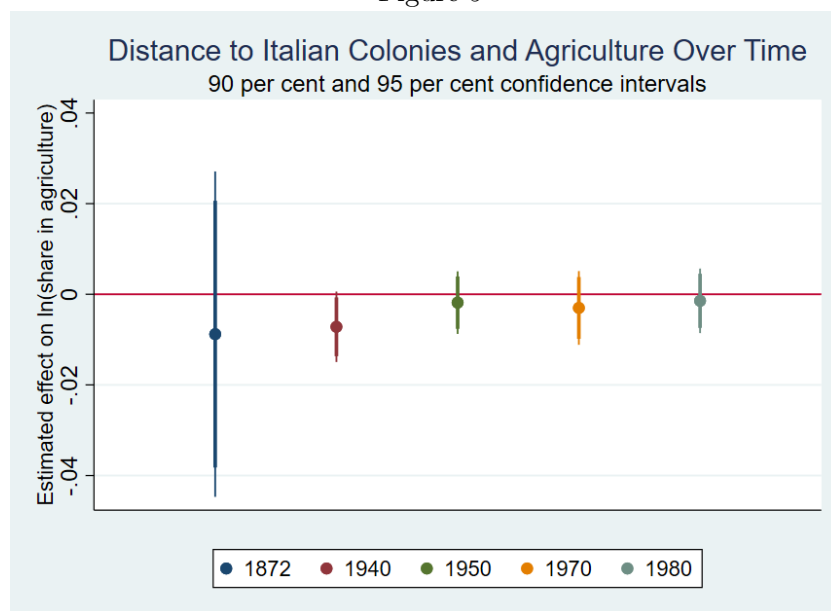Sources: See Data section.

Figure 8

Distance to Italian Colonies and Manufacturing Over Time
90 per cent and 95 per cent confidence intervals

● 1872 ● 1940 ● 1950 ● 1970 ● 1980

(a) Year-specific regression results with confidence intervals. The dependent variable is distance to São Leopoldo (10 km) and the regressions control for region fixed effects, ln population, distance to Porto Alegre (10 km), elevation, terrain, altitude, longitude, distance to the nearest coastline (10 km) and distance to the nearest river (10 km).

Figure 9

Distance to Italian Colonies and Agriculture Over Time
90 per cent and 95 per cent confidence intervals

● 1872 ● 1940 ● 1950 ● 1970 ● 1980

(a) Year-specific regression results with confidence intervals. The dependent variable is distance to São Leopoldo (10 km) and the regressions control for region fixed effects, ln population, distance to Porto Alegre (10 km), elevation, terrain, altitude, longitude, distance to the nearest coastline (10 km) and distance to the nearest river (10 km).