

Online Appendix for  
**The Effects of Immigration on Agricultural Development  
Brazil in the Age of Mass Migration**

David Escamilla-Guerrero  
University of St. Andrews & IZA

Andrea Papadia  
University of York

Ariell Zimran  
Vanderbilt University & NBER

January 2026

A. Additional Tables and Figures.....	59
B. Results with Conley Standard Errors.....	80
C. Instrument Based on Distance to Rail.....	85
D. Instrumental Variables Estimates with State-Level Heterogeneity.....	88
E. Two-Stage Least Squares Estimates.....	92
F. Additional Results for Farm Value Components.....	95
G. Additional Mechanisms.....	97

## A Additional Tables and Figures



Figure A.1: Geographic divisions of Brazil

Source: Instituto Brasileiro de Geografia e Estatística (2011)

Note: The map displays the political division of Brazil (regions, states, and municipalities) as it was in 1920 and the location of state capital cities.

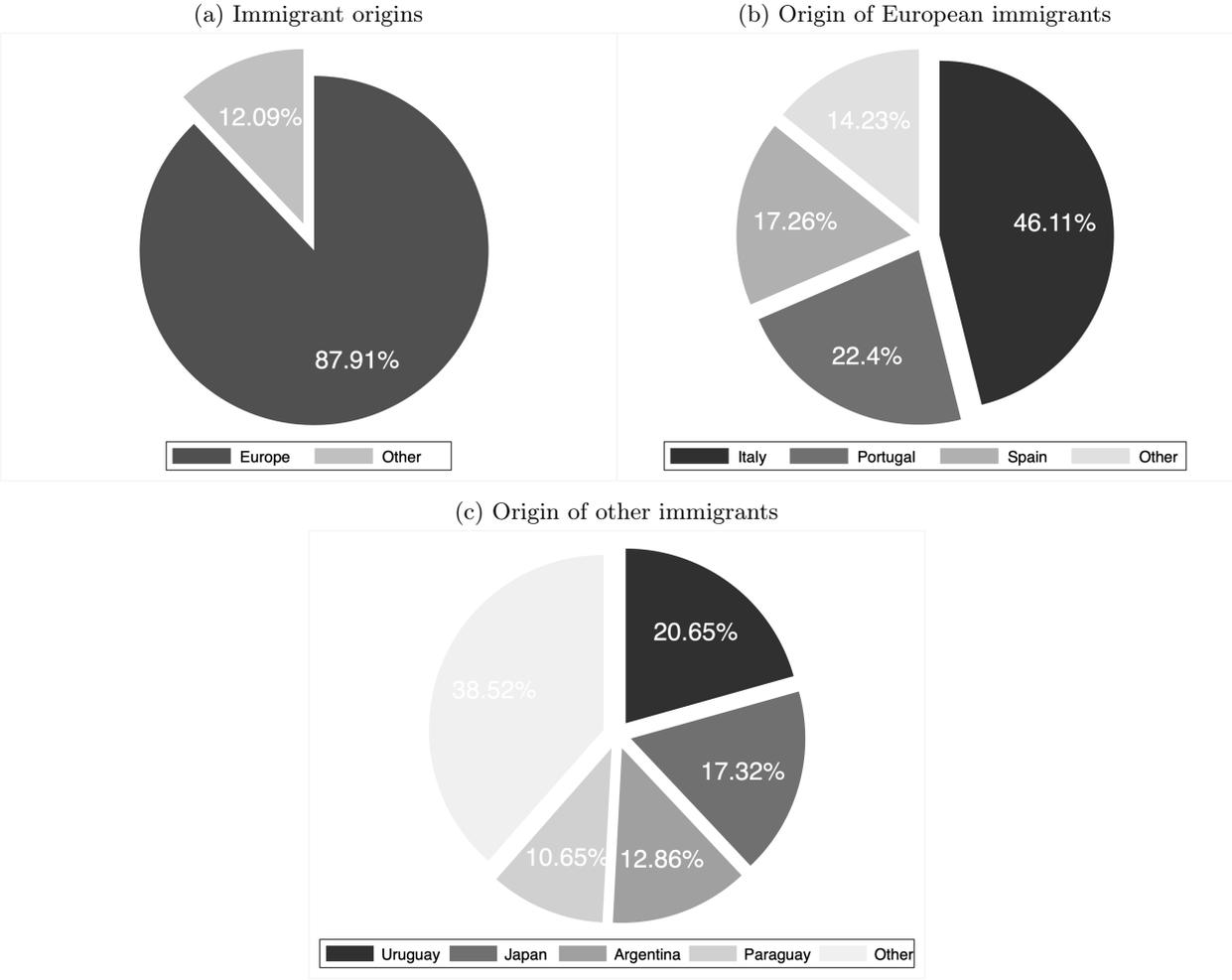


Figure A.2: Immigrant composition in 1920

Source: Directoria Geral de Estatistica (1922)

Note: The figure shows the share of immigrant population coming from each country or region of origin.

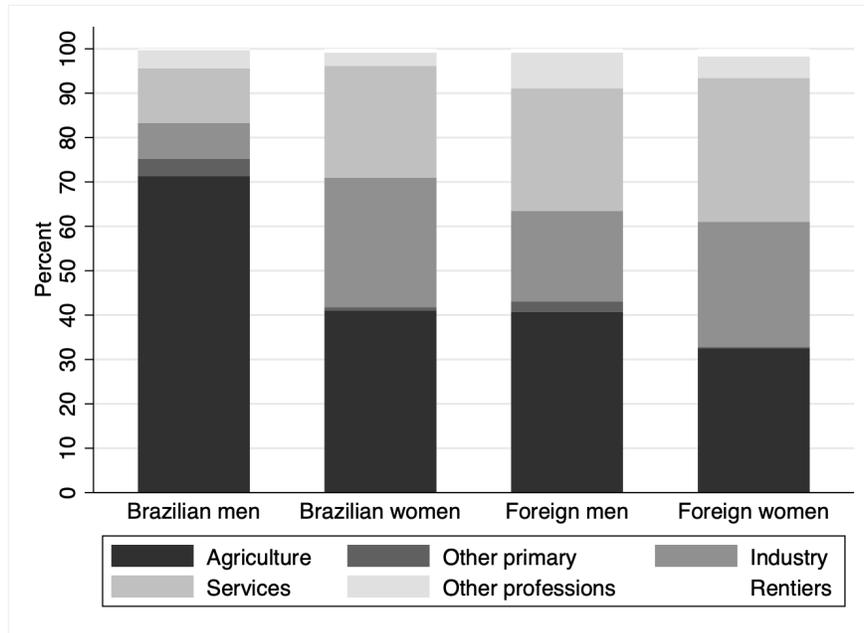


Figure A.3: Occupations of natives and immigrants

Source: Directoria Geral de Estatística (1922)

Note: The figure shows the population share of immigrants and natives broken down by occupation.

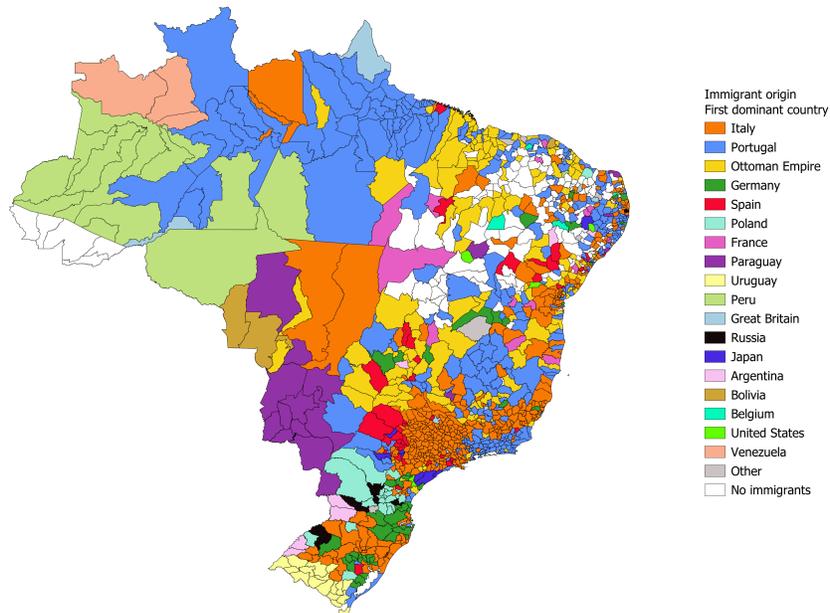


Figure A.4: Municipalities by largest immigrant group

Source: Directoria Geral de Estatística (1922)

Note: The map shows the largest immigrant-origin group for each municipality.

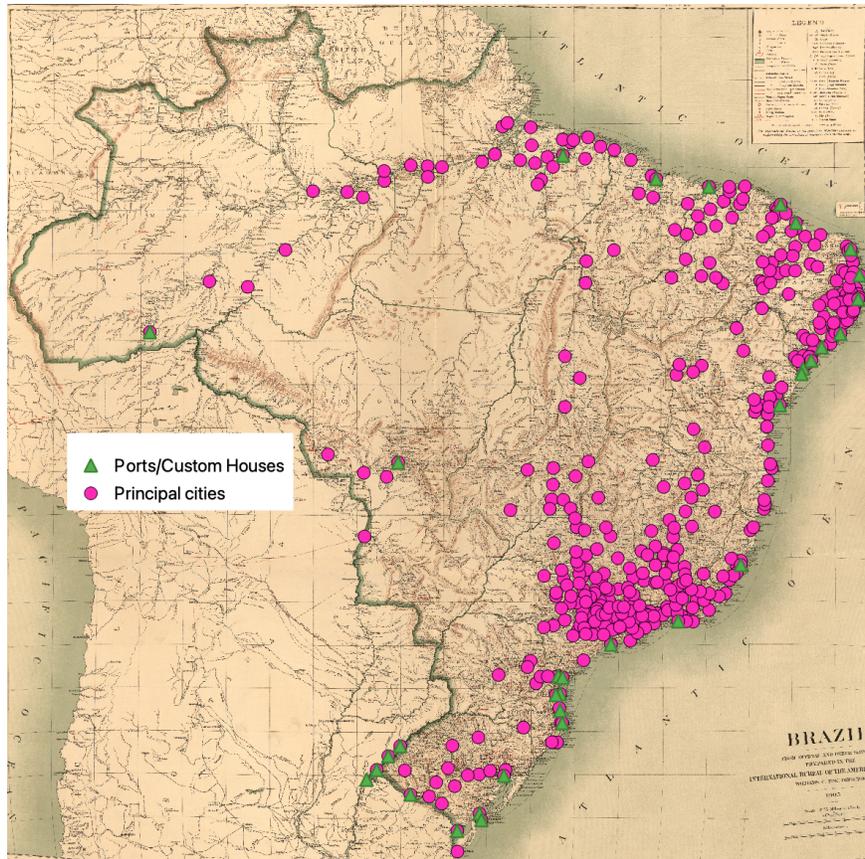


Figure A.5: Ports, custom houses, and principal cities in 1905

Source: International Bureau of the American Republics (1905).

Note: The map displays the location of principal cities, ports, and custom houses in Brazil in 1905.

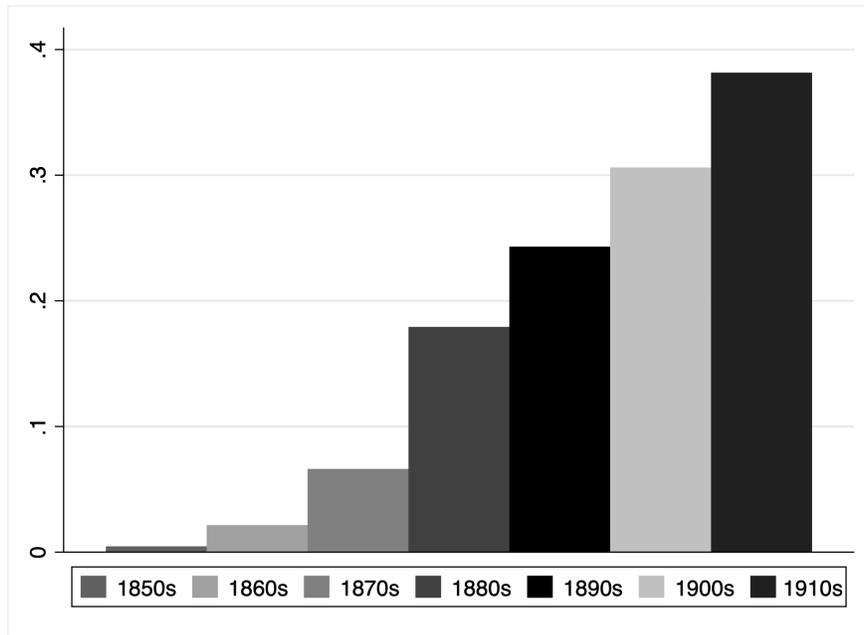


Figure A.6: Rollout of the rail network over time

Source: Giesbrecht (2023).

Note: The figure shows the share of municipalities connected to the rail network by decade.

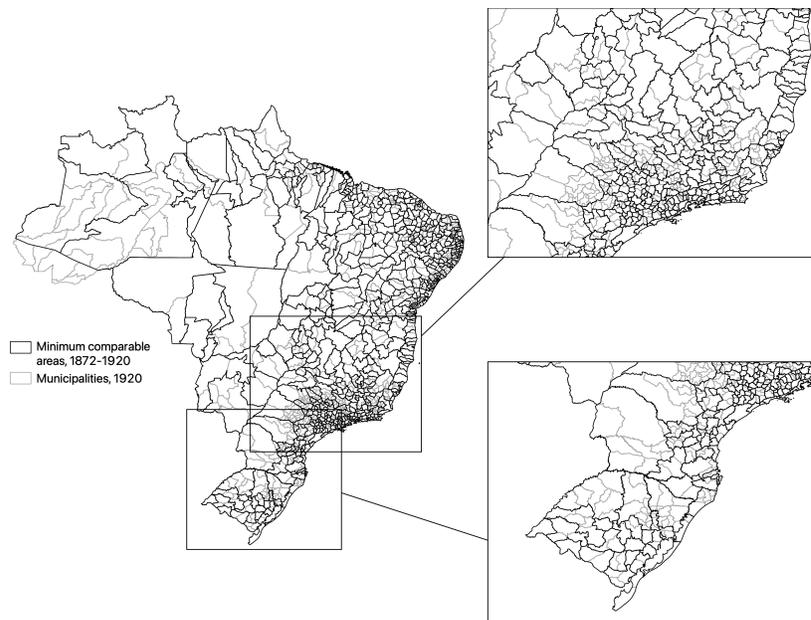


Figure A.7: Municipalities and minimum comparable areas

Source: Ehrl (2017) and Instituto Brasileiro de Geografia e Estatística (2011, 2023)

Note: The minimum comparable areas are created to maintain comparability between the 1920 municipalities on which our analysis is based and data from earlier censuses used in our balance tests.

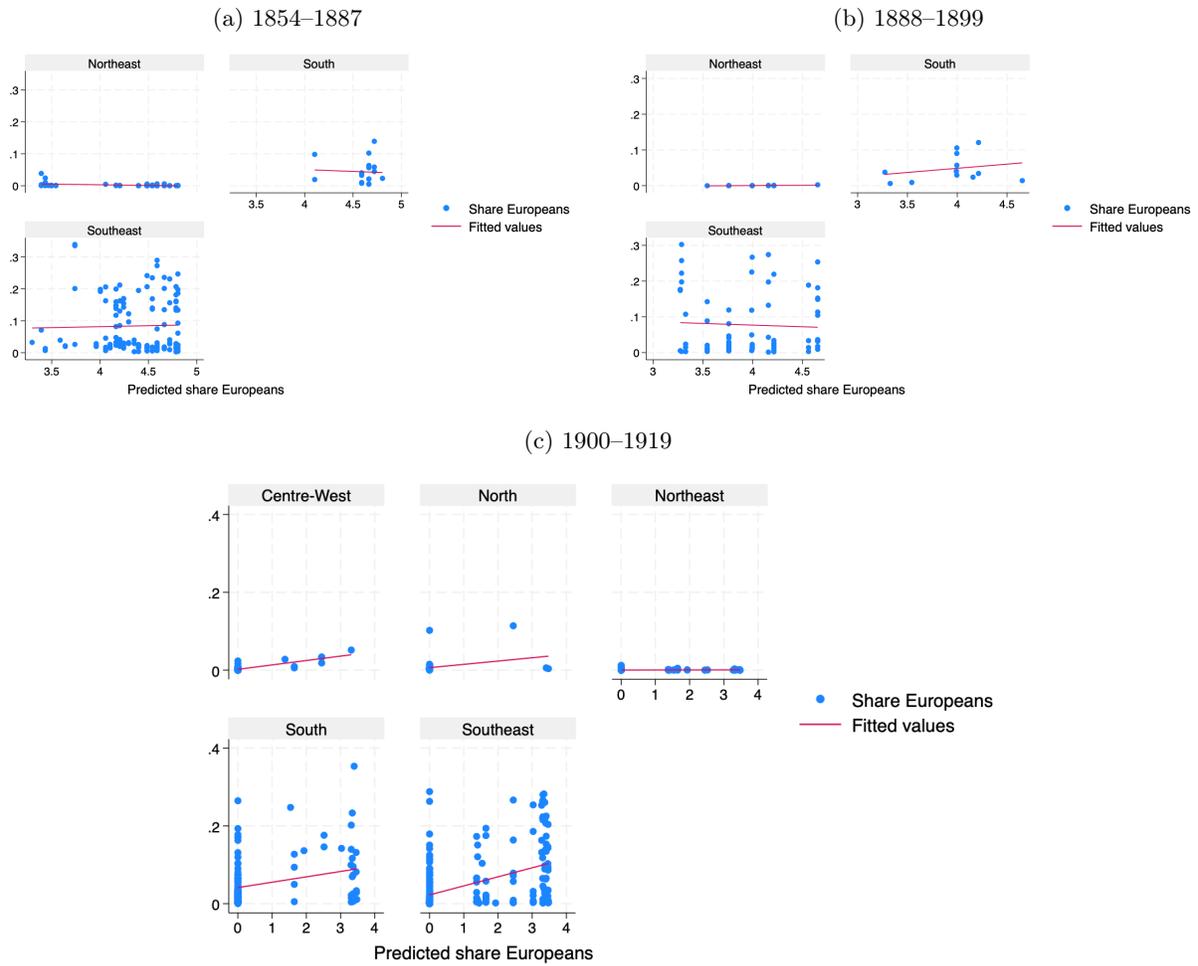


Figure A.8: Instrument and instrumented variable correlation, by period of rail linkage and region

*Note:* This figure presents the correlation between the value of the instrument and the actual share of European immigrants for various time periods and regions.

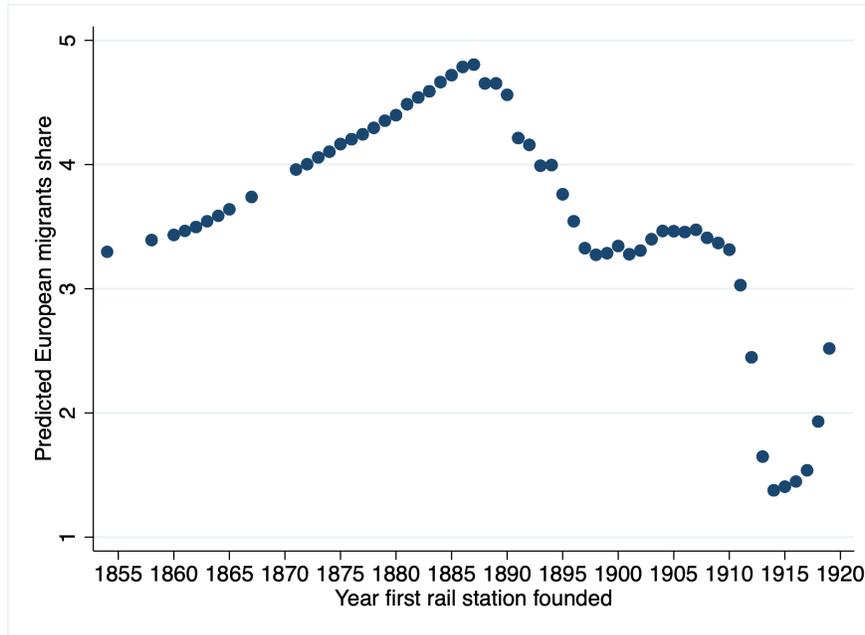


Figure A.9: Predicted European migrants by year of rail connection

*Note:* figure shows the predicted number of European immigrants (i.e., the value of the instrument) by the year in which a municipality was first connected to the rail network.

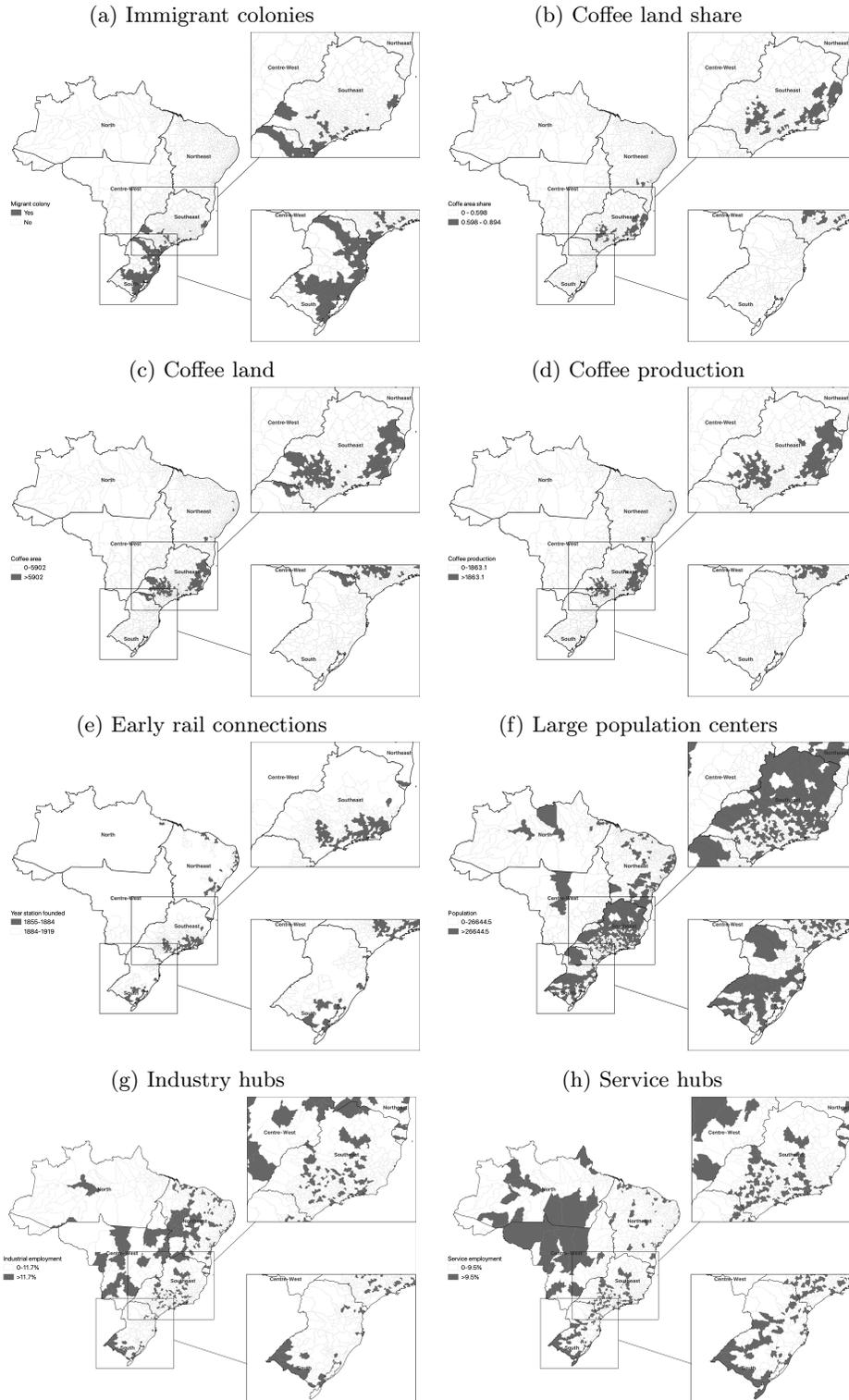


Figure A.10: Municipalities dropped in robustness checks

Note: Each map displays the municipalities excluded from the analysis as robustness checks (see Table 3).

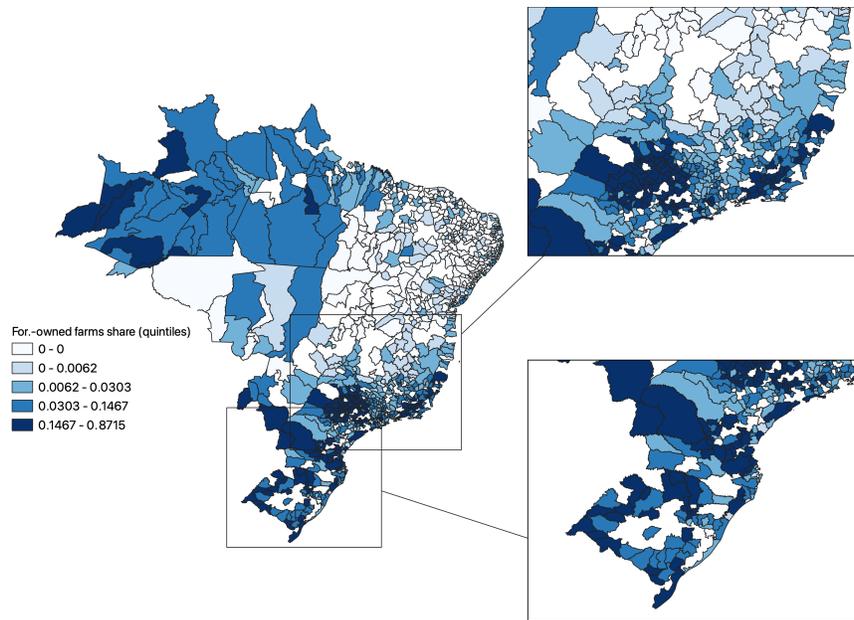


Figure A.11: Share of foreign-owned farms by municipality

*Source:* Directoria Geral de Estatística (1922)

*Note:* The maps display the spatial distribution of the share of foreign-owned farms. The zoomed-in areas illustrate the variation in these variables within and across regions.

Table A.1: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
<i>Panel A: Municipality sample</i>					
<i>Outcome variables</i>					
Farm value per hectare	1290	110.896	125.109	2.081	1148.963
Land value per hectare	1290	82.383	95.254	1.565	895.033
Infrastructure value per hectare	1290	23.989	30.644	.085	317.585
Tools value per hectare	1290	4.47	7.372	0	78.964
Literacy	1289	0.204	0.093	0.039	0.619
Brazilian Literacy	1289	0.198	0.089	0.039	0.613
Female Literacy	1289	0.157	0.09	0.027	0.606
Child Literacy	1289	0.161	0.103	0.019	0.653
Employed in Agriculture (share)	1289	.754	0.155	0.056	0.979
Employed in Industry (share)	1289	0.094	0.074	0.003	0.525
Rentiers (share)	1289	0.001	0.002	0	0.04
Employed in public admin. (share)	1289	0.002	0.003	0	0.042
Sex ratio population	1289	0.984	0.088	0.435	1.309
Sex ratio employed	1289	0.168	0.124	0.01	1.014
Sex ratio agriculture	1289	0.105	0.112	0.004	0.79
Sex ratio industry	1289	1.002	1.722	0	17.128
<i>Main explanatory variables</i>					
Share Europeans	1289	0.029	0.059	0	0.353
Foreign-owned farms	1290	0.086	0.147	0	0.872
<i>Mechanisms</i>					
Population density	1289	22.76	48.392	.02	1086.668
Agr. employment density	1290	4.438	5.365	0	55.507
Cultivated share of farm	1289	0.083	0.108	0	0.894
Coffee cultivated share	1285	0.157	0.238	0	0.894
Cash crops cultivated share	1285	0.261	0.281	0	1
Share farms with tools	1290	0.097	0.176	0	1.087
Share farms with machines	1290	0.22	0.218	0	1
<i>Controls</i>					
Railway years	1290	10.451	16.213	0	66
No rail	1290	0.617	0.486	0	1
Ruggedness	1290	103.203	70.809	4.472	482.708
Altitude	1290	445.324	314.757	5.098	1376.206
Area sq km.)	1290	4767.321	10183.36	26.067	94488.34
Latitude	1290	-15.37	8.35	-33.18	2.26
Longitude	1290	-44.585	5.93	-71.594	-34.845
Lat×Long	1290	704.994	430.851	-127.982	1759.234
Land adaptability	1290	.015	6.386	-23.102	9.502
Land quality 1	1290	-.003	1.921	-6.27	3.637
Land quality 2	1290	.002	1.855	-5.008	3.081
Port/cust. house dist. (int. markt. acc.)	1290	251.473	218.826	4.281	1460.116
Sq. of port/cust. house dist.	1290	111086.2	200183.3	18.331	2131938
City dist. (dom. markt. acc.)	1290	55.289	74.003	0.341	855.185
Sq. of city dist.	1290	8529.006	35742.93	.116	731340.8
<i>Panel B: Farm-level sample</i>					
Land value per hectare	40693	245.333	269.56	1.21	7260
Farm value per hectare	40242	1566.858	17519.86	.363	1936000
Foreign workers share	40693	0.248	0.391	0	1
Foreign owner	40693	0.17	0.375	0	1

*Note:* The table shows sample statistics for the two samples used in the study.  
Source: Directoria Geral de Estatística (1922) and Secreteria da Agricultura (1906–1910).

Table A.2: Summary statistics for selected variables by region

Variable	Brazil	North	Northeast	Center-West	Southeast	South
<i>Panel A: Outcome variables</i>						
Farm value per hectare	110.94	30.056	78.34	12.904	180.449	101.337
N	1289	81	526	70	460	153
SD	125.148	26.388	87.255	9.213	158.866	59.291
Land value per hectare	82.426	19.414	54.711	10.103	137.741	77.489
N	1289	526	70	460	153	
SD	95.278	15.424	63.832	7.187	121.399	41.838
Infrastructure value per hectare	23.987	9.791	20.019	2.616	35.58	20.083
N	1289	81	526	70	460	153
SD	30.655	12.633	25.355	2.735	38.537	19.403
Tools value per hectare	4.474	0.85	3.479	0.185	7.128	3.765
N	1289	81	526	70	460	153
SD	7.374	1.809	6.99	0.303	8.571	5.268
Literacy	0.204	0.241	0.153	0.19	0.225	0.303
N	1289	81	525	70	460	153
SD	0.093	0.067	0.066	0.079	0.08	0.111
Employed in Agriculture (share)	0.754	0.77	0.759	0.696	0.766	0.725
N	1288	81	524	70	460	153
SD	0.155	0.152	0.158	0.167	0.14	0.177
Employed in Industry (share)	0.094	0.056	0.102	0.125	0.09	0.084
N	1288	81	524	70	460	153
SD	0.074	0.037	0.082	0.104	0.063	0.062
<i>Panel B: Main explanatory variables</i>						
Share Europeans	0.029	0.008	0.001	0.005	0.061	0.051
N	1289	81	525	70	460	153
SD	0.059	0.018	0.002	0.009	0.079	0.061
Foreign-owned farms	0.086	0.074	0.006	0.019	0.153	0.194
N	1289	81	526	70	460	153
SD	0.147	0.073	0.014	0.046	0.173	0.197
<i>Panel C: Rail controls</i>						
Railway years	10.453	1.173	7.414	0.771	17.585	8.732
N	1289	81	526	70	460	153
SD	16.219	5.289	15.504	2.36	17.577	13.089
No rail	0.618	0.938	0.753	0.9	0.372	0.595
N	1289	81	526	70	460	153
SD	0.486	0.242	0.432	0.302	0.484	0.493
<i>Panel D: Mechanisms</i>						
Population density	22.776	3.425	25.317	1.536	29.186	14.753
N	1289	81	525	70	460	153
SD	48.404	7.174	59.509	1.763	46.545	21.345
Agricultural employment density	4.445	0.814	4.917	0.345	5.662	2.932
N	1288	81	526	70	460	153
SD	5.374	1.422	6.792	0.428	4.035	3.569
Cultivated share of farm	0.083	0.024	0.056	0.006	0.141	0.074
N	1289	81	525	70	460	153
SD	0.108	0.051	0.075	0.01	0.132	0.09
Coffee cultivated share	0.157	0.033	0.057	0.055	0.347	0.036
N	1285	80	522	70	460	153
SD	0.238	0.059	0.139	0.085	0.271	0.109
Cash crops cultivated share	0.261	0.389	0.451	0.095	0.11	0.079
N	1285	80	522	70	460	153
SD	0.281	0.283	0.303	0.07	0.119	0.142
Share farms with tools	0.097	0.006	0.026	0.018	0.146	0.279
N	1289	81	526	70	460	153
SD	0.176	0.019	0.094	0.041	0.174	0.273
Share farms with machines	0.22	0.089	0.195	0.193	0.293	0.166
N	1289	81	526	70	460	153
SD	0.218	0.145	0.212	0.188	0.209	0.238

*Note:* The table shows sample statistics by region for the main sample used in the study.  
Source: Directoria Geral de Estatística (1922).

Table A.3: Balance test for the 1920 European migrants share

	Share Europeans 1920		
	(1)	(2)	(2)
Population	-0.108*** (0.0345)	-0.116*** (0.0356)	0.0626** (0.0308)
Share Europeans 1872	0.0903** (0.0429)	0.129*** (0.0443)	0.178*** (0.0419)
Employed agriculture 1872	0.00910 (0.0279)	0.00230 (0.0291)	-0.110*** (0.0328)
Justice work. 1872	0.0665* (0.0360)	0.0673** (0.0319)	0.0256 (0.0245)
Slave share 1872	0.115*** (0.0427)	0.0101 (0.0472)	0.141*** (0.0486)
White share 1872	0.227*** (0.0438)	0.160*** (0.0427)	-0.00230 (0.0451)
Literacy 1872	0.0962* (0.0550)	0.127** (0.0565)	0.104* (0.0536)
School attendance 1872	-0.108** (0.0442)	-0.105** (0.0422)	-0.0850** (0.0365)
Ruggedness	-0.191*** (0.0643)	-0.210*** (0.0611)	-0.148** (0.0667)
Altitude	0.246*** (0.0582)	0.115* (0.0662)	0.100 (0.0873)
Latitude	-0.261*** (0.0422)	-0.232*** (0.0708)	0.201 (0.134)
Longitude	-0.198*** (0.0301)	-0.320*** (0.0601)	-0.477*** (0.120)
Land quality	0.214*** (0.0450)	0.131*** (0.0456)	0.245*** (0.0495)
Distance (port/custom house)	-0.133*** (0.0338)	-0.103** (0.0428)	-0.00888 (0.0391)
Distance (city)	-0.0321 (0.0330)	-0.0365 (0.0341)	-0.0228 (0.0300)
Observations	604	604	604
$R^2$	0.336	0.378	0.556
Fixed effects	No	Region	State

*Note:* \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses. The unit of observation is a minimum comparable area between 1872 and 1920. The table shows the relationship between immigrant settlement in 1920 and sociodemographic characteristics observed in 1872, as well as the relationship with geographic characteristics. Unconditional regressions are reported in column (1). Regressions conditional on region fixed effects are reported in column (2). Regressions conditional on state fixed effects are reported in column (3).

Table A.4: European immigration and farm values with alternative controls for years of rail linkage

	Farm Value per hectare					Land	Infrastructure	Tools & Machines
	(1)	(2)	(3)	(4)		(5)	(6)	(7)
<b>Panel A: Railway years exponentiated</b>								
Share Europeans	0.700*** (0.155)	0.678*** (0.117)	0.742*** (0.118)	0.745*** (0.119)	(0.120)	0.736*** (0.140)	0.577*** (0.124)	0.752***
Observations	1,289	1,289	1,289	1,289		1,289	1,289	1,289
1 <sup>st</sup> stage F-stat	16.77	22.20	22.60	22.27		22.27	22.27	22.27
<b>Panel B: Railway years squared</b>								
Share Europeans	0.667*** (0.138)	0.670*** (0.111)	0.730*** (0.109)	0.724*** (0.118)		0.713*** (0.119)	0.583*** (0.143)	0.705*** (0.142)
Observations	1,289	1,289	1,289	1,289		1,289	1,289	1,289
1 <sup>st</sup> stage F-stat	17.29	22.70	23.12	22.34		22.34	22.34	22.34
<b>Panel C: Linkage decade FE</b>								
Share Europeans	0.640*** (0.135)	0.653*** (0.112)	0.713*** (0.111)	0.701*** (0.116)		0.688*** (0.116)	0.570*** (0.139)	0.645*** (0.125)
Observations	1,289	1,289	1,289	1,289		1,289	1,289	1,289
1 <sup>st</sup> stage F-stat	14.74	19.61	20.13	24.64		24.64	24.64	24.64
Railway years	✓	✓	✓	✓		✓	✓	✓
No rail	✓	✓	✓	✓		✓	✓	✓
Geo Controls	✓	✓	✓	✓		✓	✓	✓
Region FE	✓	✓	✓					
Land adaptability		✓	✓	✓		✓	✓	✓
Land quality		✓	✓	✓		✓	✓	✓
Dom market access			✓	✓		✓	✓	✓
Int market access			✓	✓		✓	✓	✓
State FE				✓		✓	✓	✓

Note: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. The IV regressions are estimated using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. Regressions in Panel A control for years with rail connection in exponential form. Regressions in Panel B control for years with rail connection in quadratic form. Regressions in Panel C control for decade of rail connection fixed effects. *Railway years* is the number of years with railway connection. *No rail* is an indicator variable for municipalities with no railway connection. *Geo controls* include surface area, ruggedness, altitude, latitude, and longitude. *Land adaptability* is the Herfindahl-Hirschman Index (HHI) of crop suitability. *Land quality* includes the first two principal components of suitability for all major crops reported in the census. *Dom market access* includes the linear and quadratic distance to the nearest city and principal town. *Int market access* includes the linear and quadratic distance to the nearest port or custom house. All continuous variables are standardized to have mean zero and standard deviation one.

Table A.5: European immigration and agricultural production value

	Agricultural Production Value							
	Per hectare				Per cultivated hectare			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A: OLS</i>								
Share Europeans	0.301*** (0.0606)	0.306*** (0.0620)	0.301*** (0.0613)	0.312*** (0.0689)	0.0563 (0.0554)	0.0368 (0.0433)	0.0293 (0.0388)	0.0148 (0.0183)
Observations	1,287	1,287	1,287	1,287	1,285	1,285	1,285	1,285
$R^2$	0.183	0.215	0.220	0.236	0.030	0.043	0.046	0.055
<i>Panel B: IV</i>								
Share Europeans	0.289 (0.185)	0.352** (0.153)	0.360** (0.148)	0.428** (0.199)	-0.225** (0.101)	-0.123* (0.0746)	-0.116 (0.0777)	-0.113 (0.104)
Observations	1,287	1,287	1,287	1,287	1,285	1,285	1,285	1,285
1 <sup>st</sup> stage F-stat	16.88	22.25	22.62	22.30	16.89	22.25	22.62	22.28
Railway years	✓	✓	✓	✓	✓	✓	✓	✓
No rail	✓	✓	✓	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓		✓	✓	✓	
Land adaptability		✓	✓	✓		✓	✓	✓
Land quality		✓	✓	✓		✓	✓	✓
Dom market access			✓	✓			✓	✓
Int market access			✓	✓			✓	✓
State FE				✓				✓

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. The IV regressions are estimated using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *Railway years* is the number of years with railway connection. *No rail* is an indicator variable for municipalities with no railway connection. *Geo controls* include surface area, ruggedness, altitude, latitude, and longitude. *Land adaptability* is the Herfindahl-Hirschman Index (HHI) of crop suitability. *Land quality* includes the first two principal components of suitability for all major crops reported in the census. *Dom market access* includes the linear and quadratic distance to the nearest city and principal town. *Int market access* includes the linear and quadratic distance to the nearest port or custom house. All continuous variables are standardized to have mean zero and standard deviation one.

Table A.6: Randomization test for instrument

Test	Significant	Total Simulations	Actual Rate (%)
First Stage	65	1000	6.5
Reduced Form	54	1000	5.4
IV Estimate	0	1000	0

*Note:* Results from a falsification test where railway connection years are randomly reassigned across treated municipalities. We repeat the randomization process one thousand times and compute the number of times the first stage, reduced form, and IV estimates are statistically significant. Under the null hypothesis of valid identification, significance rates should approximate 5%. All municipalities are included in the analysis. All regressions are estimated by instrumental variables using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. Control variables include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects.

Table A.7: European immigration and farm value, controlling for 1872 characteristics

	Farm Value per hectare				Land	Infrastructure	Tools & Machines
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Panel A: OLS</i>							
Share Europeans	0.600*** (0.0824)	0.600*** (0.0805)	0.601*** (0.0821)	0.566*** (0.0991)	0.532*** (0.100)	0.589*** (0.0818)	0.384*** (0.0967)
Observations	604	604	604	604	604	604	604
$R^2$	0.581	0.619	0.624	0.664	0.664	0.558	0.485
<i>Panel B: IV</i>							
Share Europeans	0.850*** (0.245)	0.884*** (0.174)	0.913*** (0.182)	0.844*** (0.153)	0.811*** (0.153)	0.818*** (0.157)	0.518*** (0.144)
Observations	604	604	604	604	604	604	604
1 <sup>st</sup> stage F-stat	8.963	9.773	9.658	9.871	9.871	9.871	9.871
1872 Controls	✓	✓	✓	✓	✓	✓	✓
Railway years	✓	✓	✓	✓	✓	✓	✓
No rail	✓	✓	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓				
Land adaptability		✓	✓	✓	✓	✓	✓
Land quality		✓	✓	✓	✓	✓	✓
Dom market access			✓	✓	✓	✓	✓
Int market access			✓	✓	✓	✓	✓
State FE				✓	✓	✓	✓

*Note:* \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses. The unit of observation is a minimum comparable area. All areas are included in the analysis. Minimum comparable areas link 1872 and 1920 census data, given the creation and suppression of municipalities over time. Figure A.7 displays minimum comparable areas across the Brazilian. The IV regressions are estimated using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *1872 controls* capture early migrant settlement (population share of European immigrants), racial composition (population share of whites), human capital (literacy rate and share of children attending school), state capacity (justice workers per 1,000 inhabitants), and the presence of slavery (population share of slaves). *Railway years* is the number of years with railway connection. *No rail* is an indicator variable for municipalities with no railway connection. *Geo controls* include surface area, ruggedness, altitude, latitude, and longitude. *Land adaptability* is the Herfindahl-Hirschman Index (HHI) of crop suitability. *Land quality* includes the first two principal components of suitability for all major crops reported in the census. *Dom market access* includes the linear and quadratic distance to the nearest city and principal town. *Int market access* includes the linear and quadratic distance to the nearest port or custom house. All continuous variables are standardized to have mean zero and standard deviation one.

Table A.8: European immigration and farm value, heterogeneity by immigrant nationality (OLS estimates)

	Farm Value per hectare				Land	Infrastructure	Tools & Machines
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share Italians	0.441*** (0.0513)	0.440*** (0.0513)	0.434*** (0.0499)	0.451*** (0.0544)	0.424*** (0.0566)	0.421*** (0.0480)	0.433*** (0.0608)
Share Spanish	0.189*** (0.0696)	0.192*** (0.0698)	0.200*** (0.0685)	0.226*** (0.0722)	0.233*** (0.0716)	0.183*** (0.0686)	0.0573 (0.0607)
Share Portuguese	0.0920 (0.0634)	0.100 (0.0638)	0.0929 (0.0622)	0.0981 (0.0648)	0.0777 (0.0594)	0.154* (0.0900)	0.0228 (0.0445)
Share Others	0.666** (0.327)	0.430 (0.334)	0.552* (0.285)	0.589** (0.291)	0.421 (0.299)	0.819*** (0.311)	1.080 (0.683)
Observations	1,289	1,289	1,289	1,289	1,289	1,289	1,289
$R^2$	0.581	0.599	0.614	0.635	0.644	0.491	0.433
Railway years	✓	✓	✓	✓	✓	✓	✓
No rail	✓	✓	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓				
Land adaptability		✓	✓	✓	✓	✓	✓
Land quality		✓	✓	✓	✓	✓	✓
Dom market access			✓	✓	✓	✓	✓
Int market access			✓	✓	✓	✓	✓
State FE				✓	✓	✓	✓

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. Regressions are estimated by ordinary least squares (OLS). *Railway years* is the number of years with railway connection. *No rail* is an indicator variable for municipalities with no railway connection. *Geo controls* include surface area, ruggedness, altitude, latitude, and longitude. *Land adaptability* is the Herfindahl-Hirschman Index (HHI) of crop suitability. *Land quality* includes the first two principal components of suitability for all major crops reported in the census. *Dom market access* includes the linear and quadratic distance to the nearest city and principal town. *Int market access* includes the linear and quadratic distance to the nearest port or custom house. All continuous variables are standardized to have mean zero and standard deviation one.

Table A.9: Total immigration and farm values

	Farm Value per hectare				Land	Infrastructure	Tools & Machines
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Panel A: OLS</i>							
Share Immigrants	0.534*** (0.0751)	0.542*** (0.0780)	0.542*** (0.0740)	0.528*** (0.0777)	0.497*** (0.0760)	0.520*** (0.0761)	0.382*** (0.0576)
Observations	1,289	1,289	1,289	1,289	1,289	1,289	1,289
$R^2$	0.496	0.511	0.531	0.561	0.575	0.429	0.389
<i>Panel B: IV</i>							
Share Immigrants	0.826*** (0.185)	0.778*** (0.137)	0.825*** (0.134)	0.827*** (0.124)	0.765*** (0.122)	0.626*** (0.147)	0.709*** (0.134)
Observations	1,289	1,289	1,289	1,289	1,289	1,289	1,289
1 <sup>st</sup> stage F-stat	16.08	21.26	22.95	21.65	24.07	24.07	24.07
Railway years	✓	✓	✓	✓	✓	✓	✓
No rail	✓	✓	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓				
Land adaptability		✓	✓	✓	✓	✓	✓
Land quality		✓	✓	✓	✓	✓	✓
Dom market access			✓	✓	✓	✓	✓
Int market access			✓	✓	✓	✓	✓
State FE				✓	✓	✓	✓

*Note:* \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. The IV regressions are estimated using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of immigrants interacted with region indicators, which instrument for the actual share of immigrants. *Railway years* is the number of years with railway connection. *No rail* is an indicator variable for municipalities with no railway connection. *Geo controls* include surface area, ruggedness, altitude, latitude, and longitude. *Land adaptability* is the Herfindahl-Hirschman Index (HHI) of crop suitability. *Land quality* includes the first two principal components of suitability for all major crops reported in the census. *Dom market access* includes the linear and quadratic distance to the nearest city and principal town. *Int market access* includes the linear and quadratic distance to the nearest port or custom house. All continuous variables are standardized to have mean zero and standard deviation one.

Table A.10: Immigration, foreign ownership, farm size, and inequality

	Average farm size (1)	Land gini (2)
Share Europeans	-0.219** (0.0909)	-0.0501 (0.0590)
Foreign-owned farms	0.168** (0.0850)	0.0904 (0.0599)
Observations	1,289	1,236
$R^2$	0.463	0.115
All controls	✓	✓

*Note:* Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. Regressions are estimated by ordinary least squares (OLS). *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

Table A.11: Immigration, foreign ownership, farm values, and productivity

	(1)	(2)	(3)	(4)
	Farm value			Prod. value
	Land	Infrastructure	Tools & Machines	
Share Europeans	0.829** (0.403)	0.730** (0.358)	1.137*** (0.335)	0.826* (0.427)
Foreign-owned farms	-0.284 (0.300)	-0.139 (0.276)	-0.552** (0.255)	-0.410 (0.306)
Observations	1,289	1,289	1,289	1,287
First-stage F	9.705	9.705	9.705	9.723
All controls	✓	✓	✓	✓

*Note:* Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. Regressions are estimated using the limited information maximum likelihood (LIML) estimator. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one. Production value is per hectare.

Table A.12: Immigration, foreign ownership, and land use

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<b>Panel A: Share of total farm land</b>												
		<b>Cultivated</b>			<b>Coffee area</b>			<b>Cash crops area</b>			<b>Grain, food crops area</b>	
Share Europeans	0.464*** (0.0646)		0.480*** (0.104)	0.486*** (0.0652)		0.558*** (0.115)	0.0331 (0.0298)		0.0793* (0.0473)	0.323*** (0.0591)		0.251*** (0.0837)
Foreign-owned farms		0.334*** (0.0453)	-0.0181 (0.0744)		0.326*** (0.0456)	-0.0833 (0.0842)		0.00451 (0.0217)	-0.0536 (0.0337)		0.267*** (0.0522)	0.0834 (0.0756)
Observations	1,287	1,287	1,287	1,287	1,287	1,287	1,285	1,285	1,285	1,287	1,287	1,287
R <sup>2</sup>	0.398	0.363	0.399	0.441	0.394	0.442	0.230	0.229	0.230	0.305	0.296	0.306
<b>Panel B: Share of cultivated land</b>												
		<b>Coffee area</b>			<b>Cash crops area</b>			<b>Grain, food crops area</b>				
Share Europeans	0.258*** (0.0441)		0.250*** (0.0786)	-0.0466* (0.0240)		-0.0290 (0.0428)	-0.0480*** (0.00921)		-0.0506*** (0.0162)			
Foreign-owned farms		0.193*** (0.0316)	0.00948 (0.0600)		-0.0416** (0.0180)	-0.0204 (0.0338)		-0.0341*** (0.00800)	0.00304 (0.0142)			
Observations	1,285	1,285	1,285	1,285	1,285	1,285	1,285	1,285	1,285			
R <sup>2</sup>	0.550	0.540	0.550	0.565	0.565	0.565	0.500	0.495	0.500			
All controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. Regressions are estimated by ordinary least squares (OLS). *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

Table A.13: Implications for structural change controlling for agricultural development

<i>Panel A: Human capital formation (literacy)</i>								
	Population		Brazilians		Females		Children	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share Europeans	0.827*** (0.183)	1.188*** (0.368)	0.780*** (0.182)	1.110*** (0.357)	0.758*** (0.190)	1.158*** (0.405)	0.766*** (0.164)	1.017*** (0.315)
Farm value		-0.298 (0.189)		-0.269 (0.185)		-0.281 (0.205)		-0.199 (0.164)
Observations	1,289	1,289	1,289	1,289	1,289	1,289	1,289	1,289
1 <sup>st</sup> stage F-stat	22.31	15.05	22.31	15.05	22.31	15.05	22.31	15.05
<i>Panel B: Economic structure</i>								
	Employed Agriculture		Employed Industry		Employed Services			
Share Europeans	-0.432** (0.188)	-0.551* (0.333)	0.288 (0.194)	0.241 (0.396)	0.238 (0.228)	0.394 (0.451)		
Farm Value		0.0553 (0.173)		0.115 (0.198)		-0.0167 (0.233)		
Observations	1,288	1,288	1,288	1,288	1,288	1,288		
First-stage F (KP rk Wald)	22.34	15.13	22.34	15.13	22.34	15.13		
<i>Panel C: Specialization (probability of being a hub)</i>								
	Agriculture hub		Industry hub		Service hub			
Share Europeans	-0.190*** (0.0666)	-0.210** (0.0968)	0.213** (0.0828)	0.294** (0.139)	0.308*** (0.0736)	0.447*** (0.112)		
Farm Value		0.0424 (0.0467)		-0.0648 (0.0678)		-0.169*** (0.0581)		
Observations	1,289	1,289	1,289	1,289	1,289	1,289		
1 <sup>st</sup> stage F-stat	22.31	15.05	22.31	15.05	22.31	15.05		
<i>Panel D: Political Economy</i>								
	Rentiers		Pub. admin					
Share Europeans	-0.461*** (0.131)	-0.586*** (0.208)	-0.204 (0.145)	-0.367* (0.223)	-0.256 (0.179)	-0.392 (0.312)	0.154 (0.105)	0.215 (0.167)
Farm Value		0.162 (0.104)		0.224** (0.106)		0.161 (0.148)		-0.0883 (0.0793)
Observations	1,289	1,289	1,288	1,288	1,288	1,288	1,288	1,288
1 <sup>st</sup> stage F-stat	22.31	15.05	22.34	15.13	22.34	15.13	22.34	15.13
All controls	✓	✓	✓	✓	✓	✓	✓	✓

Note: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated by instrumental variables using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

## B Results with Conley Standard Errors

Table B.1: European immigration and farm value, Conley standard errors

	Farm Value per hectare (1)	Land (2)	Infrastructure (3)	Tools & Machines (4)
Share Europeans	0.724***	0.705***	0.576***	0.735***
<i>Distance cutoff</i>				
50 km	( 0.127)	( 0.129)	( 0.135)	( 0.130)
100 km	( 0.115)	( 0.109)	( 0.141)	( 0.145)
200 km	( 0.125)	( 0.119)	( 0.142)	( 0.158)
300 km	( 0.114)	( 0.112)	( 0.114)	( 0.145)
Observations	1,289	1,289	1,289	1,289
1 <sup>st</sup> stage F-stat				
All controls	✓	✓	✓	✓

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Conley standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated by instrumental variables using the 2SLS estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

Table B.2: Potential mechanisms and European immigration, Conley standard errors

	Agricultural share of farms (1)	Cultivated share (2)	Coffee share (no coffee) (3)	Cash crops density (4)	Population emp. density (5)	Share farms with tools (6)	Share farms with machines (7)
Share Europeans	0.728***	0.823***	0.013	-0.173	-0.061	0.675***	0.046
<i>Distance cutoff</i>							
50 km	( 0.164)	( 0.165)	( 0.108)	( 0.152)	( 0.173)	( 0.115)	( 0.177)
100 km	( 0.160)	( 0.208)	( 0.118)	( 0.170)	( 0.181)	( 0.098)	( 0.203)
200 km	( 0.128)	( 0.221)	( 0.107)	( 0.198)	( 0.204)	( 0.089)	( 0.221)
300 km	( 0.116)	( 0.255)	( 0.111)	( 0.212)	( 0.212)	( 0.099)	( 0.257)
Observations	1,287	1,285	1,285	1,296	1,296	1,289	1,289
All controls	✓	✓	✓	✓	✓	✓	✓

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Conley standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated by instrumental variables using the 2SLS estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

Table B.3: Farm value and potential mechanisms, Conley standard errors

	<b>Farm Value per hectare</b>					
	(1)	(2)	(3)	(4)	(5)	(6)
Share Europeans	0.730***	0.484***	0.547***	0.750***	0.599***	0.281*/**
<b><i>Distance cutoff</i></b>						
50 km	(0.127)	(0.110)	(0.136)	(0.120)	(0.145)	(0.129)
100 km	(0.114)	(0.124)	(0.141)	(0.111)	(0.155)	(0.155)
200 km	(0.124)	(0.124)	(0.152)	(0.117)	(0.147)	(0.124)
300 km	(0.113)	(0.104)	(0.148)	(0.0925)	(0.137)	(0.122)
			<b><i>Panel B: Land use</i></b>			
Cultivated share of farms		0.405***				0.410***
<b><i>Distance cutoff</i></b>						
50 km		(0.0814)				(0.0872)
100 km		(0.0907)				(0.0986)
200 km		(0.109)				(0.116)
300 km		(0.0960)				(0.110)
Coffee share			0.231***			0.0663/*
<b><i>Distance cutoff</i></b>						
50 km			(0.0487)			(0.0388)
100 km			(0.0504)			(0.0488)
200 km			(0.0636)			(0.0461)
300 km			(0.0699)			(0.0433)
Other cash crops share			0.130***			0.0927***
<b><i>Distance cutoff:</i></b>						
50 km			(0.0275)			(0.0214)
100 km			(0.0296)			(0.0257)
200 km			(0.0352)			(0.0268)
300 km			(0.0355)			(0.0213)
			<b><i>Panel A: Labor force</i></b>			
Pop density				0.135**		0.252***
<b><i>Distance cutoff</i></b>						
50 km				(0.0623)		(0.0716)
100 km				(0.0638)		(0.0735)
200 km				(0.0608)		(0.0732)
300 km				(0.0620)		(0.0790)
Agr. emp. density				0.155/**		-0.0413
<b><i>Distance cutoff</i></b>						
50 km				(0.0674)		(0.0402)
100 km				(0.0758)		(0.0415)
200 km				(0.0942)		(0.0394)
300 km				(0.0996)		(0.0385)

Table continues

Table B.4: Farm value and potential mechanisms, Conley standard errors (continued)

	<b>Farm Value per hectare</b>					
	(1)	(2)	(3)	(4)	(5)	(6)
Share Europeans	0.730***	0.484***	0.547***	0.750***	0.599***	0.281*/**
<b><i>Distance cutoff</i></b>						
50 km	(0.127)	(0.110)	(0.136)	(0.120)	(0.145)	(0.129)
100 km	(0.114)	(0.124)	(0.141)	(0.111)	(0.155)	(0.155)
200 km	(0.124)	(0.124)	(0.152)	(0.117)	(0.147)	(0.124)
300 km	(0.113)	(0.104)	(0.148)	(0.0925)	(0.137)	(0.122)
<b><i>Panel C: Tools &amp; Machines</i></b>						
Share farms with tools					0.178***	0.183***
<b><i>Distance cutoff</i></b>						
50 km					(0.0542)	(0.0454)
100 km					(0.0671)	(0.0539)
200 km					(0.0607)	(0.0475)
300 km					(0.0580)	(0.0510)
Share farms with machines					0.013	-0.001
<b><i>Distance cutoff</i></b>						
50 km					(0.0235)	(0.0197)
100 km					(0.0247)	(0.0206)
200 km					(0.0214)	(0.0172)
300 km					(0.0176)	(0.0155)
Observations	1,285	1,285	1,285	1,285	1,285	1,285
All controls	✓	✓	✓	✓	✓	✓

*Note:* \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Conley standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated by instrumental variables using the 2SLS estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

Table B.5: Implications for development and structural change, Conley standard errors

	(1)	(2)	(3)	(4)
<b>Panel A: Human capital formation (literacy)</b>				
	Population	Brazilians	Females	Children
Share Europeans	0.661***	0.625***	0.596***	0.647***
<b>Distance cutoff</b>				
50 km	(0.136)	(0.140)	(0.131)	(0.133)
100 km	(0.151)	(0.152)	(0.144)	(0.157)
200 km	(0.199)	(0.195)	(0.184)	(0.193)
300 km	(0.184)	(0.175)	(0.170)	(0.159)
Observations	1,296	1,296	1,296	1,296
<b>Panel B: Economic structure</b>				
	Employed Agriculture	Employed Industry	Employed Services	
Share Europeans	-0.371**/***	0.268/*	0.0138	
<b>Distance cutoff</b>				
50 km	(0.167)	(0.144)	(0.0111)	
100 km	(0.146)	(0.155)	(0.0101)	
200 km	(0.185)	(0.188)	(0.0130)	
300 km	(0.118)	(0.173)	(0.00970)	
<b>Panel C: Specialization (probability of being a hub)</b>				
	Agriculture hub	Industry hub	Service hub	
Share Europeans	-0.179**/***	0.180**/***	0.281***	
<b>Distance cutoff</b>				
50 km	(0.0754)	(0.0765)	(0.0809)	
100 km	(0.0684)	(0.0824)	(0.0933)	
200 km	(0.0499)	(0.0860)	(0.0963)	
300 km	()	(0.0679)	(0.0573)	
<b>Panel D: Political Economy</b>				
	Rentiers	Pub. admin		
Share Europeans	0.0997	-0.156		
<b>Distance cutoff</b>				
50 km	(0.175)	(0.172)		
100 km	(0.152)	(0.169)		
200 km	(0.167)	(0.194)		
300 km	(0.104)	(0.146)		
<b>Panel E: Female participation in labor markets (female to male ratio)</b>				
	Population	Employed	Agriculture	Industry
Share Europeans	-0.434*/**/***	-0.175	-0.173	0.145
<b>Distance cutoff</b>				
50 km	(0.149)	(0.147)	(0.169)	(0.116)
100 km	(0.180)	(0.170)	(0.197)	(0.132)
200 km	(0.223)	(0.159)	(0.213)	(0.144)
300 km	(0.184)	(0.152)	(0.193)	(0.173)
Observations	1,296	1,295	1,295	1,295
All controls	✓	✓	✓	✓

Note: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Conley standard errors in parentheses. The unit of observation is a municipality. Regressions in columns (1), (3), (5), and (7) include all municipalities. Regressions in columns (2), (4), (6), and (8) exclude municipalities in the top 25 percentiles of population. All regressions are estimated by instrumental variables using the 2SLS estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. All controls include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.



## C Instrument Based on Distance to Rail

In this appendix, we repeat our instrumental variables estimation, but instead of defining the instrument on the basis of a municipality's own rail connectedness, municipalities are defined to have been connected to the rail network if any municipality within 100 kilometers was connected.

Table C.1: European immigration and farm value, distance-based instrument (100 km)

	Farm Value per hectare				Land	Infrastructure	Tools & Machines
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share Europeans	2.058*** (0.486)	1.510*** (0.329)	1.363*** (0.246)	1.798*** (0.382)	1.630*** (0.296)	1.170*** (0.260)	1.168*** (0.358)
Observations	1,289	1,289	1,289	1,289	1,289	1,289	1,289
1 <sup>st</sup> stage F-stat	13.03	12.78	10.80	7.174	7.174	7.174	7.174
Railway years	✓	✓	✓	✓	✓	✓	✓
No rail	✓	✓	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓	✓
Land adaptability		✓	✓	✓	✓	✓	✓
Land quality		✓	✓	✓	✓	✓	✓
Dom market access			✓	✓	✓	✓	✓
Int market access			✓	✓	✓	✓	✓
State FE				✓	✓	✓	✓

*Note:* \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. The IV regressions are estimated using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *Railway years* is the number of years with railway connection. *No rail* is an indicator variable for municipalities with no railway connection. *Geo controls* include surface area, ruggedness, altitude, latitude, and longitude. *Land adaptability* is the Herfindahl-Hirschman Index (HHI) of crop suitability. *Land quality* includes the first two principal components of suitability for all major crops reported in the census. *Dom market access* includes the linear and quadratic distance to the nearest city and principal town. *Int market access* includes the linear and quadratic distance to the nearest port or custom house. All continuous variables are standardized to have mean zero and standard deviation one.

Table C.2: Potential mechanisms and European immigration, distance-based instrument

	Population density	Agricultural emp. density	Cultivated share of farms	Coffee share	Cash crops share (no coffee)	Share farms with tools	Share farms with machines
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share Europeans	1.368*** (0.432)	3.187** (1.504)	1.085** (0.548)	-12.66 (691.8)	0.525 (0.663)	1.570*** (0.494)	0.496 (0.424)
Observations	1,296	1,296	1,287	1,285	1,285	1,289	1,289
1 <sup>st</sup> stage F-stat	7.286	7.286	7.165	7.104	7.104	7.174	7.174
All controls	✓	✓	✓	✓	✓	✓	✓

*Note:* \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated by instrumental variables using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

Table C.3: Farm value and potential mechanisms, distance-based instrument

	<b>Farm Value per hectare</b>					
	(1)	(2)	(3)	(4)	(5)	(6)
Share Europeans	1.828*** (0.388)	1.585*** (0.493)	1.671*** (0.402)	1.511*** (0.430)	2.117** (0.946)	1.100** (0.498)
<i>Panel A: Land use</i>						
Cultivated share		0.0801 (0.157)				0.218* (0.128)
Coffee share of cultivated land			-0.0589 (0.119)			-0.0363 (0.0748)
Cash crops share of cultivated land			0.110*** (0.0363)			0.0995*** (0.0276)
<i>Panel B: Labor force</i>						
Pop density				0.0413 (0.0676)		0.114 (0.0915)
Agr. emp. density				0.113** (0.0545)		0.0476 (0.0628)
<i>Panel C: Tools &amp; Machines</i>						
Share farms with tools					0.0401 (0.0367)	0.00861 (0.0207)
Share farms with machines					-0.245 (0.276)	0.00117 (0.118)
Observations	1,285	1,285	1,285	1,285	1,285	1,285
1 <sup>st</sup> stage F-stat	7.104	4.455	7.701	3.965	3.284	2.829
All controls	✓	✓	✓	✓	✓	✓

*Note:* \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated by instrumental variables using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

Table C.4: Implications for development and structural change, distance-based instrument

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<b>Panel A: Human capital formation (literacy)</b>								
	Population		Brazilians		Females		Children	
Share Europeans	2.064*** (0.514)	1.440** (0.599)	2.076*** (0.538)	1.372** (0.594)	1.603*** (0.400)	1.125** (0.466)	1.749*** (0.438)	1.257** (0.569)
Observations	1,296	973	1,296	973	1,296	973	1,296	973
1 <sup>st</sup> stage F-stat	7.286	4.115	7.286	4.115	7.286	4.115	7.286	4.115
<b>Panel B: Economic structure</b>								
	Employed Agriculture		Employed Industry		Employed Services			
Share Europeans	-1.703*** (0.614)	-1.833 (1.307)	1.912*** (0.545)	2.114** (0.933)	1.228*** (0.416)	1.289 (0.927)		
Observations	1,295	972	1,295	972	1,295	972		
1 <sup>st</sup> stage F-stat	7.297	4.126	7.297	4.126	7.297	4.126		
<b>Panel C: Specialization (probability of being a hub)</b>								
	Agriculture hub		Industry hub		Service hub			
Share Europeans	-0.293 (0.230)	-0.182 (0.479)	0.973*** (0.312)	1.080* (0.599)	0.486*** (0.157)	0.528** (0.251)		
Observations	1,296	973	1,296	973	1,296	973		
1 <sup>st</sup> stage F-stat	7.286	4.115	7.286	4.115	7.286	4.115		
<b>Panel D: Political Economy</b>								
	Rentiers		Pub. admin					
Share Europeans	0.988** (0.436)	0.838* (0.459)	-0.0150 (0.352)	-0.649 (0.693)				
Observations	1,296	973	1,296	973				
1 <sup>st</sup> stage F-stat	7.286	4.115	7.286	4.115				
<b>Panel E: Female participation in labor markets (female to male ratio)</b>								
	Population		Employed		Agriculture		Industry	
Share Europeans	-4.092 (49.00)	-9.104 (336.4)	5.201 (11.81)	19.43 (420.5)	-4.652 (8.734)	-7.986 (27.75)	0.868* (0.446)	0.915** (0.431)
Observations	1,296	973	1,295	972	1,295	972	1,295	972
1 <sup>st</sup> stage F-stat	7.286	4.115	7.297	4.126	7.297	4.126	7.297	4.126
All controls	✓	✓	✓	✓	✓	✓	✓	✓

Note: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. Regressions in columns (1), (3), (5), and (7) include all municipalities. Regressions in columns (2), (4), (6), and (8) exclude municipalities in the top 25 percentiles of population. All regressions are estimated by instrumental variables using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

## D Instrumental Variables Estimates with State-Level Heterogeneity

Our main instrumental variables analysis allows for the relationship between the instrument and the endogenous regressor to vary at the region level. In this appendix, we allow for the relationship between the instrument and the endogenous regressor to vary at the state level.

Table D.1: European immigration and farm value, IV with state-level heterogeneity

	Farm Value per hectare			Land	Infrastructure	Tools & Machines
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Panel A: OLS</b>						
Share Europeans	0.601*** (0.0655)	0.606*** (0.0667)	0.606*** (0.0665)	0.570*** (0.0660)	0.599*** (0.0659)	0.430*** (0.0549)
Observations	1,289	1,289	1,289	1,289	1,289	1,289
$R^2$	0.572	0.587	0.595	0.605	0.463	0.404
<b>Panel B: IV</b>						
Share Europeans	0.801*** (0.127)	0.855*** (0.120)	0.877*** (0.132)	0.852*** (0.136)	0.700*** (0.108)	0.920*** (0.144)
Observations	1,289	1,289	1,289	1,289	1,289	1,289
1 <sup>st</sup> stage F-stat	22.81	17.48	18.01	18.01	18.01	18.01
Railway years	✓	✓	✓	✓	✓	✓
No rail	✓	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓	✓
State FE	✓	✓	✓	✓	✓	✓
Land adaptability		✓	✓	✓	✓	✓
Land quality		✓	✓	✓	✓	✓
Dom market access		✓	✓	✓	✓	✓
Int market access		✓	✓	✓	✓	✓

*Note:* \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. The IV regressions are estimated using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with state indicators, which instrument for the actual share of Europeans. *Railway years* is the number of years with railway connection. *No rail* is an indicator variable for municipalities with no railway connection. *Geo controls* include surface area, ruggedness, altitude, latitude, and longitude. *Land adaptability* is the Herfindahl-Hirschman Index (HHI) of crop suitability. *Land quality* includes the first two principal components of suitability for all major crops reported in the census. *Dom market access* includes the linear and quadratic distance to the nearest city and principal town. *Int market access* includes the linear and quadratic distance to the nearest port or custom house. All continuous variables are standardized to have mean zero and standard deviation one.

Table D.2: Potential mechanisms and European immigration, IV with state-level heterogeneity

	Cultivated share of farms (1)	Coffee share (2)	Cash crops share (no coffee) (3)	Population density (4)	Agricultural emp. density (5)	Share farms with tools (6)	Share farms with machines (7)
Share Europeans	0.724*** (0.122)	0.925*** (0.172)	-0.0360 (0.0731)	0.0479 (0.115)	0.000594 (0.117)	0.653*** (0.0965)	0.128 (0.101)
Observations	1,287	1,285	1,285	1,296	1,296	1,289	1,289
1 <sup>st</sup> stage F-stat	18.01	17.97	17.97	18.20	18.20	18.01	18.01
All controls	✓	✓	✓	✓	✓	✓	✓

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated by instrumental variables using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with state indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

Table D.3: Farm value and potential mechanisms, IV with state-level heterogeneity

	<b>Farm Value per hectare</b>					
	(1)	(2)	(3)	(4)	(5)	(6)
Share Europeans	0.876*** (0.116)	0.828*** (0.105)	0.697*** (0.148)	0.745*** (0.135)	0.751*** (0.136)	0.424** (0.169)
<i>Panel A: Labor force</i>						
Pop density		0.0987** (0.0491)				0.208*** (0.0662)
Agr. emp. density		0.166*** (0.0447)				-0.0176 (0.0440)
<i>Panel B: Land use</i>						
Cultivated share of farms			0.342*** (0.0626)			0.374*** (0.0660)
Coffee share				0.180*** (0.0421)		0.0418 (0.0339)
Other cash crops share				0.121*** (0.0244)		0.0906*** (0.0202)
<i>Panel C: Tools &amp; Machines</i>						
Share farms with tools					0.144*** (0.0466)	0.156*** (0.0426)
Share farms with machines					0.0223 (0.0183)	0.00355 (0.0159)
Observations	1,231	1,231	1,231	1,231	1,231	1,231
1 <sup>st</sup> stage F-stat	18.55	17.83	15.91	15.40	13.19	10.13
All controls	✓	✓	✓	✓	✓	✓

*Note:* \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated by instrumental variables using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with state indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

Table D.4: Implications for development and structural change, IV with state-level heterogeneity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A: Human capital formation (literacy)</i>								
	<b>Population</b>		<b>Brazilians</b>		<b>Females</b>		<b>Children</b>	
Share Europeans	0.503*** (0.106)	0.598*** (0.126)	0.481*** (0.107)	0.552*** (0.124)	0.489*** (0.103)	0.555*** (0.118)	0.515*** (0.101)	0.568*** (0.119)
Observations	1,296	973	1,296	973	1,296	973	1,296	973
1 <sup>st</sup> stage F-stat	18.20	21.83	18.20	21.83	18.20	21.83	18.20	21.83
<i>Panel B: Economic structure</i>								
	<b>Employed Agriculture</b>		<b>Employed Industry</b>		<b>Employed Services</b>			
Share Europeans	-0.469*** (0.138)	-0.389** (0.157)	0.310** (0.134)	0.221* (0.152)	0.257* (0.187)	0.317* (0.187)		
Observations	1,295	972	1,295	972	1,295	972		
1 <sup>st</sup> stage F-stat	18.19	21.83	18.19	21.83	18.19	21.83		
<i>Panel C: Specialization (probability of being a hub)</i>								
	<b>Agriculture hub</b>		<b>Industry hub</b>		<b>Service hub</b>			
Share Europeans	-0.233*** (0.0491)	-0.239*** (0.0603)	0.166*** (0.0556)	0.0764 (0.0696)	0.245*** (0.0557)	0.217*** (0.0730)		
Observations	1,296	973	1,296	973	1,296	973		
1 <sup>st</sup> stage F-stat	18.20	21.83	18.20	21.83	18.20	21.83		
<i>Panel D: Political Economy</i>								
	<b>Rentiers</b>		<b>Pub. admin</b>					
Share Europeans	0.213 (0.138)	0.179 (0.163)	0.0427 (0.146)	0.277 (0.183)				
Observations	1,296	973	1,296	973				
1 <sup>st</sup> stage F-stat	18.20	21.83	18.20	21.83				
<i>Panel E: Female participation in labor markets (female to male ratio)</i>								
	<b>Population</b>		<b>Employed</b>		<b>Agriculture</b>		<b>Industry</b>	
Share Europeans	-0.200* (0.104)	-0.363* (0.193)	0.0521 (0.0832)	0.0735 (0.107)	0.0821 (0.126)	0.175 (0.151)	0.0764 (0.0606)	0.0343 (0.0856)
Observations	1,296	973	1,295	972	1,295	972	1,295	972
1 <sup>st</sup> stage F-stat	18.20	21.83	18.19	21.83	18.19	21.83	18.19	21.83
All controls	✓	✓	✓	✓	✓	✓	✓	✓

*Note:* \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. Regressions in columns (1), (3), (5), and (7) include all municipalities. Regressions in columns (2), (4), (6), and (8) exclude municipalities in the top 25 percentiles of population. All regressions are estimated by instrumental variables using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with state indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

## E Two-Stage Least Squares Estimates

Our main results are estimated by instrumental variables using the limited information maximum likelihood (LIML) estimator. In this appendix, we show that our results are robust to using the two stage least squares (2SLS) estimator.

Table E.1: European immigration and farm value, two-stage least squares estimation

	Farm Value per hectare				Land	Infrastructure	Tools & Machines
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share Europeans	0.662*** (0.112)	0.659*** (0.0985)	0.707*** (0.0960)	0.724*** (0.104)	0.705*** (0.101)	0.576*** (0.132)	0.735*** (0.118)
Observations	1,289	1,289	1,289	1,289	1,289	1,289	1,289
1 <sup>st</sup> stage F-stat	16.87	22.21	22.61	22.31	22.31	22.31	22.31
Railway years	✓	✓	✓	✓	✓	✓	✓
No rail	✓	✓	✓	✓	✓	✓	✓
Geo Controls	✓	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓				
Land adaptability		✓	✓	✓	✓	✓	✓
Land quality		✓	✓	✓	✓	✓	✓
Dom market access			✓	✓	✓	✓	✓
Int market access			✓	✓	✓	✓	✓
State FE				✓	✓	✓	✓

*Note:* \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated using the two stage least squares (2SLS) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *Railway years* is the number of years with railway connection. *No rail* is an indicator variable for municipalities with no railway connection. *Geo controls* include surface area, ruggedness, altitude, latitude, and longitude. *Land adaptability* is the Herfindahl-Hirschman Index (HHI) of crop suitability. *Land quality* includes the first two principal components of suitability for all major crops reported in the census. *Dom market access* includes the linear and quadratic distance to the nearest city and principal town. *Int market access* includes the linear and quadratic distance to the nearest port or custom house. All continuous variables are standardized to have mean zero and standard deviation one.

Table E.2: Potential mechanisms and European immigration, two-stage least squares estimation

	Population density	Agricultural emp. density	Cultivated share of farms	Coffee share	Cash crops share (no coffee)	Share farms with tools	Share farms with machines
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share Europeans	0.728*** (0.134)	0.823*** (0.143)	0.0131 (0.102)	-0.173 (0.159)	-0.0614 (0.150)	0.675*** (0.103)	0.0456 (0.141)
Observations	1,287	1,285	1,285	1,296	1,296	1,289	1,289
1 <sup>st</sup> stage F-stat	22.29	22.26	22.26	22.35	22.35	22.31	22.31
All controls	✓	✓	✓	✓	✓	✓	✓

*Note:* \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated using the two stage least squares (2SLS) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

Table E.3: Farm value and potential mechanisms, two-stage least squares estimation

	<b>Farm Value per hectare</b>					
	(1)	(2)	(3)	(4)	(5)	(6)
Share Europeans	0.709*** (0.102)	0.474*** (0.104)	0.527*** (0.115)	0.727*** (0.0935)	0.571*** (0.118)	0.273** (0.110)
<i>Panel A: Land use</i>						
Cultivated share		0.406*** (0.0563)				0.410*** (0.0590)
Coffee share of cultivated land			0.235*** (0.0409)			0.0638** (0.0304)
Cash crops share of cultivated land			0.132*** (0.0246)			0.0903*** (0.0206)
<i>Panel A: Labor force</i>						
Pop density				0.140** (0.0601)		0.254*** (0.0678)
Agr. emp. density				0.151*** (0.0457)		-0.0465 (0.0385)
<i>Panel C: Tools &amp; Machines</i>						
Share farms with machines					0.190*** (0.0460)	0.188*** (0.0362)
Share farms with tools					0.0220 (0.0186)	0.00411 (0.0160)
Observations	1,232	1,232	1,232	1,232	1,232	1,232
1 <sup>st</sup> stage F-stat	22.18	15.64	15.64	24.63	16.50	11.40
All controls	✓	✓	✓	✓	✓	✓

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated using the two stage least squares (2SLS) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

Table E.4: Implications for development and structural change, two-stage least squares estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<b>Panel A: Human capital formation (literacy)</b>								
	<b>Population</b>		<b>Brazilians</b>		<b>Females</b>		<b>Children</b>	
Share Europeans	0.661*** (0.117)	0.713*** (0.131)	0.625*** (0.122)	0.683*** (0.137)	0.596*** (0.116)	0.644*** (0.128)	0.647*** (0.115)	0.720*** (0.131)
Observations	1,296	973	1,296	973	1,296	973	1,296	973
1 <sup>st</sup> stage F-stat	22.35	16.04	22.35	16.04	22.35	16.04	22.35	16.04
<b>Panel B: Economic structure</b>								
	<b>Employed Agriculture</b>		<b>Employed Industry</b>		<b>Employed Services</b>			
Share Europeans	-0.371** (0.147)	-0.337** (0.166)	0.268** (0.131)	0.187 (0.145)	0.218 (0.160)	0.241 (0.182)		
Observations	1,295	972	1,295	972	1,295	972		
1 <sup>st</sup> stage F-stat	22.38	16.09	22.38	16.09	22.38	16.09		
<b>Panel C: Specialization (probability of being a hub)</b>								
	<b>Agriculture hub</b>		<b>Industry hub</b>		<b>Service hub</b>			
Share Europeans	-0.179*** (0.0650)	-0.184** (0.0762)	0.180*** (0.0662)	0.0769 (0.0811)	0.281*** (0.0667)	0.250*** (0.0847)		
Observations	1,296	973	1,296	973	1,296	973		
1 <sup>st</sup> stage F-stat	22.35	16.04	22.35	16.04	22.35	16.04		
<b>Panel D: Political Economy</b>								
	<b>Rentiers</b>		<b>Pub. admin</b>					
Share Europeans	0.0997 (0.169)	0.00194 (0.218)	-0.156 (0.148)	-0.0175 (0.156)				
Observations	1,296	973	1,296	973				
1 <sup>st</sup> stage F-stat	22.35	16.04	22.35	16.04				
<b>Panel E: Female participation in labor markets (female to male ratio)</b>								
	<b>Population</b>		<b>Employed</b>		<b>Agriculture</b>		<b>Industry</b>	
Share Europeans	-0.434*** (0.119)	-0.463*** (0.140)	-0.175 (0.131)	-0.0576 (0.157)	-0.173 (0.139)	-0.0180 (0.166)	0.145 (0.0981)	0.116 (0.135)
Observations	1,296	973	1,295	972	1,295	972	1,295	972
1 <sup>st</sup> stage F-stat	22.35	16.04	22.38	16.09	22.38	16.09	22.38	16.09
All controls	✓	✓	✓	✓	✓	✓	✓	✓

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. The unit of observation is a municipality. Regressions in columns (1), (3), (5), and (7) include all municipalities. Regressions in columns (2), (4), (6), and (8) exclude municipalities in the top 25 percentiles of population. All regressions are estimated by instrumental variables using the two stage least squares (2SLS) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

## F Additional results for farm value components

Table F.1: European immigration and farm value, robustness tests for the components of farm value

	Land (1)	Infrastructure (2)	Tools & Machines (3)	Land (4)	Infrastructure (5)	Tools & Machines (6)
	<i>Panel A: no immigrant colonies</i>			<i>Panel B: no large coffee producers</i>		
Share Europeans	0.640*** (0.120)	0.526*** (0.146)	0.677*** (0.129)	0.650*** (0.161)	0.556*** (0.169)	0.688*** (0.142)
Observations	1,159	1,159	1,159	1,157	1,157	1,157
1 <sup>st</sup> stage F-stat	22.14	22.14	22.14	13.25	13.25	13.25
	<i>Panel C: no early railway connections</i>			<i>Panel D: no large population centers</i>		
Share Europeans	0.604*** (0.168)	0.547*** (0.204)	0.628** (0.251)	0.751*** (0.153)	0.575*** (0.152)	0.791*** (0.174)
Observations	1,111	1,111	1,111	968	968	968
1 <sup>st</sup> stage F-stat	20.37	20.37	20.37	16.06	16.06	16.06
	<i>Panel E: no industry hubs</i>			<i>Panel F: no service hubs</i>		
Share Europeans	0.849*** (0.109)	0.710*** (0.169)	0.687*** (0.129)	0.700*** (0.163)	0.470*** (0.178)	0.562*** (0.166)
Observations	965	965	965	967	967	967
1 <sup>st</sup> stage F-stat	22.53	22.53	22.53	19.71	19.71	19.71
All controls	✓	✓	✓	✓	✓	✓

*Note:* \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated by instrumental variables using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

Table F.2: Farm value components, European immigration, and potential mechanisms

	<b>Land</b>		<b>Infrastructure</b>		<b>Tools &amp; machines</b>	
	(1)	(2)	(3)	(4)	(5)	(6)
Share Europeans	0.724*** (0.119)	0.284* (0.161)	0.540*** (0.140)	0.149 (0.177)	0.746*** (0.123)	0.343** (0.172)
<i>Panel A: Land use</i>						
Cultivated share		0.373*** (0.0658)		0.463*** (0.0736)		0.216*** (0.0599)
Coffee share of cultivated land		0.0657* (0.0337)		0.0257 (0.0400)		0.119** (0.0475)
Cash crops share of cultivated land		0.0996*** (0.0202)		0.0319 (0.0283)		0.106*** (0.0320)
<i>Panel B: Labor force</i>						
Pop density		0.214*** (0.0613)		0.363*** (0.128)		0.0369 (0.0359)
Agr. emp. density		-0.0209 (0.0401)		-0.131* (0.0706)		0.0437 (0.0466)
<i>Panel C: Tools &amp; Machines</i>						
Share farms with machines		0.182*** (0.0459)		0.147*** (0.0507)		0.235*** (0.0555)
Share farms with tools		-0.00416 (0.0169)		-0.0182 (0.0170)		0.205*** (0.0276)
Observations	1,232	1,232	1,232	1,232	1,232	1,232
1 <sup>st</sup> stage F-stat	22.18	11.40	22.18	11.40	22.18	11.40
All controls	✓	✓	✓	✓	✓	✓

*Note:* \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities are included in the analysis. All regressions are estimated by instrumental variables using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

## G Additional Mechanisms

In this appendix, we present the results for three mechanisms that we find *not* to have been present or to have been less important relative to the cultivation intensity mechanism discussed in section 7.

### G.1 Labor Force

One mechanism that we test attributes the effects of immigration directly to its role in increasing the size of the agricultural labor force in a particular place. That is, immigration may have increased the number of available workers, enabling the exploitation of land, just as the immigration subsidy program intended. A greater population may also have increased local demand for land or agricultural products. We operationalize demand for land using population density (inhabitants per km<sup>2</sup>) and the agricultural labor supply using agricultural employment density (workers employed in agriculture per km<sup>2</sup>).

These mechanisms, however, are not supported by the data. In columns (4) and (5) of Table 5, we find no statistically significant relationship between immigration and either population density or agricultural employment density. Indeed, the point estimates are negative, contrary to the effect of immigration on density that would be required under such a mechanism.

While these results speak to the cumulative effect of immigration, it is possible that the absence of a cumulative effect of immigration on population density obscures a shorter-term effect that dissipated over time. To provide a glimpse into such a possible effect, Table G.1 studies the relationship between the share of immigrants in population in 1890, which was the last nationwide census prior to 1920, and population density in that year. While 1890 is very early in Brazil's mass migration experience (which began in earnest in 1888), and the 1890 census is very limited (no distinction by nationality is possible, and very few outcome variables are available), this analysis does provide some insight into the short-run effect of the initial wave of immigration. We find that the foreign share in 1890 is positively related to population density in the same year (column 1), though the coefficient is quite small (an order of magnitude less than the relationship between the 1920 European share and farm values). The effect is statistically significant at the 10 percent level. When we control for population density in 1890 in our standard IV framework, we find this to be positively related to farm values in 1920, but not to diminish the effect of the European share in 1920. If anything, the coefficient for the latter becomes slightly larger. Taken together, these results suggest that immigration had weak effects of population density in the very early stages of mass immigration to Brazil that were not present in 1920, but that population density in this early phase is unable to explain the effect

of immigration on farm values we find in 1920.

Consistent with our finding in Table 5, although column (4) of Table 6 shows that, although both measures of the availability of labor were positively related to the value of farms, the European share coefficient is only minimally changed by the introduction of the labor force variables. This result is confirmed by the decomposition in Table 7, which shows, both for farm values in total and for each individual component, that the inclusion of labor force variables has a relatively small (in comparison to the other mechanisms that we explore below) impact on the coefficient on the European share of population with virtually all of the effect operating through the value of land. These results rule out what is perhaps the most obvious mechanism for the effect of migration on farm prices, showing instead that the effect of immigration is connected to some peculiar characteristic of the immigrant labor force rather than merely its presence.

## G.2 Crop Mix

Another facet of land use that we investigate as a potential mechanism for the effect of immigration on agricultural development is that immigrants led to a change in the crop mix, and in particular to an increase in coffee cultivation. We measure the crop mix using two variables—the share of cultivated land dedicated to coffee and the share of cultivated land dedicated to other cash crops (cotton, cocoa, sugarcane, tobacco, castor beans, and rubber).<sup>97</sup> Columns (2) and (3) of Table 5 show that a greater immigrant share led to a greater share of land devoted to coffee production, but had no impact on the share of land devoted to other cash crops. Columns (3) and (4) of Table 6 control for the cultivated share of farms and for the crop mix, respectively. These factors were both associated with greater farm values, and their inclusion leads to a decline in the magnitude of the European share coefficient, which constitutes evidence that these are mechanisms through which the effect of the immigration share passed. This is confirmed by the decomposition in Table 7, which shows the production of coffee, while responsible for a non-negligible change in the immigrant share coefficient, particularly with respect to the value of tools and machines, appears to have been a less important mechanism than cultivation intensity despite the links between immigration and coffee production.

## G.3 Tools and Machines

We also consider the adoption of agricultural tools and machines as a potential mechanism for the effect of immigration. We build on our finding that immigration increased all components of farm value, including

---

<sup>97</sup>The residual category is “staple” crops (rice, maize, wheat, beans, potatoes, manioc, and coconut).

that of tools and machinery and on Lanza, Manier, and Musacchio’s (2023) findings that immigration under the subsidized immigration program in São Paulo state led to greater adoption of agricultural tools and machines.

Column (6) of Table 5 shows, consistent with our results in column (7) of Table 2, that immigration led to a greater share of farms using tools while column (7) shows that that immigration had no effect on the share of farms using machines. Such a relationship can be rationalized by considering tools as complementary to labor and machines as substitutes (Lanza, Manier, and Musacchio 2023). This result stands out in comparison to results from the United States after the end of the Age of Mass Migration (Abramitzky et al. 2023) and the Bracero program (Clemens, Lewis, and Postel 2018), which document an increase in capital intensity in agriculture with reductions in immigration.

In Table 6, we find that farms that adopted agricultural tools were indeed more valuable, though the effect is less than 0.2 standard deviations, and that the size of the European migrants share coefficient is reduced by around 0.12 standard deviations when it is included. Notably, columns (2) and (3) of Table 7 show that the tool mechanism goes beyond the mechanical channel of increasing the tools and machines component of farm values—tools are also responsible for reducing the coefficient on the immigrant share for the land and infrastructure components of value as well. The magnitude of the impact is about one-third that of the share of farms cultivated, and so we view the tools mechanism as second order in comparison to it.

Table G.1: Immigration and population density in 1890

	Pop density 1890	Farm Value per hectare	
	(1)	(2)	(3)
Share European 1920		0.721*** (0.120)	0.782*** (0.113)
Share Foreign 1890	0.0528* (0.0311)		
Pop density 1890			0.231*** (0.0598)
Observations	957	958	958
$R^2$	0.460		
1 <sup>st</sup> stage F-stat		15.16	15.97
All controls	✓	✓	✓

*Note:* \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses. The unit of observation is a municipality. All municipalities (observed in the 1890 census) are included in the analysis. All regressions are estimated by instrumental variables using the limited information maximum likelihood (LIML) estimator. The instruments are the predicted share of Europeans interacted with region indicators, which instrument for the actual share of Europeans. *All controls* include the number of years with railway connection, an indicator variable for municipalities with no railway connection, geographic characteristics (surface area, ruggedness, altitude, latitude, and longitude), land characteristics (the Herfindahl-Hirschman Index of crop suitability and the first two principal components of suitability for all major crops reported in the census), market access characteristics (linear and quadratic distance to the nearest city, principal town, port, and custom house), and state fixed effects. All continuous variables are standardized to have mean zero and standard deviation one.

## References

- Abramitzky, Ran, Philipp Ager, Leah Platt Boustan, Elior Cohen, and Casper W. Hansen (2023). “The Effects of Immigration on Local Labor Markets: Lessons from the 1920s Border Closure.” *American Economic Journal: Applied Economics* 15:1, pp. 164–191.
- Clemens, Michael A., Ethan G. Lewis, and Hannah M. Postel (2018). “Immigration Restrictions as Active Labor Market Policy: Evidence from the Mexican Bracero Exclusion.” *American Economic Review* 108:6, pp. 1468–1487.
- Directoria Geral de Estatística (1922). *Recensamento do Brazil Realizado em 1 de Setembro de 1920*. Rio de Janeiro: Officina da Estatística.
- Ehrl, Philipp (2017). “Minimum Comparable Areas for the Period 1872–2010: An Aggregation of Brazilian Municipalities.” *Estudos Econômicos* 47:1, pp. 215–229.
- Giesbrecht, Ralph Mennucci (2023). “Estações Ferroviárias do Brasil.” <http://www.estacoesferroviarias.com.br/index.html>, accessed April 25, 2023.
- Instituto Brasileiro de Geografia e Estatística (2011). “Evolução da Divisão Territorial do Brasil: 1872-2010.” <https://maps.princeton.edu/catalog/stanford-sf351hx7032>, accessed April 24, 2023. Rio de Janeiro.
- (2023). “Cidades.” <https://cidades.ibge.gov.br/>.
- Lanza, André Luiz, Manas Manier, and Aldo Musacchio (2023). “European Immigration and Agricultural Productivity in São Paulo, Brazil, 1898–1920.” In *Roots of Underdevelopment: A New Economic and Political History of Latin America and the Caribbean*. Felipe Valencia Caicedo (ed.). Springer, pp. 213–245.
- Secreteria da Agricultura, Commercio e Obras Publicas do Estado de São Paulo (1906–1910). *Estatística Agrícola e Zootécnica do Estado de São Paulo no Anno Agrícola de 1904-1905*. São Paulo: Typ. Brazil—Carlos Gerke & Rothschild.