Frontier planters, immigrants, and the abolition of slavery in Brazil*

François Seyler† Arthur Silve‡

May 6, 2021

Abstract

A protracted legislative battle culminated in the abolition of slavery in Brazil in 1888. We build a new data set of roll-call votes on 1884-1888 emancipation bills in the legislature, and connect it to local features of the districts. This allows us to unpack how the material interests of each of the 122 electoral districts coalesced into an abolitionist coalition. Our results help reconcile previous theories of labor coercion. We find slavery-intensive districts opposed emancipation. In line with a *labor demand effect*, we also find more support for emancipation where immigrants provided an alternative source of labor, and in line with an *outside option effect*, where slaves could more easily escape. A two-pronged instrumental variables strategy that uses variation in (a) history and geography and (b) heteroskedasticity with respect to the regressors supports a causal interpretation of our main results.

Key words: labor coercion, Brazil, slavery, immigration, institutional change, intra-elite conflict.

JEL codes: J15, J47, N36, N46, O54.

*We thank Daron Acemoglu, Charles Angelucci, José De Sousa, Christian Dippel, James Fenske, Claudio Ferraz, Bernard Fortin, Jeffry Frieden, Marion Goussé, Marion Mercier, Christoph Mikulaschek, Eric Rasmusen, Sulin Sardoschau, and participants to seminars at Harvard, at the Vancouver School of Economics, and Paris-1 for comments at various stages of this project. We also thank Leonardo Monasterio for the digitized version of the censuses, and the extremely kind and helpful staff at various archives in São Paulo, Rio de Janeiro, Curitiba, and Brasilia.

†Université Laval, Department of Economics. Email: francois.seyler.1@ulaval.ca

‡Université Laval, Department of Economics. Email: arthur.silve@ecn.ulaval.ca.

Art. 1.º: É declarada extinta desde a data desta lei a escravidão no Brasil.
Art. 2.º: Revogam-se as disposições em contrário.

Lei Áurea, May 13, 1888

1 Introduction

In their study of the economics of labor coercion, Acemoglu and Wolitzky (2011) proposed a theoretical argument to unify two mechanisms that scholars had long thought contradicted each other: a *labor demand effect*, by which the elite supports coercion if labor is scarce, and an *outside option effect*, by which a coercive arrangement declines if workers can easily evade it. We provide empirical evidence that in Brazil, both mechanisms played a role in building the coalition that eventually abolished slavery.

With the 1888 *Lei Áurea*, Brazil was the last country to abolish slavery in the Western Hemisphere. With the benefit of hindsight, it is maybe easy to believe that slavery was doomed to disappear eventually.[1] This teleological view of history overlooks a unique combination of obstacles to abolition in nineteenth-century Brazil. The roots of slavery are deep in the Americas. After the sixteenth-century repression of the Tupi-Guaraní-speaking populations along the Brazilian coastline (Schwartz, 1978), the success of large-scale sugar production and the massive arrival of African slaves marked the birth of a deeply coercive system that permeated through the New World. The Brazilian economy was entirely driven by agricultural exports and heavily dependent on captive labor. According to the latest estimates of the *Slave Voyages* online database (Eltis, 2007), 5,532,120 Africans arrived in Brazil between 1550 and 1866. In comparison, 'only' 472,382 slaves disembarked in mainland North America,

[1] At least, legal slavery. In the Brazilian Amazon, tens of thousands of workers (mostly internal migrants from rural areas) continue to live under coercive labor arrangements. The exact number of coerced workers is not known, but according to Repórter Brasil (2015), nearly 50,000 workers were freed between 1995 and 2015. Illegal slavery is very much a modern problem, and not only in Brazil: from the *kafala* system in Qatar to debt-bondage in India, examples abound globally. The International Labour Organization and the Walk Free Foundation (2017) estimate that in 2016, 40 million people were victims of modern slavery, 71% of whom women and girls.

nearly twelve times fewer. Maybe the real puzzle is not why Brazil abolished slavery so late, but how it came to abolish it at all.

Abolition was the result of a protracted legislative battle between members of the elite, who did not all profit from slavery equally. As the center of gravity of the economy moved from north to south, and from sugarcane to coffee, so did a massive number of slaves.[2] It is thus unsurprising that the northern elites were ready earlier for a transition to free labor, and that the elite as a whole was divided on the issue of slavery.[3] Even within the elite of the *Centro-Sul*, the coffee-growing South-Central region, frontier planters and old *latifundiários* started adopting somewhat antagonistic stances after the 1871 law. Frontier planters struggled to attract an adequate supply of labor for abundant land. Meanwhile, landowners in older settlement regions were quickly exhausting their land (Reis and Reis, 1988).[4]

This article establishes the importance of intra-elite divisions as a driver of institutional persistence and change. Our three main results correspond to three sources of variation in local elites' support for slavery, and our identification strategy supports a causal interpretation of these results.

i) Elites differed in the vested interests they had in the slaves that they or their patrons detained. According to the demographic census (see Figure 1), slaves represented between 0% and 53% of a district's population in 1872. We estimate that an increase in the prevalence of slavery in the district by one standard deviation (9%) was associated with a 27 percentage points increase in a legislator's likelihood to vote against emancipation bills.

[2]In the second half of the century, coffee exports and export prices respectively increased by 341% and 91%, while exports of sugar continued growing by a mere 33% and prices actually decreased by 11% (Viotti da Costa, 1989). Hence, whereas sugar represented 49% of the country's exports and coffee 19% in 1822, by 1913 sugar and cotton together accounted for less than 3% of Brazil's exports, while coffee had gone up to 60% (Leff, 1991). Slaves represented 23% of the northeastern population in 1823 and less than 10% in 1872, a 30% nominal decline. In 1874, more than 50% of the country's slave population was located in the Centro-Sul (Stein, 1957).

[3]Conrad (1972, p. 67) quotes the northern representative Araújo Lima who in 1854 established a parallel with the United States: "Be certain that you will have opposite interests, provinces with slaves, provinces without slaves (...). You will have the kind of struggles and antagonisms (...) which have placed the American Union in such imminent danger."

[4]The slave population decreased everywhere between 1873 and 1882, except in Rio de Janeiro's coffee *municípios* (+4.85%), in the Paulista part of the Paraíba Valley (+10.54%), and in the new coffee region of São Paulo (+45.51%). In 1882, the total slave population of Rio's coffee region comprised 156,009 individuals, 37,649 in the part of the Paraíba Valley belonging to São Paulo, and 38,242 in the new coffee region to the northwest of São Paulo (Conrad, 1972, p. 294-95).

ii) Immigrant labor could substitute for slave labor; but it was not uniformly available. According to the census, foreigners represented between 0% and 24% of a district's population in 1890. We estimate that an increase in the share of immigrants in the district by one standard deviation (4%) was associated with a 15 percentage points increase in a legislator's likelihood to vote in favor of emancipation bills.

iii) Enforcement of the coercive institution depended on how difficult it was for slaves to escape their condition. With data from Fundação Palmares (2020) and INCRA (2020), we capture the local 'distance to freedom' by constructing a collection of measures of the proximity to *quilombos*, communities of maroons that were able to escape and hide in the hinterland, and of their size. We estimate that increasing the average area of quilombola land in a district's municipalities by the interquartile range [IQR] (20km^2) was associated with an 8 percentage points increase in a high-prevalence legislator's likelihood to vote in favor of emancipation bills (and decreasing the average distance of municipalities to the nearest quilombo by one SD, i.e. 29km, with a 3 percentage points increase).

As a country with an open agricultural frontier, large stretches of uncharted territory, and where slavery played such a defining role in social relations, nineteenth-century Brazil is an ideal case-study to verify how landowners used slavery to sustain their status and, among landowners, which ones found it in their best interest to hasten the transition to free labor. We draw from a wealth of archival records, census surveys, geo-referenced spatial sources and historical maps. We retrieve roll-call votes on each emancipation-related bill in the Empire's three last legislatures, ie. between 1882 and 1889, from the annals of the *Câmara dos Deputados*, the lower chamber of the Brazilian Parliament. We match every legislator recorded in these votes to the electoral district in which they were elected, and we construct a database of relevant descriptors of each municipality in each electoral district.

An important feature of this empirical setting is that slavery was more prevalent both in the proximity of *caminhos do ouro*,[5] the trade routes that were built in the eighteenth century to transport gold, cattle and slaves, and within areas where Indigenous peoples had previously been repressed and enslaved. We exploit these sources of variation to overcome the concern that other within-province determinants of slavery in 1872 also affect voting patterns in the 1880s. The validity of this instrument hinges

[5] Gold paths, also known as *estradas reais*, i.e. royal roads.

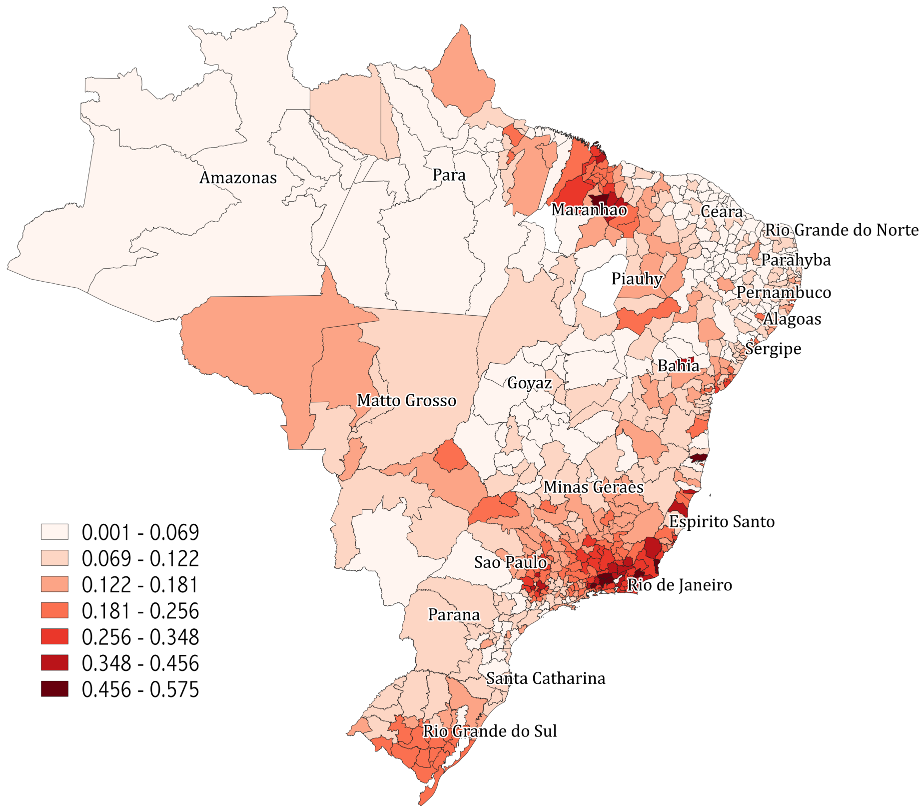


Figure 1: Enslaved population by municipalities in 1872 (Brazil, 1874). Municipality boundaries in 1872 come from IBGE (2010).

on one critical assumption: it must not have political consequences on emancipation in the 1880s beyond the higher prevalence of slavery. A particular concern is that caminhos do ouro may have durably affected economic activity through channels parallel to slavery. If this affected political decisions on emancipation, our identification assumption would be violated. In addition to citing historical evidence to the contrary, we control for a large set of determinants of economic activity in the district and unobserved heterogeneity across provinces and votes. We also run a falsification test showing that neither our instrument nor its components are significant predictors of economic activity in the long run.

To find exogenous variation in the location of immigrants, we exploit the fact that the arrival of immigrants in Brazil accelerated in the 1880s. In contrast with prior decades, the looming abolition of slavery and resulting labor demand drove a proactive immigration policy (Rocha et al., 2017; Witzel de Souza, 2019; de Carvalho Filho and Monasterio, 2012). We use the rich information about the nationality and religion of immigrants that is available in the 1872 and 1890 censuses to construct a religion-based leave-out shift-share (or 'Bartik') instrument – similar in spirit to Tabellini's (2020) approach – that predicts the inflows of immigrants between 1872 and 1890 based on preexisting migration networks in each district. The centralized nature of political decisions to attract migrants in the 1880s implies that the concern that inflows of immigrants from each sending country may have been responding to municipality-specific conditions (Jaeger et al., 2018) appears unlikely, but also that immigrant networks in 1880s Brazil certainly played a less central role than in many similar contexts, which raises the concern that identification may be weaker in this case. Nonetheless, the historiography suggests that such networks did play a role, in particular with early movers becoming credit suppliers to prospective migrants (Witzel de Souza, 2019), and we show that the association generated is sufficiently strong to dissipate these concerns. Furthermore, as Jaeger et al. (2018) point out, the observation that the instrument is not *too* strong is in fact an argument in favor of its validity, as shift-share instruments are more likely to isolate the exogenous component of immigrant inflows when the latter vary significantly over time. Another particular challenge to this identification strategy is that municipality-level conditions affecting the distribution of immigrants by religion before 1872 must be unrelated to abolitionism in the 1880s (Goldsmith-Pinkham et al., 2020). To alleviate this concern, we systematically control for migration levels in 1872, for a large set of district characteristics that may have attracted earlier migrants, and for unobserved characteristics of the province and of each vote. We also address the related concern that immigrants of specific religions or from specific countries may have selected their destination based on emancipation prospects, and show that results are robust to various ways to build and scale the instrument.

A key difference with the United States is that escaping slavery did not 'simply' mean heading north and trying to reach Canada (Allen, 2015). Instead, escaped slaves often sought refuge in the hinterland and either founded or joined existing quilombos, whose location was driven by considerations of security and remoteness. To overcome omitted variables concerns (i.e. that other within-province determinants of the location of quilombos also affect voting patterns in the 1880s), we predict the location and size of quilombos by using variation in the extent to which local features of the land were conducive to allowing slaves to make a successful escape. In particular, we exploit the variation induced by the interaction between ruggedness and remoteness. In the proximity of large settlement areas, a rugged terrain facilitated the escape of fugitives. Remoteness could substitute for ruggedness: terrain ruggedness was less important at a safe distance from cities. This instrument is only valid if within-province variation in this interaction does not affect legislators' voting decisions other than by its influence on slaves' 'outside options' (e.g. via trade and transport costs). We show that this assumption is likely to hold. A key feature of our empirical design is that quilombos do not, by themselves, influence legislators' voting decisions. We exploit this feature to show that the instrument has no placebo effects in the reduced form. Moreover, the instrument remains stable if we individually control (in addition to our full vector of controls and fixed effects) for remoteness and ruggedness, as well as when we flexibly control for nonlinearities in the un-interacted terms.

Finally, we leverage heteroskedasticity with respect to exogenous regressors. This allows us to supplement our instruments with heteroskedasticity-based instruments built as functions of the model's data, following the procedure of Lewbel (2012). These generated instruments are in general not required, but they significantly improve our estimates' precision, in addition to allowing us to run overidentification tests whenever standard instruments leave our models just-identified. Perhaps even more importantly, they allow us to compare the estimates yielded by standard 2SLS with those obtained using heteroskedasticity-based identification. Obtaining numerically similar estimates from approaches relying on different identification assumptions is a strong argument in favor of the model's robustness.

The remainder of the paper proceeds as follows. Section 2 reviews the related literature. Section 3 provides additional elements of historical context. Section 4 describes our data sources. Section 5 presents our empirical strategy. Section 6 lays out the results. Section 7 discusses briefly a possible selection effect in our data. Section 8 concludes.

2 A selective review of the literature

This paper makes a new contribution to the old debate that surrounded the so-called Nieboer-Domar hypothesis. Domar (1970) revived a conjecture first formulated by Nieboer (1900), by which we may never simultaneously find in an agrarian economy "free land, free peasants, and non-working landowners." If land is abundant and labor is scarce, individuals can only derive rents from owning land if its acquisition is restricted for workers and/or if the peasants' freedom is restricted. Brenner (1976) noted that historians had argued that the scarcity of labor after the Black Death was responsible for the progressive decline of coercive feudal arrangements in fourteenth- to sixteenth-century Western Europe, but also for the rise of the 'second serfdom' in late fifteenth-century Eastern Europe, in apparent contradiction with one another (see Aston and Philpin, 1985, for a summary of the Brenner Debate). More recently, Acemoglu and Wolitzky (2011) proposed that a shift in the land-to-labor ratio has two antagonistic effects on the prevalence of coercion: a *labor demand effect*, in line with Domar (1970), and an *outside option effect*, in line with neo-Malthusian theories of feudal decline (Postan, 1937; Le Roy Ladurie, 1969; North and Thomas, 1973). Recent empirical contributions support the Nieboer-Domar hypothesis: Fenske (2012) in southwestern Nigeria, Green (2014) in the Cape colony, and Klein and Ogilvie (2017) in Bohemia. Ashraf et al. (2018) are more specific about the labor demand effect: they show that the rise of capital-skill complementarity made the employment of free skilled workers more profitable, and thus facilitated labor emancipation in Prussia. In the second half of the nineteenth century, half of the province of São Paulo was still branded as *terreno desconhecido*, uncharted land, on historical maps: with plenty of free land and an open agricultural frontier, Brazil allows us to provide new evidence that is also consistent with the existence of the outside option effect.

We also contribute to the debate on the causes of the decline of the institution of slavery. Williams (1944) was among the first to argue that slavery had started to decline because it was simply no longer productive (his predecessors had promoted the role of humanistic sentiments). Although Fogel and Engerman (1974) defended that slavery was still productive in the antebellum South when it was abolished, other scholars have argued that its inefficiency condemned slavery to eventually disappear (see Sutch, 2018, for a related discussion). Still, this begs the question as to why the abolitionist movement managed to assemble a winning coalition at a time where slavery was still profitable to many members of the elite. Engerman (1973) suggested that once a coercion system is established, strong incentives for its perpetuation arise to avoid capital loss for the slave-holders. In this paper, we argue that even if the elite collectively profits from the institution of slavery, it may face a coordination problem that eventually leads to abolition. Brazilian coffee producers had little international competition at the time. The country became the world's largest coffee exporter in 1831 (Nützenadel and Trentmann, 2008; Klein, 2010), and it supplied 80% of the world's coffee well into the 1920s. Even in the slavery-intensive coffee-growing regions, some planters felt that slavery profited their domestic competitors comparatively more than would 'free' labor. This is not to say that slave-holders were not conscious that slavery was condemned to disappear eventually. Still, our empirical results suggest that among slave-holders, those who could count on alternative forms of labor and those who were facing higher costs of coercion found it in their best interest to hasten the transition to free labor.

Finally, our work is connected to a growing literature that examines the relationship between elite behavior and coercive institutions (and long-run development). Naidu and Yuchtman (2013) examined how labor demand shocks influenced working conditions under coercive legislation in industrial Britain. Bobonis and Morrow (2014) investigated the influence of labor coercion on human capital accumulation and show that, consistently with a reduction of plantation workers' outside options, positive shocks to coffee prices decreased public expenditure on education in nineteenth-century Puerto Rico. Dippel et al. (2020) documented how export-driven demand shocks may incentivize elites to invest in coercive institutions, and Carvalho and Dippel (2020) studied how elite identity relates to political accountability in Caribbean plantation islands. Another strand of the literature documents the persistent influence of labor coercion institutions and their extinction on development (Nunn, 2008; Nunn and Wantchekon, 2011; Dell, 2010; Acemoglu et al., 2012; Acharya et al., 2016; Bertocchi and Dimico, 2014; Markevich and Zhuravskaya, 2018). In Brazil, Summerhill (2010) found little association between colonial institutions, including slavery, and contemporary outcomes, but Naritomi et al. (2012) found that differences in sub-national colonial institutions do matter for development. More recently, Papadia (2019) investigated the influence of slavery on fiscal capacity in Brazil's main coffee provinces, and Fujiwara et al. (2019) showed that slavery had a persistent influence on contemporary income inequality. In this quickly growing literature, that focuses mostly on the mechanisms by which institutions persist, we make a contribution by documenting the mechanisms by which institutions change, using what is arguably one of the most important episodes of institutional change in modern history.

3 Abolition laws in Brazil

By 1807, the United Kingdom and the United States had abolished the Atlantic slave trade, and started pushing other countries to do the same. The oligarchical political system of the Empire of Brazil, founded in 1822, enfranchised a limited elite, and within the elite, historians argue that the interests of slaveholders dominated. Among many similar statements, we can quote Conrad (1972, p. 16): "[M]uch of the real power in the provinces was in the hands of the slaveholding landlord class," or Viotti da Costa (1989, p. 179): "[P]oliticians often represented in the Chamber, the Senate, or the Council of State the interests of plantation owners and merchants to whom they were tied by links of patronage and clientele." The young empire had accepted to ban the trade in 1831, but it took 20 years (and forceful action by the British crown) to get Brazil to effectively act against it (de Alencastro, 1979). Another 20 years later, the 1871 *Lei do Ventre Livre* (that liberated children born of enslaved mothers) temporarily placated abolitionist sentiments stirred by the American Civil War (Conrad, 1972), with little immediate effect. Despite the international and the domestic pressure, the political context was hardly favorable to abolition.

The question was brought back to the forefront of the legislative agenda in the 1880s.[6] In 1884, Emperor Pedro II—compelled to act after the rise of emancipation movements in the North—charged the liberal senator Sousa Dantas to constitute a new cabinet and to move towards emancipation. The bill he presented, known as the *Dantas project*, rallied pro-slavery interest groups (Ridings, 1994). In the ensuing parliamentary crisis (Conrad, 1972; Viotti da Costa, 1989), Sousa Dantas was ousted and replaced by a cabinet more amicable to slave-holders' interests.

The new cabinet proposed the *Saraiva-Cotegipe* bill that emancipated slaves over 60 years old. Despite opposition from some liberals disappointed that the law did not go far enough, from some conservatives (mostly from Minas Gerais and Rio de Janeiro) opposed to any change in the existing institution, and from the hardcore slave-holders from the Paraíba valley, the bill was adopted with 81% of the votes in 1885. The *Lei dos Sexagenários* marked the point where abolition started gaining supporters in the Centro-Sul (see Figure 2). Scholars disagree about the fundamental causes of São Paulo's gradual conversion to abolitionism. Morse (1958), Graham (1968) and Dean (2012) proposed that new planters were progressive and keen on turning to immigrant workers as a substitute to coerced ones. However, Conrad (1972) observes that the growth of the slave population was largest in newly cultivated areas of São Paulo (see also Lowrie, 1938). Be that as it may, and although Paulista representatives were still cautious in 1885, their conversion played a fundamental role in the abolition of slavery.

In the late 1880s, emancipation movements radicalized, with frequent rebellions,

[6] We consider the 1880s, and more precisely, the three legislatures after the electoral reform in 1881 (*Lei Saraiva*, Jan 9, 1881), during which voting rights and the electoral map are homogeneous. This period ends with a military coup on Nov 15, 1889, which established the first Brazilian Republic.

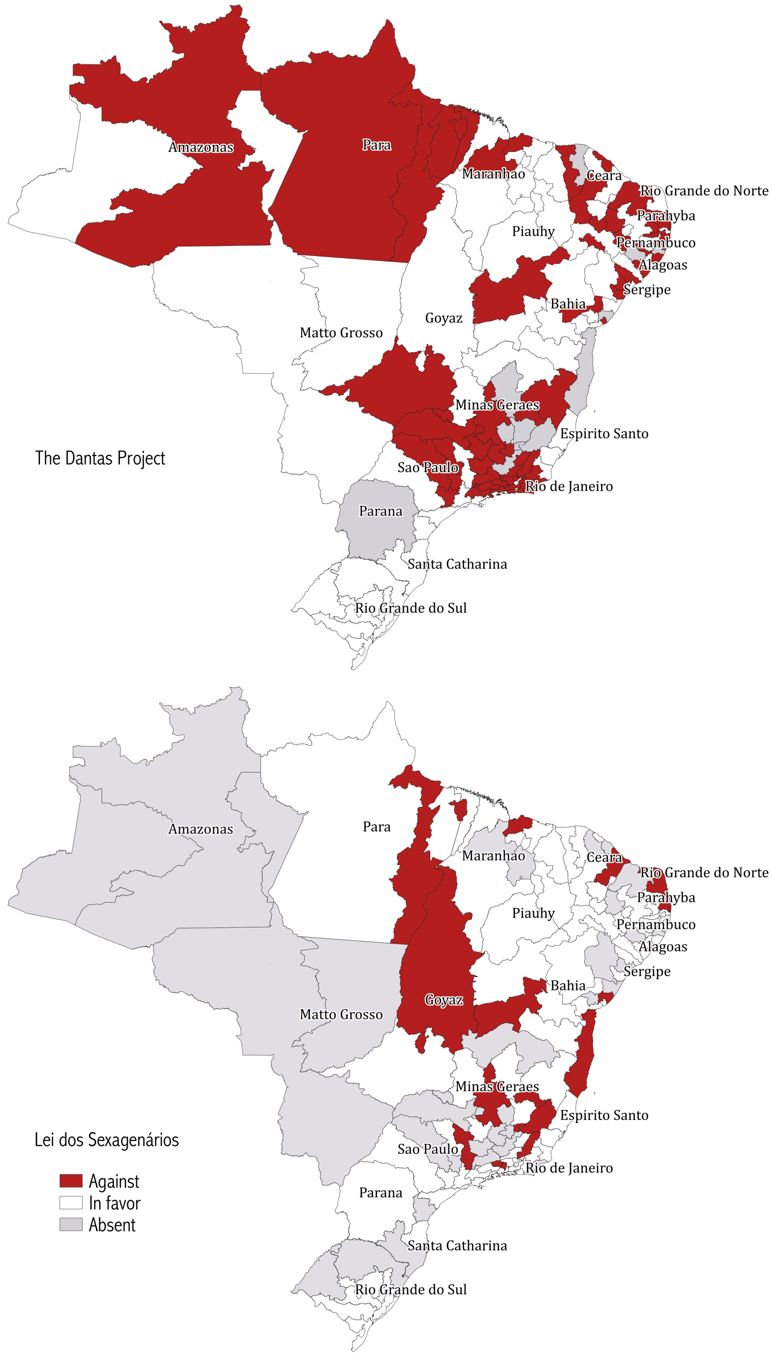


Figure 2: Vote by district on the 1884 Dantas Project and the 1885 Lei dos Sexagenários.

scenes of violence and flights from plantation (Conrad, 1972; Reis and Reis, 1988). Military aid was sent to (reluctantly) help persecute runaways (Toplin, 1969). At the same time, efforts to attract European immigrants to work on the plantations started paying off. 6,500 immigrants entered São Paulo in 1885, 32,000 in 1887, and 90,000 in 1888 (Conrad, 1972). It is under these circumstances that many Paulista planters converted to abolitionism (Luna, 1976), and liberated around 100,000 slaves in the first months of 1888. By early 1888, slavery was almost extinguished in the province (and coffee production was continuing almost unperturbed). Other provinces followed in the steps of São Paulo, leaving only planters in Rio de Janeiro and a few recalcitrant latifundiários from São Paulo and Minas Gerais to defend the coercion system.

The legislative session that opened in May 1888 had one priority: bring a definitive solution to the question of emancipation. A bill proclaiming the immediate abolition of slavery in two short articles was voted in the Chamber on May 9. As illustrated in Figure 3, Rio de Janeiro was by spring 1888 the very last bastion of slavery in the Empire. Of the nine legislators that voted against the bill, eight were representatives from the province's electoral districts. At the time of abolition, Rio had remained unaffected by immigration and its planters were threatened by bankruptcy, with little more wealth than that represented by their slaves (Conrad, 1972; Viotti da Costa, 1989). The bill was sanctioned by over 90% of representatives and passed by the Senate a few days later. It was soon approved by the Princess Regent as the Lei Áurea.

4 Data

We use archival records, census surveys, historical maps, and geo-referenced data sources to conduct our empirical analysis. In order to extract information from historical sources, we use a combination of geo-referencing, optical character recognition, text mining, and manual coding when document quality leaves no other option. Except for votes, we aggregated most of our variables from the level of the municipality to the district (642 municipalities in 1872 to match with the 122 districts after the 1881 electoral reform). Table 1 provides summary statistics of our main variables. Appendix A.5 gives more details about the construction of our variables.

Political variables. We collected the vote (or absence) of all legislators on each instance of the thirteen roll-call votes related to the emancipation of slaves from the onset of the eighteenth legislature (1882) to the end of the twentieth legislature (1889) in the Câmara dos Deputados. Together, they constitute the universe of relevant

Figure 3: Vote by district on the 1888 *Lei Áurea*.

votes,[7] starting with the first no-confidence vote against the Dantas Cabinet in 1884 and ending with the vote for the Lei Áurea in 1888. Jobim and Porto (1996) report the post-1881 district-level electoral division, which we geo-reference using IBGE (2010). Nogueira and Firmo's (1973) encyclopedia of parliamentarians allows us to match each legislator to a unique district, and helps us control for individual legislators' characteristics, such as their party affiliation.

Demographic variables. One of our main explanatory variables is the share of slaves in a district's total population in 1872, which we compute from the first nation-wide demographic census in the country's history (Brazil, 1874). We also use data on the nationality and religion of foreigners from the 1872 and 1890 censuses (Brazil, 1895). In the 1872 census, we use two variables as controls: *i)* the share of the non-captive colored population (defined as the sum of blacks, brown-skinned and mixed-race), in case they helped the cause of enslaved colored, and *ii)* the literacy rate, in case abolitionism was driven by better educated citizens. These variables are

[7]Other bills for which the vote of individual representatives is recorded were hard to connect to the issue of abolition. In particular, we identified only one roll-call vote on an issue related to immigration, on July 22, 1885. It concerned a levy to finance immigration, opposed both by moderate abolitionists and hardcore slaveholders.

Table 1: Summary statistics

Statistic	N*	Mean	St. Dev.	Min	Max
Political variables (by district × vote)					
Abolitionist vote	1,284	0.56	0.50	0	1
Absence on roll call day	1,586	0.19	0.39	0	1
Liberal	1,586	0.47	0.50	0	1
Demographic variables (by district)					
Share of slaves in 1872	122	0.147	0.091	0.013	0.534
Share of foreigners in 1872	122	0.025	0.047	0.000	0.267
Share of foreigners in 1890	122	0.019	0.040	0.000	0.237
Share of free colored in 1872	122	0.476	0.163	0.105	0.829
Share of literates in 1872	122	0.161	0.065	0.080	0.362
Colonial settlements during 1748-1800	122	0.11	0.59	0	4
Colonial settlements during 1800-1870	122	0.16	0.66	0	4
Colonial settlements during 1870-1930	122	0.94	3.64	0	36
Colonial settlements during 1930-1970	122	0.76	3.47	0	28
Coercion variables (by district)					
Av. number of quilombos	122	4.29	6.67	0.0	52.0
Av. dist. to closest 1-quilombo municipality	122	21.93	29.00	0.0	196.6
Av. dist. to closest 2-quilombos municipality	122	48.96	85.81	0.0	737.1
Av. dist. to closest 3-quilombos municipality	122	67.67	92.42	0.0	737.1
Av. dist. to closest 4-quilombos municipality	122	76.82	92.48	0.0	737.1
Av. dist. to closest 5-quilombos municipality	122	86.58	92.94	0.0	737.1
Av. dist. to closest 6-quilombos municipality	122	111.1	111.2	0.0	861.4
Av. dist. to closest 7-quilombos municipality	122	117.8	112.8	0.0	861.4
Av. dist. to closest 8-quilombos municipality	122	118.8	112.7	0.0	861.4
Av. dist. to closest 9-quilombos municipality	122	138.3	121.5	0.0	861.4
Av. area quilombola	122	112.7	729.3	0.004	7,197.7
Total area quilombola	122	622.9	3,337.5	0.01	29,712.5
Geographical variables (by district)					
Average coffee suitability	122	26.09	12.22	0.00	49.18
Average sugarcane suitability	122	24.63	8.93	6.12	52.03
Terrain ruggedness index	122	0.43	0.34	0.04	2.31
Average rainfall	122	1,409.2	374.3	708.7	2,751.4
Average latitude	122	−42.78	5.63	−60.99	−34.88
Average longitude	122	−14.60	8.04	−31.55	−1.11
Distance to the coast	122	242.6	312.0	12.1	1,825.9
Av. travel time to nearest prov. capital	122	51.61	41.68	0.0	192.3
Av. pop. density	122	19.73	39.98	0.14	232.3
Miscellaneous variables (by district)					
Av. distance to nearest gold supply road	122	229.2	276.2	2.1	1,379.9
Distance to nearest diamond mine	122	657.9	369.8	64.7	1,594.4
16th cent. Indigenous enslavement indicator	122	0.17	0.38	0	1
17th-18th cent. Indigenous enslavement indicator	122	0.49	0.50	0	1

* 122 districts × 13 laws = 1586 obs. Dist. in km, surf. in km^2, time in hr, dens. in km^{-2}.

aggregated from the municipality to the district level.

Quilombos. Quilombos were communities founded by maroons as early as the sixteenth century (most of them before the nineteenth), that offered refuge to other runaway slaves (Anderson, 1996). For each district, we compute the number of quilombos and the total area occupied by quilombos in the district with data from Fundação Palmares (2020) on all certified quilombos, as well as shapefiles from IN-

CRA (2020). We also consider the distance from each municipality's head town to the closest municipality with at least n quilombos, which we average at the level of the district to better account for the actual outside option of slaves.[8]

Geographical variables. We use IIASA/FAO (2012) to build measures of rainfall and land-suitability for sugarcane, coffee and cotton aggregated from the municipality to the district level, in combination with 1872 municipality boundaries from IBGE (2010). We compute each district's area, population density in 1872, latitude and longitude of districts' centroids, as well as measures of remoteness (notably the distance from each district's centroid to the coast). We also use Nunn and Puga's (2012) data on terrain ruggedness (TRI) at the 30"×30" grid level to construct a measure of each district's TRI (Riley et al., 1999), and the Human Mobility Index (HMI) developed by Özak (2010, 2018) to compute the average travel time across a 1km × 1km cell within a district and the average minimum travel time between each municipality's head town and the closest provincial capital. These variables are used both as individual controls and to build an instrument for the location and size of quilombos.

Miscellaneous. We geo-reference several maps from the Atlas Histórico do Brasil (CPDOC, 2016) in order to compute a number of variables: *i)* the distance to the closest supply line to eighteenth-century mining areas, averaged across municipalities, *ii)* the distance from each district's centroid to the closest eighteenth-century diamond mine, *iii)* an indicator variable capturing zones of 18th century gold mining, and *iv)* indicator variables and surfaces of areas where Indigenous peoples were repressed and enslaved between the sixteenth and the eighteenth centuries.

[8]The distribution of quilombos across districts is skewed. This explains the pattern of maximum average distances as n increases (the first district of the region of Amazonas is responsible for the maximum values of the distance for all $n \geq 2$, and the second for $n = 1$). The largest quilombo in the Fundação Palmares (2020) data set is Tambor, in the municipality of Manaus, first district of Amazonas, measuring an impressive 7197 km^2. The second one, Kalunga, in the municipality of Cavalcante, second district of Goyaz, is significantly smaller, still measuring 2618 km^2. A number of other works use quilombos, albeit in a different context. Fujiwara et al. (2019) and Papadia (2019) use quilombos to measure slavery. Closer to our interpretation (and in line with Schwartz, 1992), Lambais (2020) uses quilombos as a measure of slave resistance, whose long-run effects on economic development are the main object of his study.

5 Empirical approach

To capture the two channels discussed above, we formulate three testable hypotheses. We expect that support for slavery went hand in hand with its prevalence at the municipality/district level. This is at the heart of the labor demand effect, although it is also possible that this would simply reflect an aversion to the capital loss that emancipation would imply. A secondary but less ambiguous implication of the labor demand effect is that labor scarcity encouraged coercion. We expect that support for slavery was stronger if there were fewer immigrant workers available locally as an alternative source of manual labor. Finally, in line with the outside option effect, we expect that support for slavery in high-prevalence districts was stronger if it is harder for slaves to escape.

We summarize these hypotheses using the equations in the next section. Our baseline specifications rely on a linear probability model and within estimators. We then introduce our instrumental variable strategy to establish the plausibly causal interpretation of our results.

5.1 Empirical specifications

We start with a simple behavioral model describing legislators' decisions on abolition bills. During each vote v, a representative of district i in province j faces a binary choice: voting in favor or against abolition (respectively $P_{ijv} = 1$ and $= 0$). The relative preference of the representative for abolition is $U^1_{ijv} - U^0_{ijv} = \delta_v + \zeta_j + \tilde{\mathbf{x}}'_{ijv}\boldsymbol{\theta} - \varepsilon_{ijv}$, with $\tilde{\mathbf{x}}_{ijv}$ a vector of observables related to the representative's decision, ζ_j and δ_v province and vote fixed effects [FE], and ε_{ijv} is an unobserved idiosyncratic preference shock against abolition. Then,

$$
\begin{aligned}
\mathbb{E}[P_{ijv}|\tilde{\mathbf{x}}_{ijv}, \zeta_j, \delta_v] &= \mathbb{P}(P_{ijv} = 1|\tilde{\mathbf{x}}_{ijv}, \zeta_j, \delta_v) \\
&= \mathbb{P}(\varepsilon_{ijv} < \delta_v + \zeta_j + \tilde{\mathbf{x}}'_{ijv}\boldsymbol{\theta}) \\
&= F(\delta_v + \zeta_j + \tilde{\mathbf{x}}'_{ijv}\boldsymbol{\theta}).
\end{aligned}
$$

where we may specify F as *i)* the identity function, in which case this becomes a linear probability model with FE, or *ii)* a normal distribution, in which case this is a probit model and δ_v and ζ_j are parameters to be estimated (which should not be an issue in this setting with twenty provinces and thirteen periods). For each of the following hypotheses, we envisage both approaches.

Hypothesis 1. *The likelihood that a legislator votes in favor of abolition bills decreases with the local prevalence of slavery.*

To test hypothesis 1, we reformulate the previous equation as:

$$F^{-1}\{\mathbb{E}[P_{ijv}|\tilde{\mathbf{x}}_{ijv}, \zeta_j, \delta_v]\} = \zeta_j + \delta_v + S_{ij}^{1872}\beta + \mathbf{x}'_{ijv}\boldsymbol{\gamma}, \tag{1}$$

where S_{ij}^{1872} measures the share of slaves in the population of district i. The vector $\mathbf{x}_{ijv}$ includes both legislator-vote-level and district-level attributes. In particular, it controls for the party of the representative of district i at the time of vote v, and an array of geographic and demographic descriptors of district i (ethnicity and literacy rates, population density, crop suitability, rainfall, distance to the coast, latitude and longitude). Hypothesis 1 corresponds to $\beta < 0$.

Hypothesis 2. *The likelihood that a legislator votes in favor of abolition bills increases when there are local alternatives to slave labor.*

To test hypothesis 2, we reformulate the same equation as:

$$F^{-1}\{\mathbb{E}[P_{ijv}|\tilde{\mathbf{x}}_{ijv}, \zeta_j, \delta_v]\} = \zeta_j + \delta_v + S_{ij}^{1872}\beta + F_{ij}^{1890}\lambda + \mathbf{x}'_{ijv}\boldsymbol{\gamma}, \tag{2}$$

where F_{ij}^{1890} captures the share of foreigners in the population of district i. Our focus is on the intra-elite conflict: what we predict is that slave-holders who can rely more on immigrant labor, or expect to be able to do so in the near future, are more willing to abolish the institution of slavery. All other notations are the same as for hypothesis 1. We control for the presence in immigrants in 1872, and in an alternative specification, we use immigration between 1872 and 1890 as our main explanatory variable. Hypothesis 2 corresponds to $\lambda > 0$.

Hypothesis 3. *The likelihood that a representative from a high-prevalence area votes in favor of abolition bills increases when it is easier for slaves to escape.*

To test hypothesis 3, we proceed similarly and use the specification

$$F^{-1}\{\mathbb{E}[P_{ijv}|\tilde{\mathbf{x}}_{ijv}, \zeta_j, \delta_v]\} = \zeta_j + \delta_v + S_{ij}^{1872}\beta + O_{ij}\phi + S_{ij}^{1872} \times O_{ij}\psi + \mathbf{x}'_{ijv}\boldsymbol{\gamma} \tag{3}$$

where O_{ij} captures the proximity of freedom for slaves in district i and, correspondingly, the cost of coercion for slave-holders. Our measure of the distance to freedom for escapees relies on the location of quilombos. The marginal effect of the proximity of freedom is $\partial P_{ijv}^{1884-1888}/\partial O_{ij} = \phi + S_{ij}^{1872}\psi$. Again, our focus is on the intra-elite conflict: what we predict is that slave-holders who face higher costs of coercion have a competitive disadvantage relative to other slave-holders. Consequently, hypothesis 3 corresponds to $\psi > 0$.

5.2 A discussion of identification

Identification is complicated for each of our three main specifications by the non-random allocation of slaves, immigrants, and quilombos across districts, and by possible measurement errors. The coefficient β is a biased estimate of the effect of vested interests on voting patterns if some districts have fewer slaves because of long-standing abolitionist beliefs, or if other omitted variables (for instance deeply rooted norms) determine abolitionism and the local prevalence of slavery simultaneously. Similarly, it is apparent from the annals of the Câmara dos Deputados that questions of immigration and abolition were closely linked. From the 1880s onward, the planters themselves exerted much effort to attract immigrant workers (Conrad, 1972), so that λ may also be a biased estimate of the effect of immigrants on voting patterns. Finally, the coefficients ϕ and ψ are biased estimates of the effect of evasion on voting patterns if quilombos tend to be located in abolition-friendly districts.

To address this, our baseline specifications exploit within-province variation, which means that variations in norms, culture and other unobserved heterogeneity across provinces are absorbed by province FE. Additionally, we capture variation over bills (constant across districts) with vote FE. Hence, a causal interpretation of our coefficients of interests is equivalent to assuming that the allocation of slaves, immigrants, and quilombos *within provinces and bills* is as good as random, conditionally on a wide array of political, demographic and geographical controls. FE and controls do a great deal to alleviate endogeneity concerns, but they might not be sufficient to guarantee the validity of this assumption (e.g. there might exist within-province cultural norms driving both the allocation of slaves and legislators' voting behavior).

We address these remaining threats to identification using a two-pronged instrumental variables strategy. We build three sets of 'standard' instrumental variables (one per endogenous variable), respectively leveraging *i)* long-predating historical determinants of the location of slaves across Brazil, *ii)* preexisting migration networks, and *iii)* topographic determinants of the location of quilombos (sections 5.2.1 through 5.2.3). Following the procedure of Lewbel (2012), we then supplement these instruments with heteroskedasticity-based instruments built as simple functions of the model's data (section 5.2.4).

5.2.1 Instrument for the prevalence of slavery

Our instrument for the share of slaves in equation 1 interacts the average distance from municipalities' head towns to the nearest eighteenth-century caminho do ouro with a scaling variable. The discovery of large deposits of gold towards the end of the seventeenth century (soon followed by diamonds) in the province of Minas Gerais (literally, 'General Mines') justified, in the eighteenth century, the construction of these trade routes between the province and coastal areas throughout the country (Zemella, 1951). These routes also helped ensure the continuous supply of mining areas in slaves (from the north) and cattle (from the south). They were built by slaves, making the distance to these roads an interesting historical shifter of the prevalence of slavery into the nineteenth century (Klein and Luna, 2009). We may think of the distance to these roads as capturing historical slave-related activity. In general, proximity to caminhos do ouro is thus an indicator of higher slavery prevalence.[9] However, this relationship becomes significantly weaker within provinces, once fixed effects are added, in particular because the distance to these roads loses economic meaning within late-settlement provinces, many of which were not significantly populated at the time of the Gold Rush.

To circumvent this issue, we thus have to scale the distance to caminhos do ouro with a variable that captures, within provinces possibly distant from the roads, which municipalities were more likely to have already been settled in the 17th century. Our preferred scaling variable measures the repression and enslavement of Indigenous peoples in the 16th century. When sixteenth-century Portuguese settlers were laying the foundations of the plantation system, they first started experimenting with an enslaved Indigenous labor force. At its peak in the 1560s, Indigenous slavery counted tens of thousands of individuals. Because of a combination of widespread epidemics, continuous conflict with free Indigenous peoples, and increasing discomfort from the Crown with Indigenous enslavement after the Valladolid debate, the

[9]Figure A.5a in the Appendix A.2 provides an unconditional plot of this relationship.

Portuguese turned away from Indigenous to African slave labor (Klein and Luna, 2009). In other words, areas where Indigenous peoples where repressed and driven out during the early days of the colonization of Brazil are more likely to have received settlers and slaves.

Therefore, our instrument predicts a lower prevalence of slavery as distance to caminhos do ouro increases, except in places where Indigenous repression and enslavement were historically intensive, which—compared to other remote municipalities within provinces—were more likely to receive slaves.[10] To be valid, this instrument should not be correlated with the decision to vote in favor or against abolition in the 1880s, except through its influence on the local prevalence of slavery. We make several observations.

First, the coffee planter elite's interests that played a fundamental role in perpetuating the coercion system in the second half of the nineteenth century are plausibly orthogonal to both Indigenous repression in the sixteenth century and mining interests in the eighteenth century. The emergence of the mining interests had large consequences on the configuration of economic activity, even leading to the relocation of the colony's capital from Salvador to Rio de Janeiro in 1763. However, the economy of the region was declining in the early nineteenth century, only to be revived in its second half by the cultivation of coffee and, to a lesser extent, sugar.

Nonetheless, we may still be concerned that Indigenous repression and/or the location of the caminhos do ouro affected voting decisions in the 1880s, through other channels than the prevalence of slavery. In particular, our main concern is that caminhos do ouro may have had a persisting impact on economic activity, other than trough slavery. If this affected voting decisions, our identification assumption would be violated. However, caminhos do ouro lost their importance passed the decline of the gold rush at the end of the 18th century. They were no longer maintained and became free after Brazil became independent in 1822. In addition, Indigenous peoples were almost entirely driven out from early settlement municipalities, so that our scaling variable is unlikely to affect voting decisions other than through settlement patterns. Most importantly, we systematically control for both province and vote intercepts, in addition to a wide range of controls capturing baseline year economic activity (notably land suitability controls for the main exports crops, population density, human mobility, mining areas, and distance to mines, to rivers and to the coast). In table 12 of Appendix A.1.2, we implement a falsification test and show

[10]See Map A.9 in Appendix A.3 for a visual representation of the instrument.

that neither caminhos do ouro, nor Indigenous repression, nor their interaction seem to affect economic activity in the long run other than via slavery.

An added benefit of combining two sources of variation—aside from the fact that it produces a variation that is unlikely to influence abolitionism other than through slavery—is that we can flexibly evaluate the variation driving our instrument. In table 10, we show that the un-interacted terms themselves do not matter in explaining slavery, and that the interaction term remains almost numerically identical when un-interacted terms are dropped (in which case the instrument cannot be interpreted other than as the distance to gold paths weighted by Indigenous repression), which results in a stronger instrument (because un-interacted terms have no explanatory power). The instrument also remains consistent when un-interacted terms are only used as controls, so that the only variation underlying the instrument comes from the differential effect of proximity to gold paths in zones of 16th century Indigenous repression compared to other areas (not distance or repression themselves). We also consider alternative scaling approaches, in particular extending the Indigenous repression period to the 18th century.

Finally, we might worry that a 'preference for slavery' or a 'culture of abolitionism' predated the enslavement of Indigenous peoples and the establishment of the caminhos do ouro, and persisted into the nineteenth century. This seems implausible: it was northern districts that held the highest number of slaves during the sugar boom, well into the nineteenth century. Unlike in the United States, slavery was never defended in Brazil on the grounds that it was a positive institution (Klein and Luna, 2009). Even the slaveholding elite appeared to defend slavery mostly as a necessary evil.

5.2.2 Instrument for the distribution of immigrants

Shift-share instruments, which exploit heterogeneous exposure to common shocks, have become common to address the endogeneity of immigrants' location (Felbermayr et al., 2010; Andersen and Dalgaard, 2011; Ortega and Peri, 2014; Alesina et al., 2016; Docquier et al., 2016; Beine and Parsons, 2017; Bahar and Rapoport, 2018; Burchardi et al., 2019; Docquier et al., 2020; Tabellini, 2020). Such instruments are typically constructed by interacting past immigration disaggregated by nationality shares with inflows of immigrants from each sending country.

A particular difficulty in our case is that the 1890 census does not decompose immigration by nationality. We circumvent this difficulty thanks to the rich information on *religion* shares in each municipality, and based on the empirical regularity that the native population in 1890 was largely Roman Catholic.[11] We build a leave-out religion-based Bartik-like instrument, based on a matching of 1872-1890 predicted nationality-by-religion shares. We approximate the number of immigrants $\widehat{I}_{ir1890}$ of religion r in district i by the number of individuals of religion r in district i for every religion except Catholicism, and the number of Catholic foreigners by the difference between the number of Catholics and the native population in each district. Now, since the 1872 data provides a decomposition of the immigrant population by nationality, but only distinguishes Catholics and non-Catholics, we match 1872 nationalities with the dominant religion in each country of origin (e.g. non-Catholic German immigrants are counted as Protestant, Catholic Greeks as Orthodox etc. See Table 22 in Appendix A.2 for details). We can then compute the share $\widehat{\alpha}_{ir1872}$ of foreigners of religion r in district i as the sum for each origin country of the shares of such individuals. Finally, we write our religion-based instrument as:

$$Z_{i1890} = \frac{1}{P_{i1890}} \sum_r \widehat{\alpha}_{ir1872} \widehat{I}^{-i}_{r1890},$$

where $\widehat{I}^{-i}_{r1890}$ is the predicted number of immigrants of religion r in 1890, net of those that settled in district i, and P_{i1890} is the 1890 population of district i.

For this instrument to be valid, municipality-level conditions that may have affected the distribution of immigrants by religion before 1872 must be unrelated to abolition patterns in the 1880s (Goldsmith-Pinkham et al., 2020; Jaeger et al., 2018). To deal with this concern, in addition to our full vector of controls (which includes several baseline year characteristics that may have contributed to attracting immigrants, such as population density and other demographic covariates), we systematically control for 1872 immigration in our IV specifications, which not only mechanically predicts higher future immigration via the instrument, but may also have a distinct effect on emancipation-related voting (Tabellini, 2020). The variation used to identify the in-

[11] Roman Catholicism held a hegemonic position during the period, losing terrain only slowly and mostly after the advent of the First Republic. Article 5 of the 1824 constitution established Roman Catholicism as the state religion of the Empire of Brazil. Although Protestantism was introduced in the country only shortly after Independence, with the establishment of Swiss and German colonial settlements, these first waves of Protestant immigration remained both culturally and geographically isolated, and did not threaten the position of the Church (Mendonça, 2003). North-American missions posed a larger threat, with only relative success initially, although increasingly toward the end of the period.

fluence of 1890 immigrants is therefore in the religious composition of municipalities' foreign populations, not in the actual size of the immigrant population. In Appendix A.1.2, we also address the concern that immigrants from specific religions or nationalities selected their destinations based on the possibility that these would be more inclined to vote for emancipation by controlling for individual shares and for *núcleos coloniais* (state-sponsored settlements, also see Appendix A.4). We also consider alternative ways to construct and scale the instrument, and show that results are stable across a wide range of specifications. Finally, both the conflation of responses to immigrants' arrival put forward by Jaeger et al. (2018) and the possibility that inflows of immigrants from each sending country maybe responding to municipality-specific conditions appear unlikely in our context. Immigration increased vastly in the late 1880s, and responded to altogether different incentives, notably to a proactive immigration policy that was largely decided in Parliament as a manner to provide an alternative source of manual labor when abolition had become a probable outcome. Even the location of migrants was partly determined in Parliament: immigrants in Brazil were more often than not debt-bonded laborers, credit constrained and tied to the land (Rocha et al., 2017; Witzel de Souza, 2019).

5.2.3 Instrument for the location of quilombos

Our main instrument for the location of quilombos in equation 3 is the average terrain ruggedness, interacted with the average travel time to the provincial capital. There is little doubt that the decision of where to establish quilombos was driven by considerations of security and remoteness. Nunn and Puga (2012) showed that in Africa, terrain ruggedness discouraged slave trades and facilitated escape. In Brazil, Klein and Luna (2009, p. 195) point out that a permanent escape depended on "the existence of dense forests or inaccessible mountains within a short distance from their homes." Remoteness (which we measure as travel time from the provincial capital), in turn, made repression harder, and the establishment of successful (and surviving) quilombos more likely. We provide an unconditional plot of the relationship between the location of quilombos and the instrument in Figure A.5c in the Appendix A.2. Ruggedness and remoteness tend to act as substitutes for the successful establishment of durable quilombos: in the proximity of large settlement areas, rugged terrain was critical in providing a defensive advantage and allowing to remain hidden, whereas it played a much less deciding role at a safe distance.

To be valid, this instrument should not be correlated with the decision to vote in favor or against abolition in the 1880s, except through its influence on the location of quilombos. A concern is that the TRI tends to be negatively associated with economic outcomes at the country level, because it increases transport costs and makes trading more difficult (Nunn and Puga, 2012). This is however unlikely to be an issue within provinces, and we systematically control for determinants of economic activity (in particular population density and soil suitability) and other shifters of remoteness (in particular distance to the coast, and average human mobility index). Moreover, in Tables 18 and 19, we show that results remain qualitatively identical whether un-interacted terms are used as instruments or controls, so that only variation in the differential effect of ruggedness by remoteness is used to instrument the location of quilombos.[12] In Tables 18 and 19, we also address the concern that the instrument might be picking up nonlinearities and flexibly control for nonlinearities in un-interacted terms. Perhaps most importantly, Table 20 provides a placebo test of the instrument's exclusion restriction. Indeed, a key feature of the empirical design of Hypothesis 3 is that quilombos have no effect on legislators' voting decisions unless interacted with the prevalence of slavery. If the exclusion restriction does not hold and the instrument affects voting decisions through other channels than by determining the location and size of quilombos, we may expect direct effect of ruggedness, remoteness, and their interaction on abolition voting. In Table 20, we present reduced form results showing that this is not the case, lending further support to the instrument's validity. Finally, a remaining possible concern is that quilombos may be viewed as a 'bad control', to the extent that the number of quilombos and their size can conceivably constitute an outcome of slavery. Because we are not overly concerned with a possible reverse causation effect of abolitionism on quilombos (most quilombos predate by far the nineteenth century), we believe that this is unlikely to be an issue. Nonetheless, in Figure A.4, we show that slavery and quilombos are at best marginally correlated. Quilombos were widespread and relatively common across the country, and largely independent from the distribution of slaves in 1872 once controls and fixed effects are taken into account.

[12] One might expect ruggedness and remoteness to be strongly positively correlated, but this turns out not to be the case. Ruggedness and remoteness rather tend to be weakly negatively correlated, and have a positive but non-significant relationship within provinces.

5.2.4 Heteroskedasticity-based instruments

Instruments generated from the data are generally not a silver bullet to identification issues, but they help with within-province identification and significantly improve the efficiency of our estimations.[13] This approach is especially appropriate with triangular systems, such as what arises in classical measurement error frameworks (a common issue when dealing with nineteenth century census data and with imperfect proxies) and omitted variables (we may be concerned that deeper cultural norms influence both the prevalence of slavery and subsequent voting decisions).

In our context, the use of heteroskedasticity-based instruments provides three main advantages. First, they allow us to run overidentification tests whenever our main instruments would otherwise leave regression models just-identified. Second, they improve the precision of our estimates in cases where standard instruments are weaker and we may be concerned about weak identification. Third, and perhaps most importantly, they provide a uniform test for the validity of our main instruments. In a context with multiple instrumental variable strategies and limited possibilities to run placebos, it is helpful to compare estimates obtained from completely different identification assumptions, and reassuring that they yield numerically similar estimates (Baum and Lewbel, 2019).

To build the instruments, we run the following 'zero-stage' regressions:[14]

$$\begin{aligned} S_{ij}^{1872} &= \mathbf{x}'_{ijv}\boldsymbol{\sigma}^S + \delta_v^S + \zeta_j^S + \xi_{ijv}^S \\ F_{ij}^{1890} &= S_{ij}^{1872}\beta^F + \mathbf{x}'_{ijv}\boldsymbol{\sigma}^F + \delta_v^F + \zeta_j^F + \xi_{ijv}^F \\ O_{ij} &= S_{ij}^{1872}\beta^O + \mathbf{x}'_{ijv}\boldsymbol{\sigma}^O + \delta_v^O + \zeta_j^O + \xi_{ijv}^O \end{aligned} \tag{4}$$

from which the estimated residuals are used to create instruments as $(\mathbf{x}_{ijv} - \overline{\mathbf{x}}_{ijv})'\hat{\xi}_{ijv}$, where $\overline{\mathbf{x}}_{ijv}$ is a mean-centering vector.[15] Identification requires that the error terms

[13] Dietrich and Wright (2015) and Fails (2019) also used a similar two-pronged identification strategy to analyze economic determinants of institutional change. On the issue of slavery, Bezemer et al. (2014) used a similar identification strategy to examine the long run development outcomes associated with indigenous slavery in Sub-Saharan Africa. Also related, Depetris-Chauvin and Özak (2020) used a similar strategy to analyze the early determinants of economic specialization. On identification using moment restrictions, see Magnusson and Mavroeidis (2014) and Lewbel (2019).

[14] We call these auxiliary regressions 'zero-stage' because they are estimated as a preliminary step before running proper first-stage regressions with standard instruments. Also note that whenever a second-stage regression includes an interacted term (ie. a second endogenous variable) we run additional zero-stage regressions (and first-stage if applicable) with interactions as dependent variables.

[15] The inclusion of vote FE and time-varying controls in the second-stage regression requires the inclusion of the same FE and controls in the zero-stage regression. The heteroskedasticity-based in-

of the first-stage regressions be heteroskedastic, which we verify using Breusch-Pagan tests. It also requires that the error terms of the second-stage regressions be homoskedastic. While this is hard to justify theoretically, we can at least verify empirically that their homoskedasticity cannot be rejected using Pagan-Hall tests.

In baseline specifications, we exploit heteroskedasticity with respect to geographic regressors, which we deem more likely to be 'good controls.' Out of this already large array of possible instruments, we select those that are empirically the most relevant. We use latitude to build the corresponding instrument for the share of slaves in 1872; we use latitude and distance to the coast for the share of immigrants in 1890; and geographical coordinates, distance to the coast and/or potential yield to build the corresponding instruments for functions of location of quilombos.

6 Main results

Throughout this section, we address the issue of auto-correlation using *i)* two-way clustered standard errors and *ii)* HAC standard errors computed using Conley's (1999; 2010) approach, allowing auto-correlation within a given radius from each district's centroid.[16] In our context, there are reasons to expect observations to be both spatially *and* serially correlated, and we therefore believe that clustering at both the district and period levels is the appropriate approach. Clustering at the level of the district allows all the votes of a district's representative to be correlated non-parametrically, under the parametric assumption that observations are uncorrelated across groups. Clustering at the level of the vote allows arbitrary dependence between districts for each bill. However, because our time-dimension is 'shorter' than our spatial-dimension (i.e. there are 13 bills and 122 districts) and since the asymptotic theory underlying two-way clustering relies on clusters in the smallest dimension, we may be worried that two-way clustered standard errors are too demanding, in particular for nonlinear models. Therefore, we also report one-way clustered standard errors at the district level for these. Conley's approach constitutes an alternative to clustering. Instead of assigning observations to groups, this weighs observations

struments for each of our vote-invariant explanatory variable are therefore, maybe counterintuitively, time-varying.

[16]In baseline specifications, we select a 250 km radius for the spatial kernel (in addition to allowing observations to be serially correlated across the 13 bills). In practice, results remain qualitatively equivalent when we vary this threshold from 50 to 1000km. In addition, we adapt the code written by König et al. (2017) for Stata to compute Conley SEs in a 2SLS panel setting.

based on distance on the cross-sectional dimension, and periods on the serial dimension. Conley's estimator thus yields HAC standard errors, robust to both spatial and serial autocorrelation.

6.1 Anti-abolition voting and the prevalence of slavery

In Table 2, we report the 'uninstrumented' results for hypothesis 1. Columns 1 to 6 present the result of our baseline OLS specification with province and vote FE. In column 1, we report the 'raw' influence of the prevalence of slavery on our binary abolitionist voting outcome. In column 2 to 6, we consider how this association is affected by the progressive introduction of a series of controls. Column 7 presents the results of our GLM specification with the whole set of controls and FE.

Table 2: H1 – Prevalence of slavery and voting decisions

	𝟙(Abolition vote)						
	OLS						*GLM*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of slaves	−1.210	−1.250	−0.620	−0.642	−1.142	−0.984	−3.976
	(0.383)***	(0.433)***	(0.196)***	(0.430)	(0.391)***	(0.286)***	(1.258)***
	{0.226}***	{0.300}***	{0.202}***	{0.338}*	{0.210}***	{0.404}**	{1.500}***
							[−0.990]
Controls	None	Geo.	Pol.	Dem.	Eco.	All	All
Province & vote FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,284	1,284	1,284	1,284	1,284	1,284	1269
R^2	0.244	0.255	0.399	0.249	0.245	0.411	0.356

Note: *p<0.1; **p<0.05; ***p<0.01; District-vote two-way clustered standard errors in parentheses. Conley standard errors (with a 250km window) in curly brackets for OLS specifications, and one-way district-level clustered standard errors for GLM specification. Marginal effect in brackets. Geographical controls: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, and human mobility index. Political controls: party affiliation. Demographic controls: share of free colored and of literacy. Economic controls: gold mining and distance to closest diamond mine.

In all specifications, a higher prevalence of slavery in a district is associated with a higher probability that the district's representative votes against abolitionist bills. This relationship is robust to the introduction of all controls in the OLS and in the GLM specifications. The magnitude of the effect decreases when we introduce political controls.[17] In our sample, 46% of members of the Câmara dos Deputados are Conservatives, almost 47% are Liberals (the party of abolition), and around 2% are Republicans (the remainder's affiliation – about 5% – is unknown, sometimes

[17]In Table 9 in Appendix A.2, we report these results taking only the slaves employed in agriculture as the independent variable.

because legislators simultaneously occupy Ministerial positions and are thus unaffiliated). Since liberal representatives tended to be elected in districts with a lower prevalence of slavery (and conservatives in high-prevalence districts), it is unsurprising for part of the effect observed in column 1 to be captured by political affiliation. The magnitude of the relationship also decreases when we introduce demographic controls, which suggests that coerced workers are better represented in districts with a higher share of free colored population. The GLM in column 7 yields a comparable marginal effect of the prevalence of slavery on the abolitionist vote.

According to our estimate in column 6 of Table 2, which includes the whole set of controls and FE, a 1 percentage point increase in the share of slaves in the 1872 population is associated with a .98 percentage point increase in the probability to vote against an emancipation bill. The share of slaves ranges from 0.0% in the first district of the Amazonas province to 53.4% in the tenth district of Rio de Janeiro, with a standard deviation [SD] of 9.1%: these results imply that a SD increase in the share of slaves increases a representative's probability to vote against emancipation by 8.9 percentage points. As the political affiliation of a representative is possibly endogenously determined by both the share of slaves and the vote for or against abolition (as are, to a lesser degree, demographic variables), this is probably a conservative estimate. If we believe demographic and political variables to be colliders, column 2 suggests an alternative effect of 11.4 percentage points.

Our results are also robust to instrumenting the share of slaves in 1872 by: *i)* the standard instrument described in section 5.2.1, *ii)* the heteroskedasticity-based instrument described in section 5.2.4, and *iii)* the combination of both. Columns 1 to 3 of Table 3 present the first stage of the 2SLS specification with all controls, and province and vote FE, and columns 4 to 6 the corresponding second stage estimations. Column 7 presents the conditional maximum likelihood probit [CML] estimation with the standard instrument.

In all cases, the first stage regression shows a comfortably large association between the prevalence of slavery and our set of instruments. The strength of generated instruments depends on the degree of heteroskedasticity with respect to zero-stage regressors, which we verify using a Breusch-Pagan test. This confirms the relevance of using a heteroskedasticity-based instrument to supplement our standard instrument. 2SLS estimates are larger than their OLS counterparts in column 6 of Table 2. This is consistent with an attenuation bias in our OLS estimate, as could for instance be the case if the share of slaves was mismeasured in the 1872 census, or with a possible

Table 3: H1 - Prevalence of slavery and voting decisions – Cont.

	First stage			*Second stage*			*CML*
	Share of slaves in 1872			$\mathbb{1}$(Abolition vote)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pred. share of slaves				−3.992	−2.604	−2.974	−15.530
				(1.689)**	(1.214)**	(1.011)***	(5.305)***
				{2.179}*	{1.410}*	{1.206}**	{6.137}**
							[−0.870]
Ln 16th rep. area	1.01e-5		1.16e-5				
× Dist. Gold Paths	(2.89e-06)***		(3.17e-06)***				
	{2.05e-06}***		{2.14e-06}***				
Het. instr. (latitude)		−0.381	−0.401				
		(0.110)***	(0.106)***				
		{0.161}**	{0.152}**				
Controls and FE	All	All	All	All	All	All	All
Observations	1,284	1,284	1,284	1,284	1,284	1,284	1,269
K-P F-stat	12.405	12.058	12.964				
B-P p-value		0.000	0.000				
P-H p-value					0.631	0.636	
Hansen J p-value						0.521	
Endog. test p-value				0.092	0.150	0.045	0.081

Note: *p<0.1; **p<0.05; ***p<0.01; District-vote two-way clustered standard errors in parentheses. Conley standard errors (with a 250km window) in curly brackets for 2SLS specifications, and one-way district-level clustered standard errors for the CML. Marginal effect in brackets. Columns 1-3 present alternative 2SLS first stages, columns 4-6 the corresponding second stages, and column 7 the CML (in a single step). Columns 1, 4, and 7 only use our standard instrument, columns 2 and 5 only use het.-based instruments, and columns 3 and 6 use both. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and vote FE. The Kleibergen-Paap F-stat detects weak instruments. The Breush-Pagan test checks for heteroskedasticity in the zero-stage, and Pagan-Hall in the second. The Hansen J statistic checks the validity of overidentification restrictions. The endogeneity tests (Durbin-Wu-Hausman-like for the 2SLS, robust to violations of conditional homoskedasticity, and Wald for the ML) assess whether endogenous regressors can be treated as exogenous.

omitted variable, such as the existence of local pro- or anti-coercion norms. 2SLS estimates suggest that a SD increase in the share of slaves increases a representative's probability to vote against emancipation by 23.7 (using only generated instruments) and up to 36.3 percentage points (using only the standard instrument). The conditional ML estimate is more in line with the OLS estimate (7.9 percentage points). A Durbin-Wu-Hausman-like (resp. Wald) endogeneity test suggests that the share of slaves in 1872 is not exogenous in our 2SLS (resp. CML) specifications. Our preferred estimation includes the whole set of FE, controls, and instruments (column 6), and suggests that a SD increase in the share of slaves increases a representative's probability to vote against emancipation by 27.1 percentage points.[18]

[18] In Table 10 in Appendix A.2, we consider alternative definitions of our main instrument. These alternative definitions are weaker; they support the robustness of our qualitative result, but the confidence intervals are too wide for us to report the corresponding point estimates of the effect. In Table 11, we report the CML estimation with the same alternative definitions of the instrument. All

The fact that both types of instruments correct the OLS in the same direction and by a comparable order of magnitude is an important benefit of using heteroskedasticity-based instruments. With several instruments but, given the scarcity of historical data, limited leeway to run comprehensive placebos, obtaining comparable results under two completely different identification assumptions is reassuring. Another benefit is to verify that over-identifying restrictions, i.e. the joint validity of our instruments, cannot be rejected (per the value of the Hansen J statistic).

6.2 Anti-abolition voting and immigration

In Table 4, we report the uninstrumented results for hypothesis 2. Column 1 presents the result of our baseline OLS specification for hypothesis 2 with the share of foreigners both in 1872 and in 1890, and column 2 with the net immigration between the two dates (foreigners in 1890 − foreigners in 1872, as a proportion of the 1890 population in each district). Columns 3 and 4 replicate the analysis with a probit model.

A higher presence of immigrants in 1890 (alternatively, immigration between 1872 and 1890) is associated with a higher probability that the district's representative votes in favor of abolitionist bills. This relationship is robust to the introduction of all controls in the OLS and in the GLM specifications. Our preferred specification considers separately the shares of foreigners in 1872 and 1890: it reveals that immigrants prior to 1872 did not alleviate the reliance of slave-holders on slavery. In contrast, immigrants between 1872 and 1890 made the interests of the slave-holders less salient. This is consistent with the *labor demand effect* (and part of the Niboer-Domar hypothesis): substitution possibilities away from slaves become critical once coercive labor arrangements are threatened. This is also consistent with the Brazilian historiography, following which immigrants were substituted to slaves in some late 19th century plantation areas, in particular in São Paulo. If we take the estimate of column 1, controlling for the presence of foreigners in 1872, a 1 percentage point increase in the share of foreigners in 1890 is associated with a 3.8 percentage points increase in the probability to vote in favor of an emancipation bill. The share of foreigners ranges from 0.0% in the 4th district of the province of Parahyba to 23.7% in the districts 1 to 3 of Rio, with a SD of 4.0%: these results imply that a SD increase in the share of foreigners in 1890 increases a representative's probability to vote in favor of emancipation by 15.2 percentage points (14.8 percentage points according to

yield comparable point estimates.

Table 4: H2 – Immigration and voting decisions

	𝟙(Abolition vote)			
	OLS		*GLM*	
	(1)	(2)	(3)	(4)
1872 share of slaves	−0.895	−0.803	−3.691	−3.223
	(0.253)***	(0.258)***	(1.084)***	(1.058)***
	{0.393}**	{0.398}**	{1.328}***	{1.378}**
			[−0.905]	[−0.792]
1872 share of free foreigners	−2.941		−11.916	
	(1.029)***		(4.199)***	
	{0.773}***		{4.429}***	
			[−2.920]	
1890 share of foreigners	3.821		15.193	
	(0.885)***		(3.774)***	
	{0.838}***		{3.670}***	
			[3.723]	
$\Delta_{1890\text{-}1872}$		2.332		9.676
		(0.582)***		(2.693)***
		{0.647}***		{2.309}***
				[2.377]
Controls and FE	All	All	All	All
Observations	1,284	1,284	1,269	1,269
R^2	0.421	0.420	0.417	0.418

Note: *p<0.1; **p<0.05; ***p<0.01. District-vote two-way clustered standard errors in parentheses. Conley standard errors (with a 250km window) in curly brackets for 2SLS specifications, and one-way district-level clustered standard errors for ML specifications. Marginal effects in brackets. Columns 1 and 2 present OLS results using immigrants in 1872 and 1890 and the differential between the two respectively. Columns 3 and 4 reproduce the analysis with a probit. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and votes FE.

the probit model).

Our results are also robust to instrumenting the share of immigrants in 1890 by: *i)* the shift-share instrument described in section 5.2.2, *ii)* the heteroskedasticity-based instruments described in section 5.2.4; and *iii)* the combination of both. Columns 1 to 3 of Table 5 present the first stage of the 2SLS specification with province and vote FE, and columns 4 to 6 the corresponding second stage estimations. Column 7 presents the CML estimation with the shift-share instrument.

In all cases, the first stage regression shows a large association between the share of foreigners in 1890 and our set of instruments. We easily reject homoskedasticity with respect to exogenous regressors, which confirms the relevance of generated instruments to supplement our shift-share instrument. 2SLS estimates are comparable under two completely different identification assumptions, which is once again reassuring. We also verify that overidentifying restrictions cannot be rejected, both when we

Table 5: H2 – Immigration and voting decisions – Cont.

	First stage			*Second stage*			*CML*
	1890 Share of foreigners			$\mathbb{1}$(Abolition vote)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pred. 1890 sh. of foreigners				3.401 (1.193)*** {1.091}***	3.653 (0.725)*** {0.763}***	3.667 (0.713)*** {0.708}***	15.942 (5.393)*** {5.798}*** [2.848]
1872 share of slaves	−0.034 (0.051) {0.023}*	−0.073 (0.029)** {0.034}**	−0.060 (0.031)* {0.034}	−0.664 (0.249)*** {0.416}	−0.640 (0.228)*** {0.426}	−0.639 (0.230)*** {0.425}	−2.348 (0.919)** {1.456} [−0.608]
1872 share of free foreigners	0.874 (0.105)*** {0.061}***	0.742 (0.053)*** {0.053}***	0.786 (0.067)*** {0.063}***	-2.291 (0.873)** {0.905}**	-2.481 (0.972)** {0.771}***	-2.491 (0.923)*** {0.730}***	-11.063 (3.550)*** {5.231}** [−1.776]
Z_{1890}	-0.135 (0.027)*** {0.026}***		-0.042 (0.019)** {0.018}**				
Het. instr. (latitude)		-0.077 (0.011)*** {0.014}***	-0.074 (0.011)*** {0.015}***				
Het. instr. (dist. coast)		-0.430 (0.110)*** {0.114}***	-0.366 (0.115)*** {0.098}***				
Controls and FE	All	All	All	All	All	All	All
Observations	1,284	1,284	1,284	1,284	1,284	1,284	1,269
K-P F-stat	24.858	59.779	31.759				
B-P p-value		0.000	0.000				
P-H p-value					0.219	0.459	
Hansen J p-value					0.366	0.635	
Endog. test p-value				0.760	0.369	0.363	0.501

Note: *p<0.1; **p<0.05; ***p<0.01. District-vote two-way clustered standard errors in parentheses. Conley standard errors (with a 250km window) in curly brackets for 2SLS specifications, and one-way district-level clustered standard errors for the CML. Marginal effects in brackets. Columns 1-3 present alternative 2SLS first stages, columns 4-6 the corresponding second stages, and column 7 the CML (in a single step). Columns 1, 4, and 7 only use our standard instrument, columns 2 and 5 only use het.-based instruments, and columns 3 and 6 use both. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and votes FE. See Table 3 for details on the tests.

only use the two heteroskedasticity-based instruments, and when we use them as a supplement to the shift-share. As in the case of H1, the CML estimate is lower than the 2SLS estimate: it would suggest that a SD deviation in the share of foreigners in 1890 increases a representative's probability to vote in favor of emancipation by 11.3 percentage points, instead of around 15 percentage points for OLS and 2SLS estimates. We cannot reject the hypothesis that the share of foreigners in 1890 can be treated as exogenous, making the OLS specification (resp. GLM) consistent, and more efficient than the 2SLS (resp. CML).[19]

[19] In Table 13 in Appendix A.2, we report the 2SLS results taking net immigration as our main

6.3 Anti-abolition voting and the cost of coercion

In Table 6, we report the results for hypothesis 3. Columns 1 to 3 present the result of our baseline OLS specification with three different proxies of the distance to freedom: the number of quilombos in the district in column 1, the average area of quilombola land in the district's municipalities in column 2, and the (log of the) distance to the nearest quilombo in column 3. We replicate the same analysis limited to slaves employed in agriculture in columns 4 to 6, as hypothesis 3 and the outside option effect are more closely related to plantation slavery.

Table 6: H3 – Exit and voting decisions

	𝟙(Abolition vote)					
	All slaves			*Agric. slaves*		
	(1)	(2)	(3)	(4)	(5)	(6)
Share of slaves	-1.071 (0.326)*** {0.443}**	-1.097 (0.319)*** {0.431}**	-0.823 (0.279)*** {0.385}**	-1.415 (0.511)*** {0.687}**	-1.248 (0.482)*** {0.627}**	-0.767 (0.436)* {0.551}
Quilombos	-0.003 (0.003) {0.004}			-0.005 (0.004) {0.004}		
Sh. slaves × Quilombos	0.043 (0.028) {0.034}			0.116 (0.057)** {0.061}*		
Av area quilombola		-1.4e-3 (7.1e-4)** {7.5e-4}**			-0.001 (0.001)** {0.001}*	
Sh. slaves × Av. area quil.		0.011 (4.3e-3)** {4.2e-3}***			0.018 (0.007)*** {0.007}**	
Ln av. dist. 1 quil.			0.007 (0.008) {0.007}			0.002 (0.007) {0.006}
Sh. slaves × Ln av. dist. 1 quil.			-0.076 (0.042)* {0.043}*			-0.110 (0.090) {0.082}
Controls and FE	All	All	All	All	All	All
Observations	1,284	1,263†	1,284	1,284	1263†	1284
R^2	0.240	0.237	0.242	0.238	0.234	0.237

Note: *p<0.1; **p<0.05; ***p<0.01; † two outlier districts dropped, i.e. 26 obs. out of 1586, and 21 out of the 1284 for which a vote is registered. District-vote two-way clustered standard errors in parentheses. Conley standard errors (with a 250km window) in curly brackets. Columns 1-3 present OLS results using different measures of distance to freedom. Columns 4-6 reproduce the analysis focusing on slaves employed in agriculture. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and votes FE.

explanatory variable. In Table 14, we consider alternative definitions of our shift-share instrument. These alternative definitions are comfortably strong (except for one that relies on predicted district populations instead of the actual populations, which yields an F-stat noticeably under the usual threshold), and they all yield comparable estimates.

Distance to freedom is associated with a higher probability that a district's representative votes against abolitionist bills (in districts with a high prevalence of slavery). In all specifications, the relationship has the right sign, but the coefficient of interest is not always statistically significant (e.g. columns 1 and 6).[20] Our preferred measure is the area of quilombola land (columns 2 and 5). According to column 2, a 1 km^2 (247 acres) average increase in the area of quilombola land in a district's municipalities is associated with a $-.0014 + .0108 \times 50\% = .40$ percentage point increase in the probability to vote in favor of an emancipation bill in a district with a 50% prevalence of slavery. The average area of quilombos ranges from 0.004 km^2 (1 acre) to 7198 km^2 (1.8 million acres). As the distribution is skewed, instead of the SD, we consider the IQR (19.8 km^2) as the relevant unit of comparison. An increase in the area of quilombos by the IQR in a district with a 50% prevalence increases a representative's probability to vote in favor of abolition by 7.8 percentage points. Alternatively, according to column 3, decreasing the distance to the nearest quilombo by one SD (29 km) away from the mean is associated with a 2.6 percentage points increase in the probability to vote in favor of an emancipation bill in a district with a 50% prevalence of slavery.

Our results are also robust to instrumenting the location of quilombos by our standard instrument (terrain ruggedness interacted with remoteness) supplemented by heteroskedasticity-based instruments. Table 7 reports three 2SLS specifications: *i)* with the area of quilombola land (without the interaction term) instrumented by the standard instrument only (columns 1 and 6), *ii)* with both the area and its interaction with the share of slaves in 1872 instrumented by the standard instrument only (columns 2, 4, and 7), and *iii)* with both the area and its interaction with the share of slaves in 1872 instrumented by both standard and heteroskedasticity-based

[20] In Table 15 in Appendix A.2, we report the results of the specification that uses average area of quilombola land as the measure of the proximity to freedom with the outlier districts in logs and in levels. In Table 16, we report the results using the (log of the) distance to the nearest N quilombos, with N going from 1 to 9, as more measures of the distance to freedom. The magnitude and the significance of the estimates increases with N: a greater number of quilombos captures better prospects of freedom. Maintaining the distance constant, the promise of freedom is greater when local features of the land are more conducive to escaping durably, as captured by the number of quilombos. In Table 17, we report the results of the GLM specification for 3 measures of proximity (number of quilombos, (log of the) area of quilombola land, in a specification that includes the outliers, and area quilombola, without the outliers) and for 5 measures of distance (log of the distance to the 1, 3, 5, 7, and 9 nearest quilombos). Two-way clustering proves too heavy for these specifications, in which all coefficients have the right sign but lose statistical significance. This is most likely due to the limited number of clusters and the large number of parameters to be estimated, including both vote and province intercepts. Resorting to one-way district-level clustering solves the issue.

Table 7: H3 – Exit and voting decisions – Cont.

	First stage					Second stage		
	Av. area quilombola			*Av. area quil. × Sh. slaves*		𝟙(Abolition vote)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pred. av. area quilombola						-8.0e-4 (9.4e-4) {9.1e-4}	-3.4e-3 (1.2e-3)*** {1.6e-3}**	-2.5e-3 (9.4e-4)*** {1.1e-3}**
Pred. av. area quil. × Sh. slaves							2.3e-2 (6.7e-3)*** {7.3e-3}***	1.7e-2 (5.4e-3)*** {6.3e-3}***
1872 sh. slaves	98.711 (53.603)* {59.546}	226.42 (127.570)* {120.347}*	121.45 (71.454)* {64.870}*	-10.743 (23.750) {23.127}	2.052 (12.117) {11.387}	-0.911 (0.303)*** {0.458}***	-1.189 (0.318)*** {0.458}***	-1.143 (0.305)*** {0.457}**
Av. TRI × travel time	-1.526 (0.377)*** {0.318}***	-1.741 (0.656)*** {0.610}***	-0.996 (0.337)*** {0.273}***	-0.118 (0.090) {0.087}	-0.102 (0.065) {0.059}			
Av. TRI	49.057 (14.310)*** {13.651}***	80.215 (28.428)*** {29.645}**	36.677 (15.158)*** {13.704}***	5.958 (4.545) {4.054}	3.243 (2.643) {2.571}			
Travel time	1.330 (0.226)*** {0.178}***	1.485 (0.247)*** {0.303}***	0.635 (0.163)*** {0.172}***	0.032 (0.046) {0.047}	-0.008 (0.026) {0.024}			
Av. TRI × travel time × Sh. slaves		1.348 (3.308) {3.523}	1.823 (1.759) {1.626}	-0.620 (0.664) {0.660}	0.172 (0.343) {0.272}			
Av. TRI × Sh. slaves		-142.74 (117.070) {121.125}	-151.66 (72.674)** {65.620}**	4.766 (22.825) {23.821}	-22.998 (14.481) {12.3877}*			
Travel time × Sh. slaves		-1.456 (2.751) {2.865}	-0.096 (1.237) {0.078}	1.176 (0.589)** {0.582}**	0.642 (0.218)*** {0.253}***			
Het. instr. (latitude)			0.249 (0.029)*** {0.028}***		0.067 (0.008)*** {0.008}***			
Het. instr. (dist. coast)			0.628 (0.036)*** {0.028}***		0.049 (0.009)*** {0.008}***			
Controls and FE	All	All	All	All	All	All	All	All
Observations	1263†	1263†	1263†	1263†	1263†	1263†	1263†	1263†
K-P F-stat						12.02	4.337	25.37
B-P p-value			0.000		0.000			
P-H p-value								0.822
Hansen J p-value						0.878	0.940	0.629
Endog. test p-value						0.3427	0.177	0.241

Note: *p<0.1; **p<0.05; ***p<0.01; † two outlier districts dropped, i.e. 26 obs. out of 1586, and 21 out of the 1284 for which a vote is registered. District-vote two-way clustered standard errors in parentheses. Conley standard errors (with a 250km window) in curly brackets. Columns 1-5 present 2SLS first stages and columns 6-8 the corresponding second stages. Columns 1 and 6 use only our main instrument with only one endogenous variable: the average area of quilombola land. Columns 2,4 and 7 use only our main instrument with two endogenous variables: the area and its interaction with the share of slaves. Columns 3, 5 and 8 proceed similarly using both standard and het.-based instruments. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and votes FE. See Table 3 for details on the tests.

instruments (columns 3, 5, and 8).

Our main instrument is strong enough to instrument the area quilombola alone, but not enough to instrument both the area and its interaction with the share of slaves in 1872, while still including all controls and FE.[21] In a situation where the setup is too demanding for an instrument in which we believe nonetheless, summoning heteroskedasticity-based instruments is especially appealing. Again, we easily reject homoskedasticity with respect to exogenous regressors in the zero-stage, which confirms the relevance of generated instruments to supplement our main instrument. Again, we find reassuring that 2SLS estimates are comparable under two completely different sets of identification assumptions.[22] According to column 8, an increase in the area of quilombos by the IQR in a district with a 50% prevalence increases a representative's probability to vote in favor of abolition by 12.1 percentage points. The higher magnitude of the effect in the 2SLS estimation suggests that the OLS estimates may be downward-biased by a possible omitted variable: more abolitionist districts may have had, as a result of their culture or institutions, less need for quilombos (even controlling for the prevalence of slavery), or more abolitionist legislators may have self-selected into districts with larger/closer quilombos, which possibly had cultural norms more open to abolition. However, we cannot reject the hypothesis that the area of quilombola land and its interaction are jointly exogenous, making the OLS specfication consistent, and more efficient than the 2SLS.

7 Not a selection effect

In this section, we briefly address a possible concern that our results may be driven by a selection of legislators through their presence during the vote, or by a selection of emancipation bills. We may be worried that the decision to call the roll on an emancipation-related bill was in some way dependent on the anticipated distribution of the votes. Even if bills relating to emancipation were particularly sensitive, strongly publicized and surrounded by heated debate, so that the roll was called for every emancipation bill that passed through the Câmara dos Deputados during our sample's time frame, the bills discussed in parliament were sometimes divided in several sub-

[21] In Table 18 in Appendix A.2, we report the details of the first stage with non-linearities in the two uninteracted terms of the instrument for both the overall and the agricultural slave populations. In Table 19, we report the details of the second stage with the uninteracted terms of the instrument considered as controls instead of excluded instruments, as well as the Probit estimation. These yield similar results.

[22] In Table 21, we report the 2nd stage of the 2SLS specification with only heteroskedasticity-based instruments for various additional measures of the distance/proximity to freedom. They all paint the same story.

parts, which may individually be the object of nominative voting. To account for this possibility, in Table 8, we report our main specifications using only the three most important emancipation bills. In doing so, we reduce the size of our sample from 1,284 to 290 district × bill observations. This does not change substantially our results and suggests that our conclusions are not driven by a specific selection of bills.

Table 8: Core bills

	𝟙(Abolition vote)		
	(1)	(2)	(3)
1872 share of slaves	−1.091	−1.105	−1.261
	(0.478)***	(0.488)**	(0.517)**
	{0.545}**	{0.570}*	{0.585}**
1872 share of free foreigners		−1.687	
		(1.474)	
		{1.567}	
1890 share of foreigners		1.429	
		(0.377)***	
		{1.388}	
Av area quilombola			−0.001
			(0.001)
			(0.001)
Sh. slaves × Av. area quil.			0.014
			(0.006)**
			{0.007}**
Controls and FE	All	All	All
Observations	290	290	286†
R^2	0.386	0.388	0.392

Note: *p<0.1; **p<0.05; ***p<0.01. † two outlier districts dropped. District-vote two-way clustered standard errors in parentheses, except for column 3, which shows district-level clustered standard errors. Conley standard errors (with a 250km window) in curly brackets. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and votes FE.

We may also be worried that legislators from specific districts would display consistent patterns of absence and abstention on (scheduled) bills related to emancipation, thus biasing our results. For example, it is possible that some legislators supporting a given side were systematically absent when a bill was scheduled to be voted and they thought their side had no chance to win (or on the contrary, if they thought the matter was already won). At all times, a considerable number of legislators was absent from the Chamber (close to a fifth of the assembly). Absences were supposed to be justified, and they sometimes were. Most of the time, however, absent legislators did not provide justification. In practice, the only control that we find correlates with absenteeism is longitude (i.e. westernmost legislators within a province might be more inclined to absenteeism), but overall this does not suggest any obvious pattern in the selection of legislators.

8 Conclusions

In this paper, we investigated the determinants of the persistence and change of labor coercion institutions. We considered nineteenth-century Brazil, the last Western nation to abolish legal slavery, the largest importer of slaves through the Atlantic slave trade, and a country with (at the time) large areas of unexplored land. These three observations distinguish the Brazilian experience from any other; they also make it the ideal case study to unbundle the interests of slave-holding elites in the presence of an open agricultural frontier.

This paper also makes an original contribution to the historiography of nineteenth-century Brazil. From the archival records of the Imperial parliament, we built a district-vote-level data set documenting political decision-making on emancipation-related bills during the last decade of the Empire. We relied on census surveys, historical maps, and geo-referenced data sources to identify local variations in slave-holders' interests. We proposed a two-pronged instrumental variables strategy, leveraging historical and topographic determinants of the location of slaves and maroons across space as well as heteroskedasticity with respect to the regressors.

Our first result established the importance of slave-holders' interests in explaining voting patterns on emancipation bills. Representatives from districts with a high prevalence of slavery were less inclined to vote in favor of abolition. This accords with a large literature on the influence of elite interests and, more specifically, with the *labor demand effect*, and Engerman's (1973) argument for the persistence of the institutions of coercion. Our second result established that slave-holders' opposition to abolition was alleviated by the local availability of immigrant labor as a substitute to slave labor. This supports the importance of the land-to-labor ratio, as emphasized in the Brenner debate. Our third result established that slave-holders' interests differed between districts, depending on the local cost of enforcing the institutions of coercion. Representatives from high-prevalence districts where escaping slavery was easier were more likely to vote in favor of abolition. This nuances considerably the Nieboer-Domar hypothesis: even if the elite as a whole should favor coercion when land is abundant, the frontier planters, who apparently benefit the most from the abundance of land, favored switching to free labor. This third result is in line with the *outside option effect* (Acemoglu and Wolitzky, 2011), and in our opinion, is the main contribution of this paper. As an important corollary, this result emphasizes the role of coerced workers themselves in precipitating the collapse of the legal coercion system in Brazil. Insurrections and flights raised the costs of coercion to the planter class, which contributed to undermining the institution.

Together, these three hypotheses lay the foundations of a more general theory of institutional change in a democracy dominated by oligarchic interests. Following Stasavage (2014, p. 338), "a political regime results in the provision of property rights for a specific group, accompanied by significant barriers to entry." Individual elite members compare how they would profit under the two alternative institutional arrangements, that imply different patterns of ownership of productive assets – in this paper, a claim to owning the labor of enslaved workers. This paper revealed the importance of the *mobility* of labor (the ability for slaves to withdraw themselves from the institutional arrangement) and of the *rivalry* between slave and immigrant labor. Unbundling the interests of the elite looks like a promising way to expand this analysis to other aspects of institutional – and technological – changes.

References

Acemoglu, D., García-Jimeno, C., and Robinson, J. A. (2012). Finding Eldorado: Slavery and long-run development in Colombia. *Journal of Comparative Economics*, 40(4):534–564.

Acemoglu, D. and Wolitzky, A. (2011). The economics of labor coercion. *Econometrica*, 79(2):555–600.

Acharya, A., Blackwell, M., and Sen, M. (2016). The political legacy of American slavery. *Journal of Politics*, 78(3):621–641.

Alesina, A., Harnoss, J., and Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21(2):101–38.

Allen, T. (2015). The promise of freedom: Fertility decisions and the escape from slavery. *Review of Economics and Statistics*, 97(2):472–484.

Andersen, T. B. and Dalgaard, C.-J. (2011). Flows of people, flows of ideas, and the inequality of nations. *Journal of Economic Growth*, 16(1):1–32.

Anderson, R. N. (1996). The quilombo of Palmares: a new overview of a maroon state in seventeenth-century Brazil. *Journal of Latin American Studies*, 28(3):545–566.

Ashraf, Q. H., Cinnirella, F., Galor, O., Gershman, B., and Hornung, E. (2018). Capital-skill complementarity and the emergence of labor emancipation. Technical Report 6423, CESifo.

Aston, T. H. and Philpin, C., editors (1985). *The Brenner Debate: Agrarian Class Structure and Economic Development in Pre-Industrial Europe*. Cambridge University Press.

Bahar, D. and Rapoport, H. (2018). Migration, knowledge diffusion and the comparative advantage of nations. *Economic Journal*, 128:F273–F305.

Baum, C. F. and Lewbel, A. (2019). Advice on using heteroskedasticity-based identification. *The Stata Journal*, 19(4):757–67.

Beine, M. and Parsons, C. R. (2017). Climatic factors as determinants of international migration: Redux. *CESifo Economic Studies*, 63(4):386–402.

Bertocchi, G. and Dimico, A. (2014). Slavery, education, and inequality. *European Economic Review*, 70:197–209.

Bezemer, D., Bolt, J., and Lensink, R. (2014). Slavery, statehood, and economic development in Sub-Saharan Africa. *World Development*, 57:148–63.

Bobonis, G. J. and Morrow, P. M. (2014). Labor coercion and the accumulation of human capital. *Journal of Development Economics*, 108:32–53.

Brazil (1874). Recenseamento do Brasil em 1872. Typ. G. Leuzinger. Retrieved from `https://biblioteca.ibge.gov.br/biblioteca-catalogo?id=225477&view=detalhes`. Last accessed Dec 14, 2019.

Brazil (1895). Recenseamento geral da republica dos estados unidos do brazil em 31 de dezembro de 1890. Typ. G. Leuzinger. Retrieved from `https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=216866` and complemented with data from Leonardo Monasterio. Last accessed Dec 14, 2019.

Brenner, R. (1976). Agrarian class structure and economic development in pre-industrial Europe. *Past & Present*, 70(1):30–75.

Burchardi, K. B., Chaney, T., and Hassan, T. A. (2019). Migrants, ancestors, and foreign investments. *The Review of Economic Studies*, 86(4):1448–86.

Carvalho, J.-P. and Dippel, C. (2020). Elite identity and political accountability: A tale of ten islands. *The Economic Journal*. (in press).

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of econometrics*, 92(1):1–45.

Conley, T. G. (2010). Spatial econometrics. In Durlauf, S. N. and Blume, L. E., editors, *Microeconometrics*, pages 303–313. Springer.

Conrad, R. E. (1972). *The destruction of Brazilian slavery, 1850-1888*. University of California Press.

CPDOC (2016). Atlas histórico do Brasil. Centro de Pesquisa e Documentação de História Contemporânea do Brasil, Fundação Getúlio Vargas. Retrieved from `https://atlas.fgv.br/` on Dec. 14, 2019.

de Alencastro, L. F. (1979). La traite négrière et l'unité nationale brésilienne. *Outre-Mers. Revue d'Histoire*, 66(244):395–419.

de Carvalho Filho, I. and Monasterio, L. (2012). Immigration and the origins of regional inequality: Government-sponsored european migration to southern brazil before world war i. *Regional Science and Urban Economics*, 42(5):794–807.

Dean, W. (2012). *The Industrialization of São Paulo, 1800-1945*, volume 17. University of Texas Press.

Dell, M. (2010). The persistent effects of Peru's mining mita. *Econometrica*, 78(6):1863–1903.

Depetris-Chauvin, E. and Özak, O. (2020). The origins of the division of labor in pre-industrial times. *Journal of Economic Growth*, 25(3):297–340.

Dietrich, S. and Wright, J. (2015). Foreign aid allocation tactics and democratic change in Africa. *The Journal of Politics*, 77(1):216–34.

Dippel, C., Greif, A., and Trefler, D. (2020). Outside options, coercion, and wages: Removing the sugar coating. *Economic Journal*. (in press).

Docquier, F., Lodogiani, E., Rapoport, H., and Schiff, M. (2016). Emigration and democracy. *Journal of Development Economics*, 120:209–223.

Docquier, F., Turati, R., Valette, J., and Vasilakis, C. (2020). Birthplace diversity and economic growth: Evidence from the US states in the post-World War II period. *Journal of Economic Geography*, 20(2):321–54.

Domar, E. D. (1970). The causes of slavery or serfdom: A hypothesis. *The Journal of Economic History*, 30(1):18–32.

Eltis, D. (2007). 'A brief overview of the trans-Atlantic slave trade,' Slave voyages: The trans-Atlantic slave trade database. Retrieved from `http://www.slavevoyages.org/estimates/kCBqlBWr` and `http://www.slavevoyages.org/estimates/r3lNeZuD` on Feb. 2, 2020.

Engerman, S. L. (1973). Some considerations relating to property rights in man. *The Journal of Economic History*, 33(1):43–65.

Fails, M. D. (2019). Fuel subsidies limit democratization: Evidence from a global sample, 1990-2014. *International Studies Quarterly*, 63(2):354–63.

Felbermayr, G. J., Hiller, S., and Sala, D. (2010). Does immigration boost per capita income? *Economics Letters*, 107(2):177–9.

Fenske, J. (2012). Land abundance and economic institutions: Egba land and slavery, 1830-1914. *The Economic History Review*, 65(2):527–555.

Fogel, R. W. and Engerman, S. L. (1995 [1974]). *Time on the cross: The economics of American Negro slavery*, volume 1. WW Norton & Company. Reissue edition.

Fujiwara, T., Laudares, H., and Valencia Caicedo, F. (2019). Tordesillas, slavery and the origins of Brazilian inequality. unpublished.

Fundação Palmares (2020). Comunidades quilombolas certificadas. `http://www.palmares.gov.br/?page_id=37551`. Retrieved on June 21, 2020.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–624.

Graham, R. (1968). *Britain and the Onset of Modernization in Brazil 1850-1914*. Cambridge University Press.

Green, E. (2014). The economics of slavery in the eighteenth-century Cape colony: Revising the Nieboer-Domar hypothesis. *International Review of Social History*, 59(1):39–70.

IBGE (2010). Evolução da divisão territorial do Brasil - conteúdo digital. Retrieved from `https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estruturaterritorial/` on Dec. 14, 2019.

IIASA/FAO (2012). Global agro-ecological zones (gaez v3.0). IIASA, Laxenburg, Austria and FAO, Rome, Italy.

INCRA (2020). Acervo fundiário: áreas de quilombolas. `http://acervofundiario.incra.gov.br/acervo/acv.php`. Retrieved on June 24, 2020.

IPEA (2006). Ipeadata. *Instituto de Pesquisa Econômica Aplicada*, `http://www.ipeadata.gov.br/Default.aspx`. Retrieved on February 12, 2021.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018). Shift-share instruments and the impact of immigration. NBER Working Paper 24285, National Bureau of Economic Research.

Jobim, N. and Porto, W. C. (1996). Legislação eleitoral no Brasil : do século XVI a nossos dias. Brasília : Senado Federal, Subsecretaria de Biblioteca. Retrieved from `http://www2.senado.leg.br/bdsf/handle/id/224494` on Dec. 14, 2019.

Klein, A. and Ogilvie, S. (2017). Was Domar right? Serfdom and factor endowments in Bohemia. CAGE Online Working Paper Series 344, Competitive Advantage in the Global Economy (CAGE).

Klein, H. S. (2010). *The Atlantic slave trade*. Cambridge University Press.

Klein, H. S. and Luna, F. V. (2009). *Slavery in Brazil*. Cambridge University Press.

König, M. D., Rohner, D., Thoenig, M., and Zilibotti, F. (2017). Networks in conflict: Theory and evidence from the great war of africa. *Econometrica*, 85(4):1093–1132.

Lambais, G. (2020). Slave resistance, cultural transmission, and Brazil's long-run economic development. Retrieved from `https://gbrlambais.github.io/resistance_longrun.pdf` on Oct. 28th, 2020.

Le Roy Ladurie, E. (1969). *Les paysans de Languedoc*. Flammarion.

Leff, N. H. (1991). Economic development in Brazil, 1822-1913. Technical Report 92-02, Graduate School of Business, Columbia University.

Lewbel, A. (2012). Using heteroskedasticity to identity and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30(1):67–80.

Lewbel, A. (2019). The identification zoo: Meanings of identification in econometrics. *Journal of Economic Literature*, 57:835–903.

Lowrie, S. (1938). O elemento negro na população de São Paulo. *Revista do Arquivo Municipal*, 48:27–32.

Luna, L. (1976). *O negro na luta contra a escravidão*. Livraria Editora Cátedra.

Magnusson, L. M. and Mavroeidis, S. (2014). Identification using stability restrictions. *Econometrica*, 82(5):1799–851.

Markevich, A. and Zhuravskaya, E. (2018). The economic effects of the abolition of serfdom: Evidence from the Russian empire. *American Economic Review*, 108(4-5):1074–1117.

Martins, J. d. S. (1973). *A imigração e a crise do Brasil agrário*. Pioneira.

Mendonça, A. G. (2003). República e pluralidade religiosa no Brasil. *Revista USP*, 59:144–63.

Milliet, S. (1941). *Roteiro do café: e outros ensaios*, volume 25. Coleção do Departamento de Cultura, 3rd edition. Digitized on Oct. 2, 2007.

Morse, R. M. (1958). *From Community to Metropolis: A Biography of São Paulo, Brazil*. University of Florida Press.

Naidu, S. and Yuchtman, N. (2013). Coercive contract enforcement: law and the labor market in nineteenth century industrial Britain. *American Economic Review*, 103(1):107–144.

Naritomi, J., Soares, R. R., and Assunção, J. J. (2012). Institutional development and colonial heritage within Brazil. *The journal of economic history*, 72(2):393–422.

Nieboer, H. J. (2011 [1900]). *Slavery as an industrial system: Ethnological researches*. Cambridge University Press.

Nogueira, O. and Firmo, J. S. (1973). *Parlamentares do Império (obra comemorativa do sesquicentenário da instituição parlamentar no Brasil)*, volume 1. Centro Gráfico do Senado Federal.

North, D. C. and Thomas, R. P. (1973). *The rise of the western world: A new economic history*. Cambridge University Press.

Nunn, N. (2008). The long-term effects of Africa's slave trades. *The Quarterly Journal of Economics*, 123(1):139–176.

Nunn, N. and Puga, D. (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics*, 94(1):20–36.

Nunn, N. and Wantchekon, L. (2011). The slave trade and the origins of mistrust in Africa. *American Economic Review*, 101(7):3221–52.

Nützenadel, A. and Trentmann, F. (2008). *Food and globalization: consumption, markets and politics in the modern world*. Berg.

Ortega, F. and Peri, G. (2014). Openness and income: The roles of trade and migration. *Journal of international Economics*, 92(2):231–51.

Özak, Ö. (2010). The voyage of homo-economicus: Some economic measures of distance. *mimeo*.

Özak, Ö. (2018). Distance to the pre-industrial technological frontier and economic development. *Journal of Economic Growth*, 23(2):175–221.

Papadia, A. (2019). Slaves, migrants and development in Brazil, 1872-1923. Technical Report 2019/05, European University Institute.

Postan, M. (1937). The chronology of labour services. *Transactions of the Royal Historical Society*, 20:169–193.

Reis, E. J. and Reis, E. P. (1988). As elites agrárias e a abolição da escravidão no Brasil. *Dados, Revista de Ciéncias Sociais*, 31(3):309–341.

Repórter Brasil (2015). Dados sobre trabalho escravo no B rasil. `https://reporterbrasil.org.br/dados/trabalhoescravo`. Retrieved in Mar. 20, 2020.

Ridings, E. (1994). *Business interest groups in nineteenth-century Brazil*. Cambridge University Press.

Riley, S. J., DeGloria, S. D., and Elliot, R. (1999). A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of sciences*, 5(1-4):23–27.

Rocha, R., Ferraz, C., and Soares, R. R. (2017). Human capital persistence and development. *American Economic Journal: Applied Economics*, 9(4):105–36.

Scarato, L. C. (2009). Caminhos e descaminhos do ouro nas minas gerais: adminstração, territorialidade e cotidiano. Master's thesis, Universidade Estadual de Campinas.

Schwartz, S. B. (1978). Indian labor and new world plantations: European demands and indian responses in northeastern Brazil. *The American Historical Review*, 83(1):43–79.

Schwartz, S. B. (1992). Rethinking Palmares: Slave resistance in colonial Brazil. In Schwartz, S. B., editor, *Slaves, peasants and rebels*. University of Illinois Press.

Stasavage, D. (2014). Was Weber right? the role of urban autonomy in Europe's rise. *American Political Science Review*, 108(2):337–54.

Stein, S. J. (1986 [1957]). *Vassouras, a Brazilian coffee county, 1850-1900: the roles of planter and slave in a plantation society*. Princeton University Press.

Summerhill, W. (2010). Colonial institutions, slavery, inequality, and development: Evidence from São Paulo, Brazil. MPRA Paper 22162, University Library of Munich, Germany.

Sutch, R. C. (2018). The economics of African American slavery: The cliometrics debate. NBER Working Papers 25197, National Bureau of Economic Research, Inc.

Tabellini, M. (2020). Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. *The Review of Economic Studies*, 87(1):454–486.

The International Labour Organization and the Walk Free Foundation (2017). *Global Estimates of Modern Slavery: Forced Labour and Forced Marriage*. ILO. Retrieved form `https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/documents/publication/wcms_575479.pdf`. Last accessed Jul 1, 2020.

Toplin, R. B. (1969). Upheaval, violence, and the abolition of slavery in Brazil: the case of São Paulo. *The Hispanic American Historical Review*, 49(4):639–655.

Viotti da Costa, E. (1989). 1870-1889. In Bethell, L., editor, *Brazil: Empire and Republic, 1822-1930*, volume 3. Cambridge University Press.

Williams, E. (1994 [1944]). *Capitalism and slavery*. UNC Press Books. with a new introduction by Colin A. Palmer.

Witzel de Souza, B. G. (2019). The rationale of sharecropping: immigrant bonded laborers and the transition from slavery in Brazil (1830-1890). Technical Report 239, Ibero America Institute for Econ. Research (IAI).

Zemella, M. P. (1990 [1951]). *O abastecimento da capitania das Minas Gerais no século XVIII*, volume 19 of *Coleção Estudos Históricos*. Hucitec.

A Appendix

A.1 Robustness checks

This section provides a number of additional robustness checks. Subsection A.1.1 focuses on our first hypothesis and *i)* examines agricultural slavery, *ii)* tests alternative formulations of the instrument, and *iii)* explores the validity of the instrument. Subsection A.1.2 focuses on our second hypothesis and *i)* examines the stability of instrumenting immigration as a flow rather than a stock, *ii)* tests alternative ways to scale the main instrument and *iii)* explores the stability of the main instrument. Subsection A.1.3 focuses on our third hypothesis and *i)* shows that results are robust to alternative ways of dealing with outliers, *ii)* explores alternative measures of distance to freedom, *iii)* examines the stability of our main instrument, *iv)* proposes a placebo test of our main instrument, *v)* considers instrumenting distance to freedom with heteroskedasticity-based instruments only, and *vi)* examines the relationship between slavery and quilombos.

A.1.1 Anti-abolition voting and the prevalence of slavery

Table 9 tests an alternative version of our Hypothesis 1, focusing explicitly on plantation slavery, as measured by the share of slaves employed in agriculture in 1872. This shows that results remain qualitatively similar to Table 2.

Table 9: H1 – Prevalence of agricultural slavery and voting decisions

	𝟙(Abolition vote)						
	OLS						*GLM*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of agri. slaves	-1.536**	-1.662**	-0.759**	-0.450	-1.461**	-0.880*	-3.947**
	(0.689)	(0.752)	(0.356)	(0.739)	(0.699)	(0.492)	(1.976)
							[-0.988]
Controls	None	Geo.	Pol.	Dem.	Eco.	All	All
Province & vote FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,284	1,284	1,284	1,284	1,284	1,284	1269†
R^2	0.237	0.250	0.397	0.247	0.239	0.408	0.353‡

Note: *p<0.1; **p<0.05; ***p<0.01. District-vote two-way clustered standard errors in parentheses. Marginal effect in brackets. We consider alternative groups of controls in columns 1 to 5, and include them all in column 6. Column 7 is a probit specification with all controls. Geographical controls: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, and human mobility index. Political controls: party affiliation. Demographic controls: share of free colored and of literacy. Economic controls: gold mining and distance to closest diamond mine.

Table 10 tests the stability of our results to different formulations of the main slavery instrument. Column 1 corresponds to a baseline specification, using only our preferred interaction instrument in a just-identified model (with the instrument having no other interpretation than the distance to gold paths scaled by 16th cent. Indigenous repression). Column 2 includes uninteracted terms as instruments, and shows that the latter have no explanatory power, and that results remain stable despite the F-stat dropping below usual levels. Column 3 instead includes uninteracted terms as controls, which results in a larger impact of slavery, but is consistent with previous results in showing that OLS estimates likely underestimate the true influence of the prevalence of slavery. Columns 4-6 reproduce these steps using 16th to 18th century repression instead of 16th century only. This results in a slightly weaker instrument but similar results. Table 11 proceeds analogously with an IV probit estimated in a single step. Overall, results in Tables 10 and 11 show that our results remain qualitatively stable across formulations, and always consistent with estimates obtained from baseline specifications, increasing our confidence in the latter.

Table 10: H1 – Slavery and voting decisions – Alternative instr.

	Share of slaves					
Panel A: 1st stage	(1)	(2)	(3)	(4)	(5)	(6)
Ln 16th rep. area × Dist. Gold Paths	1.01e-5***	1.14e-5***	1.14e-5***	1.06e-5***	8.73e-06**	8.73e-6**
	(2.89e-6)	(4.02e-6)	(4.02e-6)	(2.86e-06)	(4.11e-06)	(4.11e-06)
Dist. Gold Paths		-2.6e-5	-2.6e-5		6.33e-06	6.33e-06
		(4.44e-5)	(4.44e-5)		(5.38e-5)	(5.38e-5)
Ln 16th rep. area		1.855e-4	1.855e-4		0.002	0.002
		(1.736e-3)	(1.736e-3)		(0.001)	(0.001)
Ln 17-18th rep. area × Dist. Gold Paths				-8.30e-06	-7.28e-06	-7.28e-06
				(6.53e-06)	(8.09e-06)	(8.09e-06)
Ln 17-18th rep. area					-0.001	-0.001
					(0.002)	(0.002)
Ln 16th rep. area × Ln 17-18th rep. area					-1.86e-4	-1.86e-4
					(2.48e-4)	(2.48e-4)
Panel A: 2nd stage			1(Abolition vote)			
Pred. share of slaves	-3.992**	-4.251**	-8.499**	-4.140***	-2.609**	-8.235***
	(1.689)	(1.791)	(3.313)	(1.522)	(1.319)	(3.075)
Dist. Gold Paths			2.79e-4			6.33e-06
			(3.69e-4)			(5.38e-5)
Ln 16th rep. area			0.028			0.040***
			(0.017)			(0.014)
Ln 17-18th rep. area						-0.014
						(0.017)
Ln 16th rep. area × Ln 17-18th rep. area						-0.003
						(0.003)
Controls & FE	All	All	All	All	All	All
Observations	1,284	1,284	1,284	1,284	1,284	1,284
K-P F-stat	12.405	4.705	8.063	7.577	3.150	3.920
Hansen J p-value	N/A	0.166	N/A	0.887	0.291	0.314

Note: *p<0.1; **p<0.05; ***p<0.01. District-vote two-way clustered standard errors in parentheses. Panel A presents alternative 1st stage specifications, and panel B the corresponding 2nd stages. Column 1 only uses our preferred standard instrument. Column 2 includes the non-interacted terms as instruments. Column 3 includes the non-interacted terms as flexible controls. Columns 4-6 reproduce the former using 16th to 18th century repression. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and vote FE.

Table 12 proposes a falsification test to assess the exclusion restriction underlying our preferred instrument. The validity of the latter crucially rests on the assumption that the distance to gold paths, scaled using 16th century Indigenous repression, does not affect emancipation voting other than through its influence on the geographical distribution of slaves across municipalities, within provinces. An important concern is that gold paths and the repression of Indigenous peoples may have influenced economic activity through channels distinct from slavery alone (e.g. urbanization). This economic activity might have in turn affected abolition voting at the end of the 19th century. This is unlikely to be the case, in particular because gold paths appear to have ceased to matter significantly passed the decline of the Gold Rush at the end of the 18th century. In addition, Indigenous peoples were driven out from settlement areas passed the 16th century. We only use Indigenous repression to predict more accurately areas within provinces that were more likely to receive African slaves after the 16th century, and our results are robust to using alternative scaling variables. Nevertheless, in order to formally test this assumption, we examine the influence of our instrument and its main components on economic outcomes in Table 12. Ideally, we would examine this relationship using 'pre-treatment' outcomes. There are however no

Table 11: H1 – Prevalence of slavery and voting decisions – Cond. ML

	𝟙(Abolition vote)					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of slaves	-15.530***	-18.885***	-23.721***	-15.528***	-13.947*	-25.700***
	(5.305)	(7.061)	(4.614)	(4.954)	(8.261)	(4.623)
	[-0.870]	[-0.816]	[-0.815]	[-0.831]	[-0.847]	[-0.873]
Dist. Gold Paths			7.568e-4			2.135e-4
			(0.001)			(0.001)
Ln 16th rep. area			0.068			.140***
			(0.054)			(0.045)
Ln 17th-18th rep. area						-.0432
						(0.054)
Ln 16th rep. area × Ln 17th-18th rep. area						-0.013
						(0.009)
Controls & FEs	All	All	All	All	All	All
Observations	1,269	1,269	1,269	1,269	1,269	1,269
Wald test p-value	0.0881	0.1178	0.0537	0.0634	0.3145	0.0127

Note: *p<0.1; **p<0.05; ***p<0.01. District-vote two-way clustered standard errors in parentheses. Marginal effect in brackets. Column 1 only uses our preferred standard instrument. Column 2 includes the non-interacted terms as instruments. Column 3 includes the non-interacted terms as flexible controls. Columns 4-6 reproduce the former using 16th to 18th century repression. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and vote FE.

available data that would allow us to do so, and we therefore resort to using economic outcomes closest to the 'treatment.' Specifically, we evaluate the association between our instrument and the (log) value of industrial production in 1907 (columns 1-4) and the (log) GDP per capita in 1920 (both from the IBGE and digitized by the IPEA), at the cross-section and controlling for the prevalence of slavery in addition to our usual vector of controls and province fixed effects. If the distance to gold paths and Indigenous repression did influence economic activity in their own right, we would expect to find placebo effects in Table 12. The coefficients associated with the instrument and its components are reassuringly never significant, lending further support to the validity of the instrument.

Table 12: H1 – Instrument and further economic outcomes

	Ln prod. value 1907				Ln per cap. GDP 1920			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dist. Gold Paths	0.001		-0.001		-1.56e-4		2.03e-4	
	(0.002)		(0.003)		(1.99e-4)		(3.45e-4)	
Ln 16th rep. area		-0.081	-0.084			0.012	0.014	
		(0.059)	(0.059)			(0.009)	(0.009)	
Ln 16th rep. area × Dist. Gold Paths			2.61e-4	1.84e-4			-4.54e-5	-2.6e-5
			(4.19e-4)	(2.88e-4)			(3.75e-5)	(2.56e-5)
Observations	334	237	237	237	1,277	956	956	956
R-squared	0.298	0.347	0.350	0.344	0.567	0.546	0.547	0.546
Controls and Prov. FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors clustered at the 1872 municipality level in parentheses. Columns 1-4 and 5-8 examine the influence of our instrument and its components on (log of the) value of industrial production in 1907 and the (log of the) GDP per capita in 1920. Controls and FE: 1872 share of slaves, population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, share of free colored and of literacy, gold mining and distance to closest diamond mine, and province FE.

A.1.2 Anti-abolition voting and immigration

Table 13 proposes an alternative formulation of our religion-based immigration instrument, whereby variables are expressed in differences rather than in levels. Results are reassuringly similar to instrumenting the stock of 1890 foreigners.

Table 13: H2 – Immig. and voting decisions – 2SLS

	First stage			*Second stage*		
		$\Delta_{1890\text{-}1872}$			$\mathbb{1}$(Abolition vote)	
	(1)	(2)	(3)	(4)	(5)	(6)
Pred. $\Delta_{1890\text{-}1872}$				2.398***	2.366***	2.374***
				(0.813)	(0.545)	(0.550)
$Z_{1872\text{-}1890}$	-0.519***		-0.130**			
	(0.129)		(0.053)			
Het. instr. (latitude)		-0.070***	-0.065***			
		(0.008)	(0.009)			
Het. instr. (dist. coast)		-0.349***	-0.319***			
		(0.080)	(0.075)			
Controls & FE	All	All	All	All	All	All
Observations	1,284	1,284	1,284	1,284	1,284	1,284
K-P F-stat	16.330	71.447	58.254			
B-P p-value		0.000	0.000			
P-H p-value					0.475	0.474
Hansen J					0.181	0.395

Note: *p<0.1; **p<0.05; ***p<0.01. District-vote two-way clustered standard errors in parentheses. This table reproduces the results of Table 5 using the inflow of immigrants instead of the stock. Accordingly, we adapt the instrument of section 5.2.2 as $Z_{i1872\text{-}1890} = \frac{1}{\widehat{P}_{i1890}} \sum_r \widehat{\alpha}_{ir1872} \widehat{\Delta}^{-i}_{r1890\text{-}1872}$. Columns 1-3 present alternative 2SLS first stages, and columns 4-6 the corresponding second stages. Columns 1 and 4 only use our standard instrument, columns 2 and 5 only use het.-based instruments, and columns 3 and 6 use both. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and votes FE. See Table 3 for details on the tests.

Table 14 assesses the stability of our instrument across various specifications. Column 1 corresponds to our baseline specification, where the instrument uses predicted Roman Catholic population and is scaled using actual district population. In columns 2 and 3, we address the possible concern that district population may itself be an outcome of immigration, and use instead total population and predicted district population respectively as scaling variables. Results remain remarkably stable, although the instrument becomes slightly weaker in the latter case. Columns 4 and 5 evaluate the robustness of the instrument to relaxing one of the main assumptions underlying its construction, namely that the number of foreign Roman Catholics can be approximated using overall Roman Catholics and native population. Coefficients remain again reassuringly stable when we use instead the actual Roman Catholic population, whether the instrument is scaled using district population (column 4) or total population (column 5).

The main identification assumption underlying the validity of our instrument is that municipality level conditions that may have influenced the inflow of immigrants of any given religion must not affect abolition patterns in the 1880s (within provinces). This leads us to systematically controlling for immigration in 1872, which (in addition to being mechanically correlated with immigration in 1890) possibly has a distinct effect on emancipation-related voting (so that the variation we use to identify the influence of 1890 immigrants is in the religious composition of municipalities' foreign populations, not in the actual size of the immigrant population). One might still be concerned that immigrants from specific religions or nationalities selected their destinations based on the possibility that these would be more inclined to vote for emancipation. We control for this possibility in column 6 of Table 14, and also add controls for colonial enclaves (núcleos coloniais) in column 7. Results remain qualitatively similar in both cases.

Table 14: H2 – Foreigners and voting decisions – Alternative instruments and instruments validity

	1890 share of foreigners						
Panel A: 1st stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Z_{1890}	-0.135***	36.087***	-0.075**	-0.004***	1.088***	-0.176***	-0.115***
	(0.027)	(0.027)	(0.030)	(0.001)	(0.142)	(0.029)	(0.028)
Share of slaves	-0.034	-0.060	-0.066	-0.043	-0.023	-0.041	-0.039
	(0.051)	(0.044)	(0.050)	(0.050)	(0.042)	(0.036)	(0.045)
1872 share of free foreigners	0.874***	0.054	0.848***	0.872***	-0.007	0.849***	0.850***
	(0.105)	(0.186)	(0.136)	(0.112)	(0.124)	(0.084)	(0.094)
Panel B: 2nd stage	𝟙(Abolition vote)						
Pred. 1890 share of foreigners	3.401***	3.641***	4.617*	3.224**	3.555***	2.676**	2.588**
	(1.193)	(1.242)	(2.676)	(1.271)	(0.979)	(1.412)	(1.271)
Share of slaves	-0.664***	-0.641***	-0.548	-0.681***	-0.649***	-1.092***	-0.660***
	(0.249)	(0.219)	(0.352)	(0.260)	(0.211)	(0.266)	(0.253)
1872 share of free foreigners	-2.291***	-2.471**	-3.207*	-2.158**	-2.407**	-1.885**	-1.650**
	(0.873)	(1.249)	(1.770)	(0.901)	(0.987)	(0.928)	(0.958)
Controls & FE	All	All	All	All	All	All	All
Observations	1,284	1,284	1,284	1,284	1,284	1,284	1,284
K-P F-stat	24.86	22.21	6.05	58.89	58.89	37.76	16.42

Note: *p<0.1; **p<0.05; ***p<0.01. District-vote two-way clustered standard errors in parentheses. In columns 1 to 3, the instrument uses the predicted foreign Roman Catholic population and is weighted using district, total and predicted district population respectively (predicted district population is given by the formula $\widehat{P}_{i1890} = P_{ij1872}(1 + (P^{-i}_{1890} - P^{-i}_{1872})/P^{-i}_{1872}))$. In columns 4 and 5, the instrument uses the actual Roman Catholic population and is weighted using district and total population respectively. Column 6 controls for individual country shares and Column 7 controls for núcleos coloniais.

A.1.3 Anti-abolition voting and the cost of coercion

In Table 15, we consider alternative ways to dealing with outliers, both for overall (columns 1-4) and agricultural slavery (columns 5-8). Columns 1 and 5 keep the explanatory variable as originally measured by the INCRA (2020), columns 2 and 6 use the average area of quilombola land in logs to penalize extreme values, columns 3 and 7 drop the two outlier districts, and columns 4 and 8 do both. Results remain stable across these alternatives, to the exception of column 5, where the interacted term loses significance.

Table 15: H3 – Exit and voting decisions – outliers

	𝟙(Abolition vote)							
	All slaves				*Agric. slaves*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1872 Share of slaves	-1.079***	-1.125***	-1.097***	-1.119***	-1.161**	-1.427***	-1.248***	-1.422***
	(0.319)	(0.310)	(0.319)	(0.313)	(0.497)	(0.542)	(0.482)	(0.547)
Av. quilombo area	-0.0001**	-0.045*	-0.001**	-0.046*	-0.0001***	-0.043**	-0.001**	-0.044**
	(0.0001)	(0.025)	(0.001)	(0.025)	(0.0000)	(0.018)	(0.001)	(0.018)
Sh. slaves × Av. quil. area	0.004*	0.146*	0.011**	0.143*	0.008	0.369**	0.018***	0.371**
	(0.003)	(0.078)	(0.004)	(0.078)	(0.005)	(0.156)	(0.007)	(0.158)
Outliers excluded	No	No	Yes	Yes	No	No	Yes	Yes
In logs	No	Yes	No	Yes	No	Yes	No	Yes
Controls and FE	All	All	All	All	All	All	All	All
Observations	1,284	1,284	1,263†	1,263†	1,284	1,284	1,263†	1,263†
R^2	0.241	0.242	0.237	0.237	0.237	0.240	0.233	0.235

Note: *p<0.1; **p<0.05; ***p<0.01; † two outlier districts dropped, i.e. 26 obs. out of 1586, and 21 out of the 1284 for which a vote is registered. District-vote two-way clustered standard errors in parentheses. This table compares different ways to deal with outlier districts for both overall (columns 1-4) and agricultural (5-8) slavery. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and vote FE.

In Table 16 we extend our measures of distance to freedom by considering the distance to the N closest quilombos, $N \in \{1; 9\}$. As expected, the magnitude and the significance of the estimates increases with N: a greater number of quilombos captures better prospects of freedom. Maintaining the distance constant, the promise of freedom is greater when local features of the land are more conducive to escaping durably, as captured by the number of quilombos. For distances relative to a high number of quilombos, the support for coercive institutions is entirely driven by municipalities for which distance to freedom is large, i.e. which faced lower coercion costs. Again, this suggests that heterogeneity with respect to the cost of enforcing coercion plays an important role in determining voting behavior.

Table 17 provides a summary of the results of hypothesis 3 when the framework is instead that of a probit model, using a number of distance to freedom measures (number of quilombos in column 1, average quilombo area with and without outliers in columns 2 and 3, and different distances to closest quilombos in columns 4-8). Overall, this shows that results are robust to using a generalized linear model instead of a linear probability model.

Table 16: H3 – Exit and voting decisions – more distances

	𝟙(Abolition vote)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: All slaves				
1872 Share of slaves	-0.823	-0.705	-0.731	-0.733	-0.686	-0.637	-0.532	-0.536	-0.465
	(0.279)***	(0.314)**	(0.315)**	(0.315)**	(0.334)**	(0.337)*	(0.337)	(0.339)	(0.382)
Ln av. dist. N quil.	0.007	0.003	0.000	0.004	0.006	0.018	0.024	0.024	0.029
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.010)*	(0.013)*	(0.013)*	(0.013)**
Sh. slaves × Ln av. dist. N quil.	-0.076	-0.057	-0.050	-0.056	-0.068	-0.093	-0.111	-0.110	-0.135
	(0.042)*	(0.047)	(0.047)	(0.045)	(0.053)	(0.060)	(0.063)*	(0.064)*	(0.068)**
R^2	0.242	0.241	0.242	0.240	0.240	0.240	0.240	0.240	0.240
					Panel B: Agricultural slaves				
Share of agri. slaves	-0.767	-0.643	-0.665	-0.711	-0.694	-0.330	-0.325	-0.330	0.017
	(0.436)*	(0.486)	(0.486)	(0.507)	(0.491)	(0.491)	(0.486)	(0.485)	(0.621)
Ln av. dist. N quil.	0.002	-0.001	-0.001	-0.000	0.001	0.020	0.022	0.022	0.033
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.010)**	(0.011)**	(0.011)**	(0.013)***
Sh. agri. slaves × Ln av. dist. N quil.	-0.110	-0.096	-0.101	-0.074	-0.087	-0.226	-0.226	-0.225	-0.278
	(0.090)	(0.098)	(0.099)	(0.087)	(0.093)	(0.084)***	(0.081)***	(0.081)***	(0.109)**
R^2	0.237	0.238	0.239	0.236	0.236	0.239	0.239	0.239	0.237
N	1	2	3	4	5	6	7	8	9
Controls and FE	All	All	All	All	All	All	All	All	All
Observations	1,284	1,284	1,284	1,284	1,284	1,284	1,284	1,284	1,284

Note: *p<0.1; **p<0.05; ***p<0.01. District-vote two-way clustered standard errors in parentheses. In columns 1-9, we measure distance to freedom by the distance to the closest N quilombos, with $N \in \{1; 9\}$. Panel A uses all slaves, whereas panel B focuses on slaves employed in agriculture. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and vote FE.

Tables 18 (first stages) and 19 (second stages and IV probit) examine the validity and stability of our preferred instrument across various specifications, for both overall and agricultural slavery. The column 1 of both tables provides the results of a specification with no interaction term, with average quilombola land the only endogenous variable instrumented by our standard instrument. This shows that quilombos have no influence in and of themselves, despite the instrument being strong enough following the usual standards. The columns 2 and 3 of Table 18 provide the first stages of a specification in which the average area of quilombola land and its interaction with slavery are instrumented by our standard instrument and its interaction with slavery. The column 2 of Table 19 provides the second stage of this specification. Results are strongly consistent with our hypothesis three, despite the instruments becoming weaker. In columns 4 and 5 of Table 18 (column 3 of Table 19 for the second stage), we proceed similarly but also flexibly control for nonlinearities in the uninteracted terms. Results remain remarkably stable, which answers the possible concern

Table 17: H3 – Exit and voting decisions – GLM

	𝟙(Abolition vote)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: All slaves							
1872 Share of slaves	-4.274	-4.439	-4.298	-3.320	-3.045	-2.676	-2.109	-0.469
	(1.418)***	(1.398)***	(1.419)***	(1.078)***	(1.103)***	(1.163)**	(1.185)*	(2.108)
	{1.578}***	{1.565}***	{1.561}***	{1.554}**	{1.673}*	{1.661}	{1.843}	{2.739}
Dist. to freedom	-0.011	-0.173	-0.004	0.018	-0.003	0.024	0.095	0.221
	(0.017)	(0.113)	(0.113)	(0.037)	(0.036)	(0.042)	(0.064)	(0.135)
	{0.019}	{0.095}*	{0.002}**	{0.029}	{0.033}	{0.036}	{0.061}	{0.169}
Sh. slaves × Dist. to freedom	0.179	0.561	0.043	-0.230	-0.165	-0.282	-0.438	-0.859
	(0.110)	(0.375)	(0.029)	(0.206)	(0.207)	(0.219)	(0.272)	(0.504)*
	{0.144}	{0.297}*	{0.016}***	{0.163}	{0.180}	{0.203}	{0.287}	{0.583}
Pseudo R^2	0.3576	0.3584	0.3543	0.3576	0.3576	0.3570	0.3571	0.3573
	Panel B: Agricultural slaves							
1872 Share of slaves	-6.379	-5.959	-5.294	-3.516	-3.112	-3.040	-1.560	1.948
	(2.379)***	(2.422)**	(2.108)**	(1.652)**	(1.819)*	(1.778)*	(1.610)	(3.279)
	{2.622}**	{2.655}**	{2.483}**	{2.386}	{2.519}	{2.457}	{2.462}	{4.154}
Dist. to freedom	-0.018	-0.163	-0.005	0.006	-0.008	0.004	0.082	0.222
	(0.017)	(0.081)**	(0.004)	(0.032)	(0.034)	(0.037)	(0.047)*	(0.123)*
	{0.019}	{0.075}**	{0.002}**	{0.022}	{2.519}	{0.031}	{0.039}**	{0.153}
Sh. slaves × Dist. to freedom	0.497	1.424	0.073	-0.405	-0.399	-0.415	-0.916	-1.597
	(0.219)**	(0.680)**	(0.045)	(0.429)	(0.475)	(0.433)	(0.367)**	(0.777)**
	{0.268}*	{0.638}**	{0.025}***	{0.294}	{0.343}	{0.361}	{0.323}***	{0.883}*
Pseudo R^2	0.3564	0.3570	0.3570	0.3547	0.3561	0.3540	0.3564	0.3553
Controls and FE	All	All	All	All	All	All	All	All
Observations	1,269	1,269	1,248‡	1,269	1,269	1,269	1,269	1,269

Note: *p<0.1; **p<0.05; ***p<0.01. ‡ two outlier districts dropped, i.e. 26 obs. out of 1586, and 21 out of the 1284 for which a vote is registered. District-vote two-way clustered standard errors in parentheses. District-level one-way clustered standard errors in curly brackets. This table presents the results of a probit model using (in the corresponding columns) the following measures of distance to freedom: (1) number of quilombos in the district, (2) Ln av. quilombo area, including outliers, (3) Av. quilombo area, excluding outliers; (4) Ln av. distance to the nearest quilombo, (5) to the 3 nearest quilombos, (6) to the 5 nearest quilombos, (7) to the 7 nearest quilombos, and (8) to the 9 nearest quilombos. H3 corresponds to a positive coefficient for the interacted variable in columns (1) to (3), and negative in columns (4) to (8). Panel A uses all slaves, whereas panel B focuses on slaves employed in agriculture. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and vote FE.

that the instrument may be accidentally picking up nonlinearities. Column 4 of Table 19 provides the results of a specification in which the uninteracted terms are used as controls rather than included instruments, i.e. the influence of distance to freedom is identified using only the differential effect of ruggedness by the degree of remoteness, not ruggedness or remoteness themselves (and the first stages are equivalent to columns 2 and 3 of Table 18). Again, results remain reassuringly very stable. Finally, the column 5 of Table 19 is analogous to column 2, but we estimate an IV probit by conditional maximum likelihood rather than proceeding with 2SLS.

In Table 20, we run a placebo test of our main instrument's validity. A key feature of our empirical design for Hypothesis 3 is that by themselves, quilombos have no influence on legislator's voting decisions. If our instrument is valid and indeed only impacts voting decisions through its influence on the location and size of quilombos, we should therefore observe no effect of the instrument in the reduced form regression with no interaction term. If, on the contrary, ruggedness, remoteness, and their interaction influenced voting decisions on abolition-related roll calls through different channels than that of quilombos (e.g. trade), than we would expect a statistically significant effect of the instrument in the reduced form. Columns 1 and 2 of Table 20 respectively provide results for the baseline OLS and 2SLS specifications (for the sake of comparison), whereas column 3 presents results of the reduced form. Reassuringly, none of the coefficient associated with the instrument's component are statistically significant, which strongly supports the validity of the exclusion restriction.

Table 18: H3 – Exit and voting decisions – 2SLS 1st stage

	Av. quilombo area				
	Uninteracted	*Interacted*		*Interacted*	
Panel A: Overall slaves	(1)	(2)	(3)	(4)	(5)
1872 Share of slaves	98.711*	226.416*	-10.743	131.400	-26.735
	(53.603)	(127.570)	(23.750)	(130.802)	(26.173)
Av. terrain ruggedness	49.057***	80.215***	5.958	67.032**	5.013
	(14.310)	(28.428)	(4.545)	(29.039)	(4.266)
Travel time to nearest cap.	1.330***	1.485***	0.032	0.471	-0.110
	(0.226)	(0.297)	(0.046)	(0.548)	(0.078)
Ruggedness × Travel time	-1.526***	-1.741***	-0.118	-1.559**	-0.097
	(0.377)	(0.656)	(0.090)	(0.671)	(0.092)
Ruggedness × Share of slaves		-142.741	4.766	-99.716	13.625
		(117.072)	(22.825)	(121.175)	(25.083)
Travel time × Share of slaves		-1.456	1.176**	0.345	1.440**
		(2.751)	(0.589)	(2.745)	(0.604)
Ruggedness × Travel time × Share of slaves		1.348	-0.620	1.209	-0.620
		(3.308)	(0.664)	(3.144)	(0.642)
Ruggedness × Ruggedness				0.465	-0.696
				(7.862)	(1.451)
Travel time × Travel time				0.004**	0.001***
				(0.002)	(0.000)
Panel B: Agricultural slaves					
1872 Sh. of agri. slaves	173.813*	36.840	-37.099	25.951	-37.882
	(98.096)	(239.813)	(26.030)	(209.206)	(25.527)
Av. terrain ruggedness	42.688***	40.233*	-0.338	33.652	-0.455
	(13.489)	(22.503)	(1.685)	(25.244)	(1.790)
Travel time to nearest cap.	1.286***	1.064***	-0.015	0.176	-0.059**
	(0.220)	(0.302)	(0.022)	(0.444)	(0.026)
Ruggedness × Travel time	-1.447***	-1.146*	0.008	-0.918	0.020
	(0.363)	(0.618)	(0.036)	(0.600)	(0.033)
Ruggedness × Share of agri. slaves		69.372	40.286	64.264	40.341
		(237.496)	(25.929)	(220.177)	(26.059)
Travel time × Share of agri. slaves		4.144	1.828***	4.791	1.861***
		(5.055)	(0.558)	(4.512)	(0.541)
Ruggedness × Travel time × Share of slaves		-3.748	-1.428**	-3.580	-1.423**
		(6.228)	(0.619)	(5.856)	(0.610)
Ruggedness × Ruggedness				0.737	-0.075
				(6.522)	(0.444)
Travel time × Travel time				0.004***	0.000**
				(0.002)	(0.000)
Controls and FE	All	All	All	All	All
Observations	1,263†	1,263†	1,263†	1,263†	1,263†

Note: *p<0.1; **p<0.05; ***p<0.01; † two outlier districts dropped, i.e. 26 obs. out of 1586, and 21 out of the 1284 for which a vote is registered. District-vote two-way clustered standard errors in parentheses. This table presents alternative 2SLS 1st stage specifications for H3, using only standard instruments. Column 1 uses only our main instrument with only one endogenous variable: the average area of quilombola land. Columns 2-3 use only our main instrument with two endogenous variables: the average area of quilombola land (col 2) and its interaction with the share of slaves (col 3). Columns 4-5 proceed similarly but control for nonlinearities in uninteracted terms. Panel A uses all slaves, whereas panel B focuses on slaves employed in agriculture. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and vote FE.

(And they are also non-significant if added individually.)

In Table 21, we instrument our various measures of distance to freedom (number of quilombos in column 1, average quilombo area with and without outliers in columns 2 and 3, and different distances to closest quilombos in columns 5-8) using heteroskedasticity-based identification only. Again, this provides and informal test of our instrument's validity: the fact that we obtain similar results under completely different identification assumptions is strongly reassuring.

Finally, a remaining possible concern is that quilombos and slavery might be colliders, as the number of quilombos and their size can conceivably constitute an outcome of the prevalence of slavery. For this to be an issue, there would however need to be a reverse relationship between abolition voting and quilombos, which seems improbable in this setting. Moreover, quilombos are widespread across districts irrespective of the prevalence of slavery

Table 19: H3 – Exit and voting decisions – 2SLS 2nd stage & CMLE

	2SLS				CMLE
	$\mathbb{1}$(Abolition vote)				
Panel A: Overall slaves	(1)	(2)	(3)	(4)	(5)
1872 Share of slaves	-0.911***	-1.189***	-1.214***	-1.215***	-4.682***
	(0.303)	(0.318)	(0.312)	(0.312)	(1.499)
Av. quilombo area	-0.0008	-0.003***	-0.003	-0.003**	-0.013**
	(0.001)	(0.001)	(0.002)	(0.001)	(0.006)
Sh. slaves × Av. quil. area		0.023***	0.023***	0.024***	0.093***
		(0.007)	(0.008)	(0.007)	(0.035)
Ruggedness × Ruggedness			-0.003		
			(0.022)		
Travel time × Travel time			-0.000		
			(0.000)		
Av. terrain ruggedness				0.028	
				(0.069)	
Travel time to nearest cap.				-0.001	
				(0.001)	
K-P F-stat	12.022	6.648	2.483	4.018	
Hansen J	0.878	0.940	0.806	0.736	
Endog. test	0.343	0.177	0.253	0.162	0.013
Panel B: Agricultural slaves					
1872 Share of agri. slaves	-0.714	-1.481***	-1.503***	-1.600***	-6.429***
	(0.555)	(0.486)	(0.431)	(0.458)	(2.209)
Av. quilombo area	-0.001	-0.003**	-0.002	-0.002	-0.011**
	(0.001)	(0.001)	(0.002)	(0.002)	(0.006)
Sh. agri. slaves × Av. quil. area		0.035***	0.032***	0.034***	0.140***
		(0.010)	(0.011)	(0.011)	(0.048)
Ruggedness × Ruggedness			-0.002		
			(0.021)		
Travel time × Travel time			-0.000		
			(0.000)		
Av. terrain ruggedness				0.045	
				(0.070)	
Travel time to nearest cap.				-0.000	
				(0.001)	
K-P F-stat	11.939	6.812	4.119	7.271	
Hansen J	0.792	0.594	0.493	0.275	
Endog. test	0.492	0.763	0.842	0.854	0.024
Controls and FE	All	All	All	All	All
Observations	1,263†	1,263†	1,263†	1,263†	1,248†

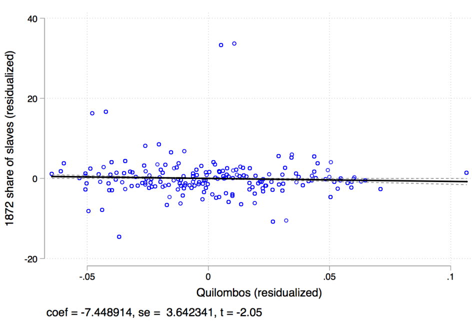
Note: *p<0.1; **p<0.05; ***p<0.01; † two outlier districts dropped, i.e. 26 obs. out of 1586, and 21 out of the 1284 for which a vote is registered. District-vote two-way clustered standard errors in parentheses. This table presents alternative 2SLS 2nd stage specifications for H3, using only standard instruments. Column 1 uses only our main instrument with only one endogenous variable: the average area of quilombola land (1st stage: T18 col 1). Column 2 uses only our main instrument with two endogenous variables: the average area of quilombola land and its interaction with the share of slaves (1st stages: T18 cols 2-3). Column 3 proceeds similarly but controls for nonlinearities in uninteracted terms (1st stages: T18 cols 4-5). Column 4 uses uninteracted terms as controls rather than instruments (1st stages are the same as column 2 T18 cols 2-3). Column 4 is analogous to column 2 but proceeding by CMLE in one step. Panel A uses all slaves, whereas panel B focuses on slaves employed in agriculture. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and vote FE. See Table 3 for details about the statistics.

in 1872, and once we control for our usual vector of covariates and fixed effects, quilombos and slavery are at most marginally correlated, as illustrated in Figure A.4.

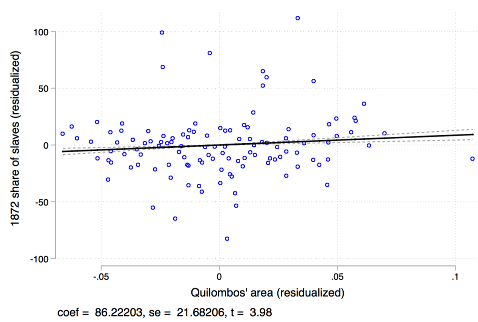
Table 20: H3 – Exit and voting decisions – Placebo test

	𝟙(Abolition vote)		
	(1)	(2)	(3)
1872 Share of slaves	-0.972***	-0.911***	-0.967***
	(0.307)	(0.303)	(0.309)
Av. quilombo area	-0.000	-0.001	
	(0.000)	(0.001)	
Av. terrain ruggedness			-0.051
			(0.102)
Travel time to nearest cap.			-0.001
			(0.001)
Ruggedness × Travel time			0.001
			(0.002)
Observations	1,263†	1,263†	1,284
R-squared	0.007	0.005	0.008
Controls and FE	All	All	All

Note: *p<0.1; **p<0.05; ***p<0.01; † two outlier districts dropped, i.e. 26 obs. out of 1586, and 21 out of the 1284 for which a vote is registered. District-vote two-way clustered standard errors in parentheses. This table presents a placebo test of Hypothesis 3. Column 1 is the OLS, Column 2 is the second stage of the 2SLS (first stage is the same as in Table 18), and column 3 is the reduced form. For the instrument to be valid, we should not observe any placebo effect of the instrument's components in column 3. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and vote FE.



(a) Quilombos



(b) Quilombos' area

Figure A.4: Slavery and quilombos

Table 21: H3 – Exit and voting decisions – 2SLS 2nd stage with het.-based instruments and various measures of distance to freedom

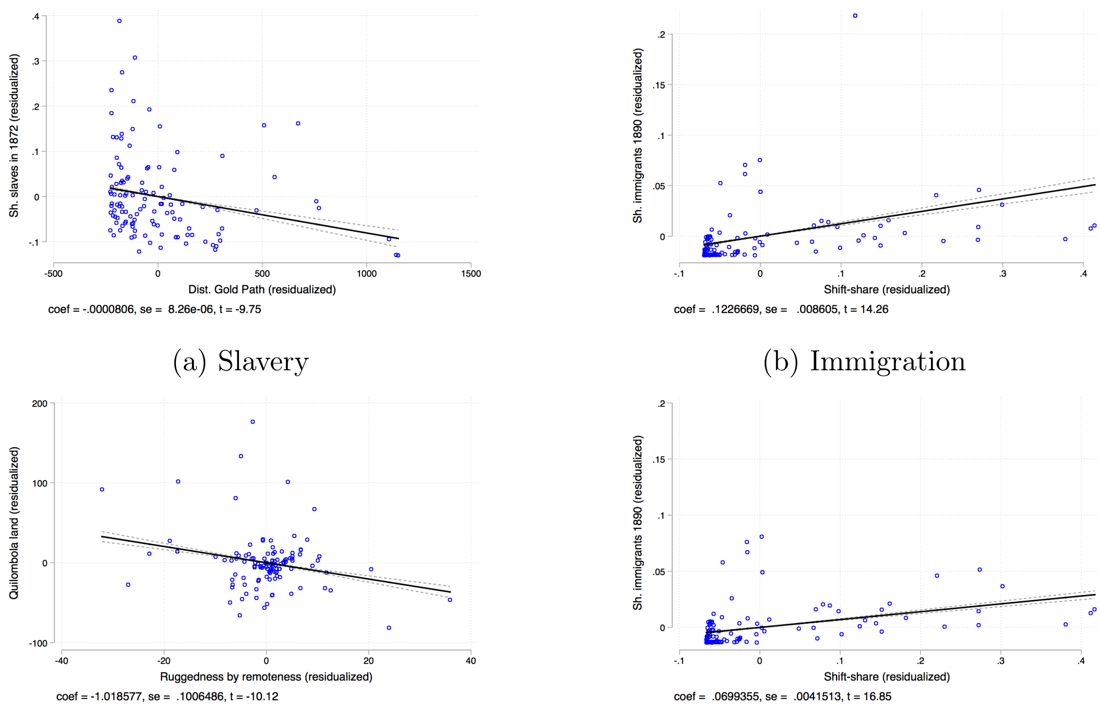
	𝟙(Abolition vote)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: All slaves				
1872 Share of slaves	-1.192***	-1.303***	-1.085***	-0.645**	-0.232	-0.482	0.333	0.169
	(0.316)	(0.323)	(0.303)	(0.316)	(0.542)	(0.398)	(0.696)	(0.682)
Distance	-0.011**	-0.171**	-0.002**	0.019**	0.015	0.009	0.059***	0.051**
	(0.005)	(0.086)	(0.001)	(0.008)	(0.017)	(0.012)	(0.022)	(0.022)
Sh. slaves × Distance	0.088**	0.461**	0.013**	-0.168***	-0.167*	-0.112*	-0.335**	-0.301**
	(0.041)	(0.231)	(0.006)	(0.046)	(0.099)	(0.061)	(0.134)	(0.126)
R^2	0.239	-0.0241	0.237	0.238	0.237	0.239	0.234	0.235
K-P F-stat	38.990	4.402	18.648	22.612	16.172	24.702	28.951	116.555
Hansen J	0.231	0.532	N/A	0.304	0.359	0.530	0.765	0.780
B-P p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
H-P p-value	0.495	0.770	0.816	0.511	0.430	0.199	0.227	0.194
				Panel B: Agricultural slaves				
Share of agri. slaves	-1.518***	-2.258***	-1.220***	-0.666	-0.289	-0.291	0.337	0.936
	(0.474)	(0.567)	(0.458)	(0.448)	(0.585)	(0.644)	(0.804)	(1.092)
Distance	-0.010**	-0.163***	-0.002*	0.007	0.018	0.029**	0.037***	0.056**
	(0.005)	(0.056)	(0.001)	(0.010)	(0.013)	(0.013)	(0.013)	(0.023)
Sh. agri. slaves × Distance	0.138**	1.112***	0.019**	-0.203*	-0.282**	-0.339*	-0.506***	-0.599***
	(0.067)	(0.316)	(0.009)	(0.110)	(0.120)	(0.176)	(0.183)	(0.230)
R^2	0.236	-0.040	0.233	0.236	0.232	0.224	0.232	0.231
K-P F-stat	78.100	9.520	30.115	10.315	11.419	37.300	39.495	28.139
Hansen J	0.5879	0.8384	N/A	0.1605	0.3727	0.3664	0.6451	0.6953
B-P p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
H-P p-value	0.1543	0.808	0.719	0.1331	0.1056	0.1699	0.2264	0.2198
Controls & FE	All	All	All	All	All	All	All	All
Observations	1,284	1,284	1,263†	1,284	1,284	1,284	1,284	1,284

Note: *p<0.1; **p<0.05; ***p<0.01; † two outlier districts dropped, i.e. 26 obs. out of 1586, and 21 out of the 1284 for which a vote is registered. District-vote two-way clustered standard errors in parentheses. This table presents 2SLS results using only heteroskedasticity-based instruments with (in the corresponding columns) the following measures of distance to freedom: (1) number of quilombos in the district, (2) Ln av. quilombo area, including outliers, (3) Av. quilombo area, excluding outliers; (4) Ln av. distance to the nearest quilombo, (5) to the 3 nearest quilombos, (6) to the 5 nearest quilombos, (7) to the 7 nearest quilombos, and (8) to the 9 nearest quilombos. H3 corresponds to a positive coefficient for the interacted variable in columns (1) to (3), and negative in columns (4) to (8). Het-based instruments are for each measure based on a subset of: share of free colored, average rainfall, distance to the coast, latitude, longitude, and sugarcane suitability. Controls and FE: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to closest river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to closest diamond mine, province and vote FE. See Table 3 for details about the statistics.

A.2 Additional graphs and tables

Table 22: 1872 nationality-religion matching

Nationality	Cath. VS N.-Cath.	Predicted religion	Nationality	Cath. VS N.-Cath.	Predicted religion
German	Cath.	CR	Italian	Cath.	CR
	N.-Cath.	Prot.		N.-Cath.	Prot./NA
Austrian	Cath.	CR	Japanese	Cath.	CR
	N.-Cath.	Prot.		N.-Cath.	Ath./NA
Argentinian	Cath.	CR	Mexican	Cath.	CR
	N.-Cath.	Prot./NA		N.-Cath.	Prot./NA
Belgian	Cath.	CR	North-American	Cath.	CR
	N.-Cath.	Prot./NA		N.-Cath.	Prot.
Bolivian	Cath.	CR	Oriental	Cath.	CR
	N.-Cath.	Prot./NA		N.-Cath.	M
Chinese	Cath.	CR	Paraguayan	Cath.	CR
	N.-Cath.	Ath./NA		N.-Cath.	Prot./NA
Danish	Cath.	CR	Persian	Cath.	CO
	N.-Cath.	Prot.		N.-Cath.	M
French	Cath.	CR	Peruvian	Cath.	CR
	N.-Cath.	Prot.		N.-Cath.	Prot./NA
Greek	Cath.	CO	Portuguese	Cath.	CR
	N.-Cath.	Prot./NA		N.-Cath.	Prot./NA
Spanish	Cath.	CR	Russian	Cath.	CO
	N.-Cath.	Prot./NA		N.-Cath.	M./NA
Dutch	Cath.	CR	Swiss	Cath.	CR
	N.-Cath.	Prot.		N.-Cath.	Prot.
Hungarian	Cath.	CR	Swedish	Cath.	CR
	N.-Cath.	Prot.		N.-Cath.	Prot.
English	Cath.	CR	Turkish	Cath.	CO/NA
	N.-Cath.	Prot.		N.-Cath.	M.



(a) Slavery

(b) Immigration

(c) Quilombos

(d) Immigration, w/o Rio de Janeiro

Figure A.5: First stage plots

A.3 Additional maps

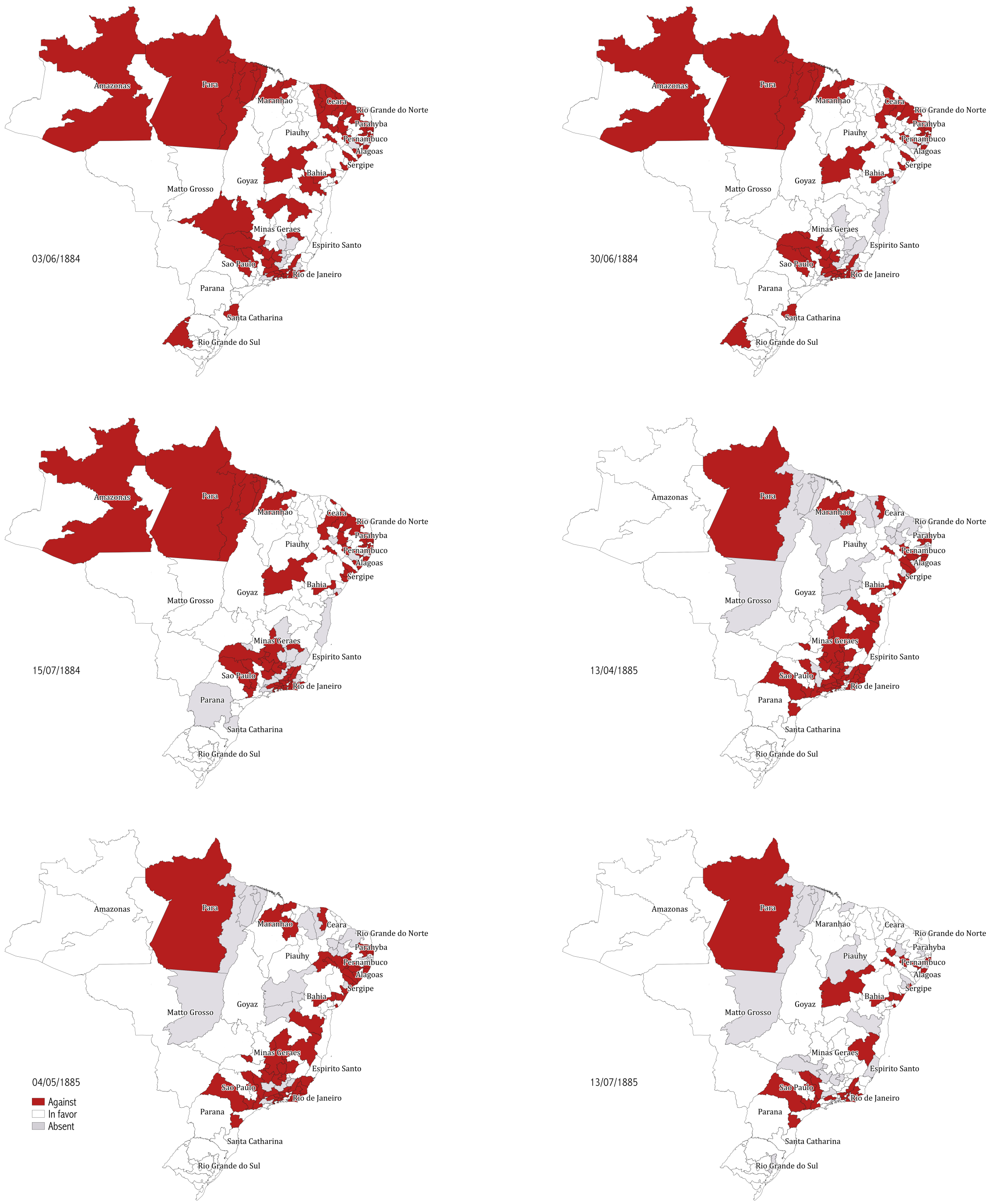


Figure A.6: Abolition maps – other bills (1)

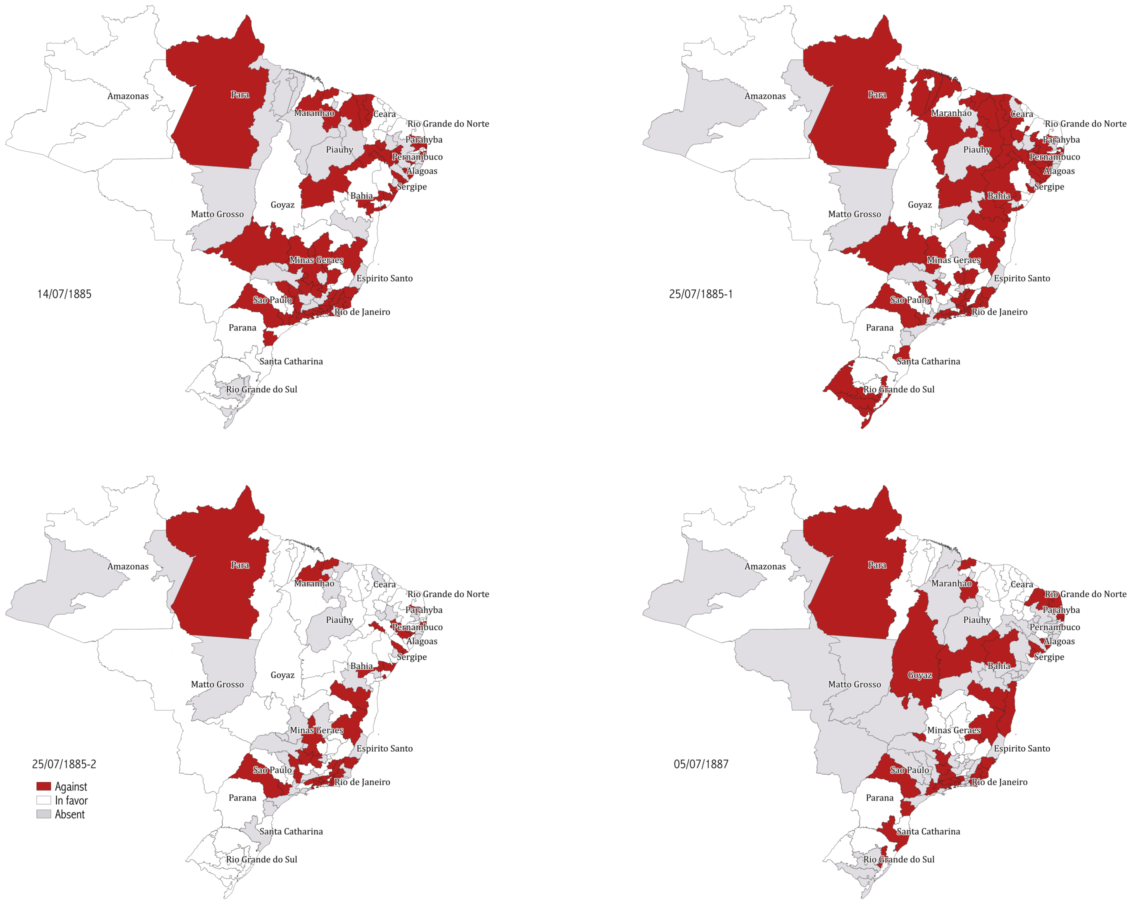


Figure A.7: Abolition maps – other bills (2)

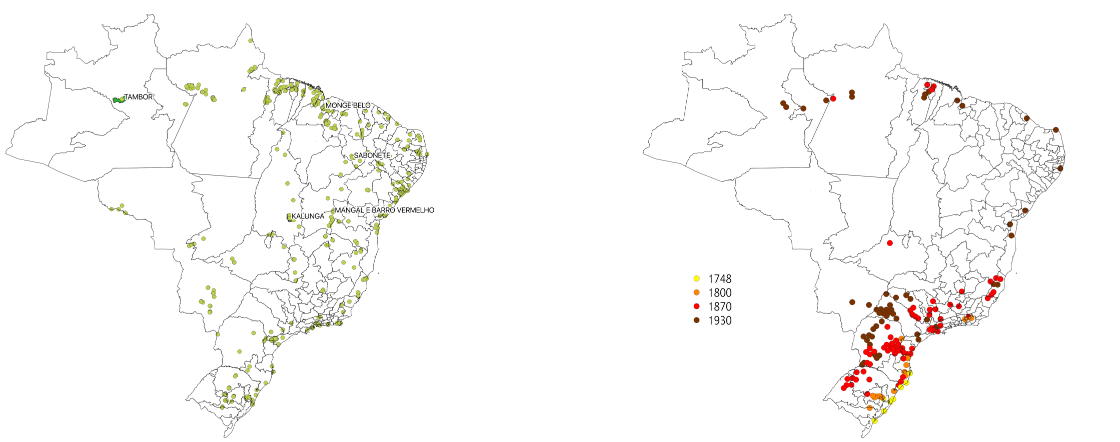


Figure A.8: Quilombos and núcleos coloniais with district boundaries

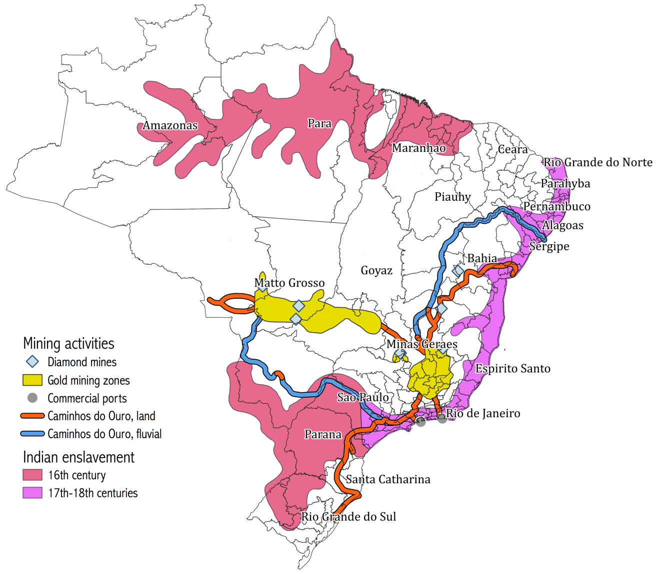


Figure A.9: Eighteenth-century mining activities and Indigenous enslavement (CPDOC, 2016).

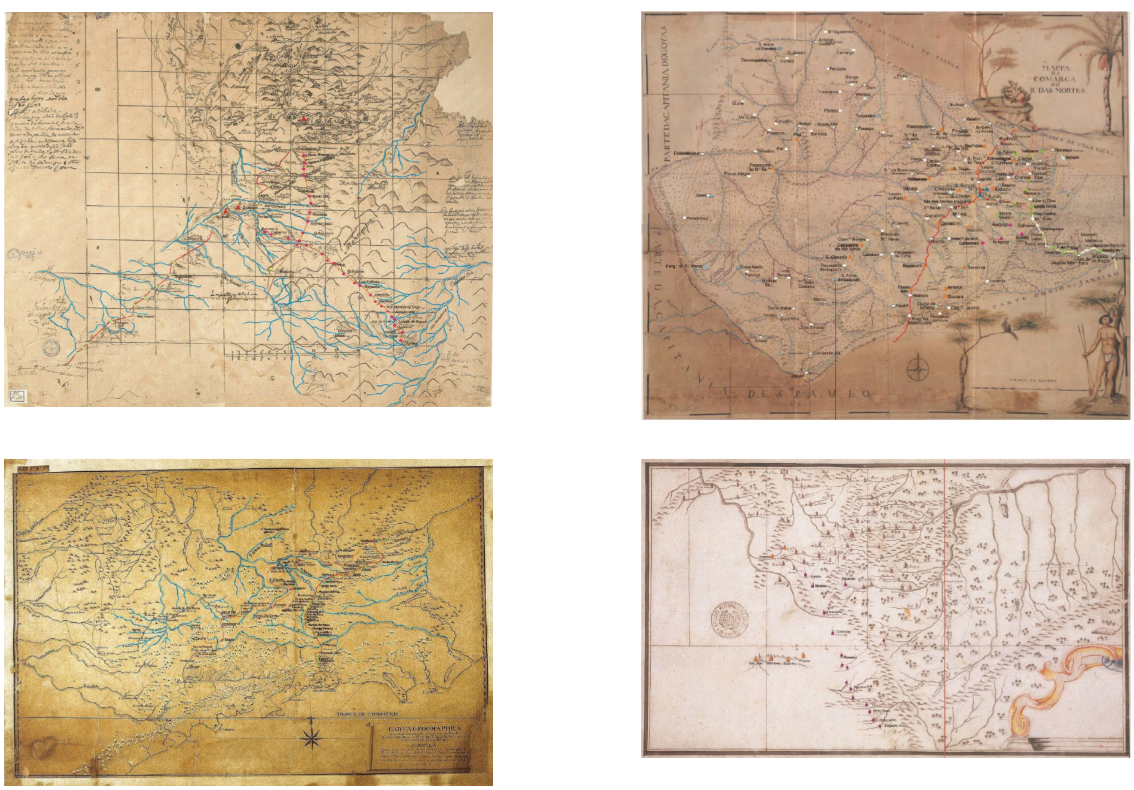


Figure A.10: Some historical *Caminhos do Ouro* maps from Scarato (2009)

A.4 Núcleos coloniais

The *Atlas Histórico do Brasil* (CPDOC, 2016) describes immigration in Brazil in a succession of waves: 1748-1800 (Azorean immigration only), 1800-1870, 1870-1930 and 1930-1970. Núcleos coloniais became the preferred alternative to address the shortage of labor from the 1870s onward (Martins, 1973; Rocha et al., 2017). They allow us to verify that it was indeed immigration in the corresponding time period that facilitated the vote for emancipation. There are two issues with using núcleos coloniais as a measure of immigration. First, only a fraction of immigrants settled in núcleos coloniais, and we cannot control for a selection effect in our data. Second, parliamentary records establish that their location was determined in the light of emancipation, i.e. endogenously, and even possibly as rewards or punishment for voting patterns among districts.

We use the following specification with the establishment of state-sponsored settlements as another measure of the local presence of immigrants:

$$F^{-1}\{\mathbb{E}[P_{ijv}|\tilde{\mathbf{x}}_{ijv}, \zeta_j, \delta_v]\} = \zeta_j + \delta_v + S_{ij}^{1872}\beta + \sum_t N_{ij}^t \lambda^t + \sum_t N_{ij}^t \times S_{ij}^{1872}\mu^t + \mathbf{x}'_{ijv}\boldsymbol{\gamma}, \qquad (2')$$

where N_{ij}^t is the number of núcleos coloniais located in district i during period t, with $t \in \{1748\text{-}1800, 1800\text{-}1870, 1870\text{-}1930\}$. The marginal effect of a núcleo colonial during our main period of interest is $\partial P_{ij}^{1884-1888} / \partial N_{ij}^{1870-1930} = \lambda^{1870-1930} + S_{ij}^{1872}\mu^{1870-1930}$, and an alternative formulation of hypothesis 2 corresponds to $\mu^{\text{1870-1930}} > 0$.

Table 23: H2' – Núcleos coloniais and voting decisions

	$\mathbb{1}$(Abolition vote)
1872 Share of slaves	-0.976***
	(0.276)
1748-1800 settlements	0.087
	(0.063)
1800-1870 settlements	-0.051
	(0.041)
1870-1930 settlements	-0.017
	(0.016)
Sh. of slaves × 1748-1800 settlements	-0.225
	(0.303)
Sh. of slaves × 1800-1870 settlements	0.284**
	(0.129)
Sh. of slaves × 1870-1930 settlements	0.293*
	(0.162)
Controls and FE	All
Observations	1,284
R^2	0.416

Note: *p<0.1; **p<0.05; ***p<0.01. District-vote two-way clustered standard errors in parentheses. Controls: population density, coffee, sugar, and cotton suitability, rainfall, longitude and latitude, distance to coast and to river, human mobility index, party affiliation, share of free colored and of literacy, gold mining and distance to diamond mine.

In line with our hypothesis, a larger number of immigrant settlements close to the abolition period in a district with a high prevalence of slavery is associated with a higher probability that the district's representative votes in favor of abolition bills. Although the effect associated with the interaction term is slightly less statistically significant for the 1870-1930 wave than the 1800-1870 wave, the former is arguably much more significant in magnitude (despite similar coefficients). A SD increase in the number of núcleos coloniais during 1800-1870 (0.66) is associated with a $0.66 \times (-0.05 + 0.28 \times 14.7\%) = 0.58$ percentage points decrease in the probability to vote in favor of emancipation in an average district. In turn, a SD increase in the number of núcleos coloniais during 1870-1930 (3.64) is associated with a $3.64 \times (-0.02 + 0.29 \times 14.7\%) = 8.2$ percentage points increase in the probability to vote in favor of emancipation in an average district.

A.5 Database construction and variables creation

This section provides additional details on how we proceed to build our legislative database and generate our variables of interest.

Abolitionist voting behavior. Our main dependent variable captures the voting behavior on emancipation-related bills of legislators in the lower house of the Brazilian parliament (the *Câmara dos Deputados*, Chamber of Deputies). To build this variable, we explored the archival records of parliamentary debates from the onset of the eighteenth legislature (1882) to the end of the twentieth legislature (1889) and identified every occurrence of roll-call vote relative to slavery bills. Across this three legislatures, we identified thirteen such events of nominative voting. In order to make it into the final data set, occurrences of nominative voting had to be *i)* clearly related to the emancipation of slaves and *ii)* discussed in the parliamentary records in a way that allowed clear-cut identification of how the vote related to emancipation.[23] The votes we retain in our final data set are the following:

- 03/06/1884: First motion of no confidence against the new *Presidente do Conselho*, nominated by the Emperor with the known objective to push forward the gradual emancipation of slaves. (Abolitionists vote against.)
- 30/06/1884: The roll is called on the decision to delay discussions on the Empires' finances until after the government's project related to labor emancipation is presented to the Chamber. (Abolitionists vote against.)
- 15/07/1884: Dantas (*Presidente do Conselho*, in charge to build a proposal for the gradual abolition of slavery) finally presents his project to the chamber. This occasions a large upheaval, in particular because the proposal fails to mention any compensation to slave owners. This leads the President of the Chamber (Moreira de Barros) to offer his resignation, on which the roll is called. (Abolitionists vote in favor.)
- 28/07/1884: New motion of no confidence against Dantas. The Chamber rejects the government's proposal on the emancipation of slaves and withdraws its confidence from Dantas. (Abolitionists vote against.)
- 13/04/1885: New motion of no confidence, this time ending in a tie. (Abolitionists vote against.)

[23] In other words, the annals must have made clear that voting in favour/against the bill meant being in favour/against the emancipation of slaves.

- 13/07/1885: The roll is called on a bill designed to increase the State valuation of slaves aged between 60 and 65 years old. (Abolitionists vote against.)
- 14/07/1885: The roll is called on an amendment related to the unconditional abolition of slaves aged 60 years or older. (Abolitionists vote in favor.)
- 27/07/1885 (1): The roll is called on a project regarding slave prices. (Abolitionists vote against).
- 27/07/1885 (2): The roll is called on a project regarding the manumission of disabled slaves. (Abolitionists vote in favor).
- 13/08/1885: The roll is called on the final version of the bill that would later become the *Lei dos Sexagenários*. (Abolitionists vote in favor.)
- 04/05/1885: New motion of no confidence. (Abolitionists vote against.)
- 05/05/1887: First project regarding the complete abolition of slavery. The roll is called on allowing it to proceed without hindrance. (Abolitionists vote in favor.)
- 09/05/1888: The roll is called on the project later known as the *Lei Áurea*. (Abolitionists vote in favor.)

Among these votes, we retain three into our core legislative data set, one by legislature: the motion of no confidence against Dantas on July 28, 1884 (this precipitates the end of the eighteenth legislature and the annals could hardly be clearer about it being related to emancipation), the vote on the *Lei dos Sexagenários*, and the vote on the *Lei áurea* (these two are the only two bills that become laws in their own right). Our goal in building this secondary data set (which we use only to assess selection issues) is to abstain—as much as possible—from any discretionary judgment, and we thus only keep the most important vote within each legislature.

We believe that together, the thirteen occurrences of nominative voting outlined above constitute, for the three legislatures considered, the universe of roll call votes with clear cut interests related to emancipation. In each of these cases, the annals of the *Câmara dos Deputados* provide the nominative list of legislators voting in favor or against (yeas and nays). We record these names and votes and match each legislator with the electoral district she represents. This is mostly done using records of *Juntas Verificadoras de Poderes* (special councils that occur during the early months of each legislature), double checked using Nogueira and Firmo (1973). We then link each district with the municipalities and parochias (the lowest administrative unit) that comprise them using the transcript of the 1881 electoral reform in Jobim and Porto (1996), which details the Empire's electoral division after the 1881 Saraiva law. Ultimately, we are thus left with a mapping of 1881 parochias and municipalities into 1882-1889 district-legislator-vote observations.

Demographic variables. Our demographic variables come from Brazil's first nation-wide demographic census in 1872 (Brazil, 1874). The main challenge here comes from matching 1872 municipalities with 1881 municipalities. For increased precision, we implement a first matching procedure at the parochia level. Although less than ten years separate the census from the legislation, a significant number of parochias were created in this time frame. From the 1441 parochias the country counted in 1872, there were 1662 enumerated in the 1881 legislation. Most of these 221 parochias are added to existing municipalities, but several municipalities were also created in the meantime and several 1872 parochias had become municipalities by 1881. To improve the precision of our matching, we exploit (whenever possible) the IBGE (2010) Evoluçao da Divisao Territorial to (painstakingly) manually trace the genealogy of municipalities. Matched parochias are then aggregated at the level of the municipality and the district.

Geocoded data. We exploit several sources of geocoded data, most notably from IIASA/FAO (2012), Nunn and Puga (2012), Özak (2010, 2018) and INCRA (2020). Information provided by these sources are used in combination with the IBGE's municipality-level boundaries for the year 1872. This allows us to compute zonal statistics and distance measures at the municipality and district levels. In particular, for each administrative unit, we compute: land suitability, climatic and topographic measures, major towns' and centroids' geographical coordinates, remoteness measures (distance to the coast, to provincial capitals and to Rio de Janeiro), population density (using Brazil (1874)), and 'frontier openness/distance to freedom' proxies (e.g. number and size of quilombos and, in combination with Fundação Palmares (2020), average distance to the closest municipalities with at least k quilombos).

Linking our matched 1872-1881 municipalities-to-districts data to the 1872 municipality grid allows to draw district-level maps of voting patterns on the bills we consider. Aggregating municipality boundaries into districts offers an additional matching challenge, as 1872 municipalities do not perfectly map into 1881 districts. Two issues may arise: *1)* Some municipalities are actually comprised of several districts. This occurs for some large cities, e.g. Salvador (two districts, BA1-2) and Rio de Janeiro (three districts, RJ1-3). *2)* Some 1881 districts include municipalities that were created/whose territory was altered after 1872, in which case apparent inconsistencies may occur. For example, the province of Amazonas is comprised of two districts in 1881, but the second district maps into two non-contiguous polygons based on the 1872 territorial division. This most likely occurs because the territorial division of the two (very large) municipalities that comprise the first ditrict of Amazonas (Manaus and Barcellos) shrunk (the former in particular) before 1881 as newer municipalities expanded. Hence, the resulting maps do not offer a perfect representation of the 1881 electoral division by any means, but they do illustrate the geographical configuration of abolition votes.

Finally, we also georeference a number of existing maps, in particular from Milliet (1941) and CPDOC (2016). Most notably, this allows us to approximate the location of state-sponsored settlements in a succession of waves and to assess proximity to Caminhos do Ouro, mining sites, insurrection sites, Tupi-Guarani-speaking populations' areas of enslavement, slavery/abolitionist interest groups and abolitionist journals.