Exit, Voice and Political Change: Evidence from Swedish Mass Migration to the United States*

Mounir Karadja
IIES, Stockholm University

Erik Prawitz
IIES, Stockholm University

JOB MARKET PAPER

December 13, 2015

Please click here for the most updated version.

Abstract

During the Age of Mass Migration, 30 million Europeans immigrated to the United States. We study the long-term political effects of this large-scale migration episode on origin communities using detailed historical data from Sweden, a major sending country in the period. To instrument for emigration cumulated over several decades, we exploit severe local frost shocks that sparked an initial wave of emigration, interacted with within-country travel costs. Because Swedish emigration was highly path dependent, the initial shocks strongly predict total emigration over the whole period. Our estimates show that emigration substantially increased membership in local labor organizations, the strongest political opposition groups at the time. Furthermore, emigration caused greater strike participation, and mobilized voter turnout and support for left-wing parties in national elections. We interpret these findings as an increase in citizens' bargaining power and demand for political change. Emigration also had effects on the supply of political change. Municipalities with more emigration exhibit higher welfare expenditures per capita, both before and after the introduction of democracy. In addition, local governments become more likely to adopt inclusive formal institutions. Together, our findings indicate that large-scale emigration can achieve significant and long-lasting effects on the political equilibrium in origin communities.

*We are indebted to Torsten Persson and David Strömberg for their invaluable commitment and guidance throughout this project. We thank Ran Abramitzky, Ingvild Almås, Konrad Burchardi, Björn Tyrefors Hinnerich, Supreet Kaur, Suresh Naidu, Arash Nekoei, Peter Nilsson, Per Pettersson-Lidbom, Imran Rasul, Jakob Svensson, Anna Tompsett and seminar participants at IIES, Stockholm University, the 8th International Conference on Migration and Development, the 2015 Summer School in Development Economics in Garda, and the 2015 PODER/CEPR Conference for very helpful comments.

1 Introduction

Institutions are widely believed to be important determinants of long-run development. Yet, less is known about what causes them to change over time. This paper proposes and empirically verifies that large-scale emigration can be a mechanism leading to political change in origin countries. Using one of the largest migration episodes in human history, the Age of Mass Migration, we estimate the long-run effects of emigration on local political outcomes.

Starting in the mid 19^{th} century, the Age of Mass Migration saw 30 million Europeans settle in the United States. We focus our attention on Sweden, which had one of the highest exit rates over the period. A quarter of its population, or about 1.3 million citizens, emigrated in the course of sixty years, mostly to the United States. Swedish economic and political elites were highly concerned about the newfound mobility of ordinary citizens. As a result, proposals to restrict emigration were continually made during the period, but were never put in place. Instead, the Age of Mass Migration coincided with a period of political development in Sweden. The dominant force in Swedish 20^{th} century politics, the Social Democratic Party, as well as the powerful labor union movement, were founded during the period and became key actors in reforming Swedish policy and political institutions.

It is unclear how emigration related to this development, however. Theoretically, the effect of emigration on an autocratic origin country is ambiguous. If democratic dissidents choose to exit the country rather than push for reforms, the result may be a lower level of political development (Hirschman, 1970). On the other hand, the option to emigrate may encourage activism if it can serve as an insurance against repression by elites.[1] Moreover, this effect is amplified if the elite's responsiveness to collective action increases with labor mobility, thus increasing the incentives for organizing.[2]

Motivated by this discussion, we deploy a wide range of data sources spanning the mass migration period to study the long-term political effects of emigration across Swedish municipalities. We organize our empirical analysis in two sections, broadly examining citizens' demand for and elites' supply of political change. On the demand side, we start by investigating the relationship between emigration and the local political organization of citizens. Our main outcome variable is per capita membership in the labor movement, defined as labor unions and the Social Democratic Party, which we observe yearly starting in 1890. To further probe the role of the labor movement, the membership data are supplemented with measures of the movement's strength and influence. First, participation in the general

[1] Section 2.2 elaborates on the risks involved in activism in our setting. A similar dynamic mechanism to the one described above is emphasized in the literature on migration-driven "brain gain", i.e. that the possibility of high-wage emigration may induce more people to pursue higher education (Beine et al., 2001).

[2] For example, regarding the out-migration of blacks from the US south, Myrdal (1944) writes that "the experience [suggests] that emigration of a significant number of Negroes is one of the surest ways of stimulating the Southern whites to give more consideration to the Negroes that remain in the South".

strike of 1909 involving 300,000 workers is used as a measure of direct, individually costly engagement. Second, we study voter turnout and vote shares for left-wing parties in national elections 1911–1921, allowing us to identify political mobilization as well as labor-related political preferences.

If emigration induces citizens to demand political change, a natural question is if this is reflected in local government actions. In the second part of our analysis, we therefore turn to estimating the effect of emigration on the supply of political change, using data on local redistribution policy and political institutions. We use data on welfare expenditures per capita to test if emigration resulted in changed patterns of redistribution across municipalities. Using data on the form of democracy chosen by local governments, we then test the hypothesis that emigration leads municipalities to adopt more inclusive formal institutions.

To establish causality, we exploit the fact that Sweden's mass emigration was sparked by a series of severe agricultural shocks in the 1860s, caused by unusually cold temperatures (Beijbom, 1995; Barton, 1994; Sundbärg, 1913). Using daily temperature data from this period, we measure the incidence of growing-season frost shocks 1864–1867, just prior to the onset of early mass migration. We then construct an instrument which only captures variation in the intensity of migration push factors: the interaction between frost shocks 1864–1867 and the shortest distance from a municipality to one of the two major emigration ports.[3] Using only the interaction term as our instrument allows us to control for both port distance and frost shocks themselves, which avoids picking up any confounding direct effects of severe economic shocks on political outcomes. Importantly, because Swedish emigration was highly path dependent, which we show, the instrument strongly predicts cross-sectional variation in total emigration across the 50-year sample period.[4] The instrument passes several exogeneity tests, including a randomization test on pre-determined covariates and placebo treatments using shocks in other periods. Shocks occurring in the *non-growing* seasons 1864–1867 have no effect on emigration or second-stage outcomes.

Our results show that municipalities that experienced more emigration during the Age of Mass Migration exhibit significantly increased demand for political change. Membership in local labor organizations is significantly higher starting in 1900, which provides a link between Sweden's mass emigration and the growth of its influential labor movement. This relationship is also reflected in our measures of organizational strength, as emigration leads to higher participation in the major general strike of 1909. Furthermore, we find increased voter turnout in national elections 1911 to 1921, as well as higher vote shares for left-wing

[3]See e.g. Morten and Oliveira (2014) and Quigley (1972) regarding the importance of travel costs for migration decisions.

[4]High degrees of path dependence in migration patterns is a canonical finding in the migration literature and has been found in numerous settings, see e.g. Massey et al. (1993), Hatton and Williamson (2002), Bryan et al. (2014), Giulietti et al. (2014) and McKenzie and Rapoport (2007).

parties in those same elections. Rather than inhibiting their use of voice, higher emigration led to more political coordination and left-wing preferences among ordinary citizens, and arguably resulted in a greater bargaining power vis-à-vis local elites.

Emigration also had an impact on policy and political institutions, in line with the preferences of the labor movement. Welfare expenditures per capita are significantly higher in municipalities with more emigration, both before and after the introduction of democracy in 1919.[5] In 1918, a weighted voting system in local elections gave wealthy individuals up to 40 votes, biasing decision making power towards economic elites. The observed increase in expenditures is therefore unlikely to have been caused by changes in the preferences of ordinary citizens. Rather, it is consistent with concessions being made by elites in favor of citizens.

During this period, local governments were organized either as direct or representative democracies. Recent evidence has shown that municipalities under representative democracy provided higher welfare expenditures, likely due to direct democracies being more easily captured by local elites (Hinnerich and Pettersson-Lidbom, 2014). We find that municipalities with greater emigration are more likely to adopt the more inclusive political institution between 1919 and 1938. As such transitions were required to last at least five years, and often lasted longer in practice, this finding is in line with the theory of institutional change as a commitment device (Acemoglu and Robinson, 2000, 2006).

Lastly, we test for the long-run persistence of the effect of emigration. A number of studies in economics and political science have found political preferences to be persistent within individuals as well as correlated across generations.[6] We find that labor movement membership remains higher at least until 1940 in municipalities with more emigration. More strikingly, we also find that left-wing voting patterns persist in contemporary elections, using data at both the municipal and national level between 1998 and 2014.

Our identification strategy has several attractive features. First, identification solely relies on the interaction term between frost shocks and the distance to emigration ports. This allows us to control for any potential direct effects on our second-stage outcomes related to frost shocks and port distance alone, which would otherwise violate the exclusion restriction and bias the IV estimates.[7] Second, it allows us to credibly estimate the long-term effect of an economic phenomenon taking place over several decades, rather than year-to-year effects. Third, it allows us to control for the differential impact of bad weather shocks between locations that are have more or less access to markets, which may be correlated with port

[5]These results are not explained decreased population, as results hold in expenditure levels as well.

[6]See e.g. Madestam et al. (2011), Alford et al. (2005) and Jennings et al. (2009).

[7]For example, it is possible that municipalities which were more affected by frost shocks developed more extensive social insurance systems as a result. However, even if such effects are persistent, they would be captured by controlling for the direct effect of frost shocks.

distance. We do so by interacting our measure of frost shocks with the distance to the nearest major *trade* port and the nearest town, respectively. Including these additional controls has an negligible effect on the estimates.

Emigration may have affected political outcomes through other channels than labor organization. Using the data sources available to us, including censuses, we evaluate the plausibility of a number of additional mechanisms. We first find that labor movement membership and strike participation are not driven by changes in the employment composition of municipalities towards industry. We also find no changes at the municipal level in terms of sex ratios, marriage rates, household size or in-migration from other municipalities. Moreover, there are no effects on the share of voting eligible citizens 1911–1921. These tests do not suggest any significant shifts in the composition of the high-emigration municipalities that can explain our findings. A bounding exercise also shows that even if emigration was highly skewed in terms of ideology, the effect of emigration of left-wing voting remains sizable and significant.

The existing migration literature has emphasized the potential transfer of attitudes from host to origin countries. We find indirect evidence against this mechanism in our setting, as there are no positive effects on membership in two types of organizations that were highly influenced by the US: non-state free churches and temperance lodges.

This paper relates to a nascent literature on the political effects of emigration in origin countries. One set of papers studies the effects of migration to democratic countries, finding a positive effect of such migration on democratization (Spilimbergo, 2009) and voting for an opposition party (Omar Mahmoud et al., 2015; Pfutze, 2012). Batista and Vicente (2011) find that households in Cap Verde with more migration experience exhibit a higher demand for good governance. Docquier et al. (2014) use cross-country data for developing countries to measure the effect of migration flows on the institutional quality of origin countries. In the context of the Age of Mass Migration, Hirschman (1978) notes the co-occurrence of emigration to the United States and a wave of European democratization.

Our contribution is three-fold. First, we exploit plausibly exogenous variation in order to identify the causal effects of emigration on political outcomes. Although current studies are aware of the challenges in identifying a causal effect, there is still a lack of well-identified estimates. Second, we focus on the mechanism of emigration improving the organizational and bargaining strength of citizens rather than a transfer of attitudes. Third, we contribute by following-up on citizens' response to emigration and showing how policies and institutions change within the same setting.

We also relate to the literature on institutions in economic development, and political change in particular. There is a large literature investigating transitions into and from democ-

racy at the national level.[8] Besley et al. (2015) study the effect of incumbents' probability of staying in power on institutional reform. A growing set of papers investigates the effects of trade on institutional change (Acemoglu et al., 2005; Dippel et al., 2015; Puga and Trefler, 2014; Sánchez de la Sierra, 2015). The importance of factor mobility for institutional change has been studied theoretically, though mainly focusing of the mobility of capital rather than labor (Boix, 2003). Acemoglu and Wolitzky (2011) show that greater outside options for workers improves their equilibrium outcomes under coercive institutions. This paper also relates to the literature on groups and voter turnout, in finding a co-occurrence of labor movement size and voting (Morton, 1991).

The remainder of the paper proceeds as follows. Section 2 provides an overview of Swedish mass emigration and describes the historical evidence regarding the cause of its onset in the 1860s. The labor movement and its relationship to emigration are also described. Section 3 describes our data. Section 4 introduces the econometric framework and our identification strategy. Sections 5 to 7 discusses the first-stage relationship as well as the effects of emigration on our political outcomes. Section 8 considers alternative mechanisms that could explain our findings, while Section 9 performs a series of robustness checks on our main specifications. Section 10 tests for longer term persistence in labor movement membership and left-wing voting. Section 11 discusses our results and presents some simple cross-country evidence on US emigration and union density.

2 Background

2.1 Swedish mass emigration to the United States

Starting in 1850, the Age of Mass Migration saw 30 million Europeans settle in the United States. Under its policy of free immigration, individuals from all over the world were allowed permanent residency in the United States. Sweden was one of the biggest sending countries in per capita terms, along with Ireland, Italy and Norway (Taylor and Williamson, 1997). A total of 1.3 million Swedes emigrated from 1860 to 1920, corresponding to one quarter of the average population over the period.

Swedish emigration took off abruptly at the end of the 1860s. In the peak year of 1869 alone, nearly 1 percent of the population emigrated and in the years between 1867 and 1879, 200,000 Swedes left their home country. We refer to the sharp increase in emigration 1867–1879 as the *first wave of mass emigration*.[9] The spike of the first wave is evident in Figure

[8] See Persson and Tabellini (2009) and studies cited therein.

[9] Earlier emigration was uncommon – in 1865 the Swedish American population was estimated at 25,000 (Barton, 1994). Poor communications may have held back potential emigrants, as crossing the Atlantic was expensive and time-consuming. Sailing was the predominant means of transport and traveling from Sweden

1, which displays per capita emigration rates over the period. The causes and timing of the Swedish mass emigration episode have been widely discussed by historians. Central to the existing accounts is the series of bad harvests in the 1860s, caused by unusually poor weather conditions, which led to widespread poverty and served as a catalyst for emigration on a large scale (see e.g. Beijbom, 1995; Barton, 1994; Sundbärg, 1913).[10] In particular, cold weather led to a high incidence of frost as nighttime temperatures fell below zero degrees Celsius, even during the regular growing season. The spring of 1867 saw the most extreme weather, in some cases lasting well into the summer months.[11]

The famine years were particularly harshly felt, as agriculture was the main source of food and income for most citizens: in 1865, 83 percent of the population lived in rural areas and only 11 percent of the labor force worked in manufacturing (Edvinsson, 2005). Cities and towns were affected indirectly, however, as the supply of food and the demand for goods and services dropped (Beijbom, 1995). In our data set, 28.3 percent of emigrants 1867–1920 are from urban areas. Figure 2 displays Swedish GDP and real wages at the national level 1800–1900. There is a visible trough for wages and production around 1867.[12]

Later emigration waves occurred during the 1880s and at the turn of the century, as seen in Figure 1. This pattern was common throughout Europe and has been linked to inversely developing business cycles across the Atlantic during this period (Hatton, 1995). In particular, differences in growth rates between the United States and Sweden predict aggregate emigration flows between 1870 and 1910 (Bohlin and Eurenius, 2010). Social networks were also crucial drivers of emigration in the later waves. First-hand accounts of Swedes in the US reveal that many would not have emigrated if it were not for having family members overseas (Sundbärg, 1913). Having emigrants in one's network reduced uncertainty and lowered the costs associated with traveling to the United States and finding an occupation once there (Runblom and Norman, 1976). Postal communication was well-developed and emigrants frequently sent home pre-paid tickets for family members to join

to North America took up to two months.

[10]Sweden's case is similar to that of Ireland, whose first emigration wave was caused by a famine (Hatton and Williamson, 1993).

[11]The month of May 1867 is the coldest known May in Swedish history and the meteorological summer (five days in a row with temperatures above 10 C) started only in mid-June in many parts of Sweden (SMHI, 2013). In Finland, the temperatures observed during the spring of 1867 have a 1 in 500 probability of occurring (Jantunen and Ruosteenoja, 2000).

[12]Several factors are likely to have interacted with the poor harvests in sparking the first wave of mass emigration to the United States. The introduction of steam ship technology led to a shift away from sailships in the late 1860s and the cost of migration fell considerably. The US Homestead Act of 1862, which offered free land to immigrants, together with the end of the US Civil War in 1865 are also considered to have contributed to the large number of emigrants observed (Barton, 1994).

them in America.[13] Pre-paid tickets accounted for up to half of all travelers.[14]

The mass emigration of Swedes did not go unnoticed among elites. Policies to reduce emigration were proposed throughout the period, and applied a mix of carrot and stick strategies: proposals to restrict emigration were common, as were calls for improving living standards so as to induce citizens to stay. In 1869, at the height of the first emigration wave, several motions were raised in parliament by MPs from high-emigration counties. Even at this very early stage, the awareness of and concern about emigration was high (Kälvemark, 1972). The central government later surveyed governors about their counties' experiences with emigration. A majority of governors then agreed that emigration was a net bad for the country (Kälvemark, 1972). When asked for policy proposals to reduce emigration, governors suggested measures to both make emigration more difficult and to improve the conditions in Sweden, for example by facilitating the procurement of small land plots by landless farmers.[15] However, emigration remained essentially unrestricted throughout the mass migration period.

The return of high emigration rates in Sweden in the early 20th century brought the strongest political reactions yet. Landowners and agrarian interest groups worried about labor scarcity and identified emigration as the main culprit.[16] Others were concerned about the emigration of young men who would otherwise perform military service, and worried about a deterioration in national defense (Kälvemark, 1972). These concerns eventually lead to the appointment of a large public commission, assigned the task of finding measures to end the mass emigration. When its 21 volume report was published in 1913, it recommended political reforms to improve the conditions of ordinary citizens to induce them to stay, rather than suggesting emigration restrictions. The large-scale emigration of Swedes ended in the 1920s, as the United States introduced quotas on immigration.

[13]Data from Denmark, which had a much smaller number of emigrants than Sweden, has shown that up to 1.8 million letters were sent yearly to Denmark from the US (Beijbom, 1995).

[14]Studies of the archives of the Larsson Brothers emigration agency in Gothenburg have shown that around half of their clients traveled using pre-paid tickets (Runblom and Norman, 1976). Beijbom (1995) also reports that half of the Swedish emigrants traveled on pre-paid tickets at the beginning of the 1880s, and around 40 percent by the end of that decade. Pre-paid tickets also accounted for 40 percent of Norwegian travelers (Hvidt, 1975).

[15]The survey was carried out in 1882. Governors also identify family ties to emigrants as a chief determinant of emigration.

[16]Noting that landowners were less likely to emigrate, the state was encouraged to sell smaller plots of land and provide a transfer to enable poor farmers to acquire land. The plot size was a key parameter, however, as transfers were not intended to allow farmers to become self-sufficient but to remain attached to major landowners' farms. In a parliamentary debate in 1904, the Minister of Agriculture openly discussed the central point of contention: should the subsidy be so large that it allowed a farmer to be self-sufficient or should it be smaller, so that "owners would invariably need to seek employment with others in order to earn a living (Kälvemark, 1972).

2.2 The labor movement and emigration

The Social Democratic Party was a dominant actor in Swedish politics during the 20th century and long garnered a near majority of votes in national elections. Founded in 1889, it entered government for the first time in 1917 and remained in government for most of the 20th century.[17] The Social Democrats were closely linked to the Swedish Trade Union Confederation (LO), founded in 1898 as a central organization for the many smaller unions that existed at the time. Both organizations championed the right to organize, the 8 hour workday and universal suffrage (Lundkvist, 1977).[18]

The labor movement was regularly in conflict with employers and was known to use emigration as a tool. In Stockholm, labor unions held English courses and helped colleagues emigrate. The Social Democrats' main newspaper updated readers about prospects in the US labor market (Tedebrand, 1983). After the general strike in 1909, which was considered to be a defeat for the labor movement, a socialist newspaper called upon workers to emigrate (Beijbom, 1995). Many emigrated labor activists continued their work overseas, founding labor organizations in the United States (Bengston and Brook, 1999; Nordahl, 1994).

Emigration might have been useful for the labor movement because of the high risks involved in labor activism. Workers could be fired, evicted and blacklisted for being union members. Until 1885, an anti-loitering law made striking illegal and punishable by forced labor (Westerståhl, 1945).

A case study of the town of Ljusne elucidates the conflicted interactions between labor, elites and emigration. In 1906, more than a hundred workers emigrated from Ljusne, following a clash between the local Social Democratic club and the main employer, who owned all buildings in town and disallowed political and union organizing among workers. After the Social Democrats sent an incendiary telegram to the Swedish King, leading figures were fired while others were intimidated via the local police to stop their activities or be evicted. Rather than complying, many opted to emigrate. The option of emigrating was facilitated by the town's history of US migration – it had experienced large participation in the emigration waves of the 1860s and 1880s. Regarding the choice of emigrating rather than relocating within the country, one of the central activists later commented that "strangely enough, there were only two places for us in the world then, Ljusne or America". The news of Ljusne's "mass emigration" became widely spread in national media at the time and severely hurt the reputation of the owner and first chamber parliamentarian Count Walther von Hallwyl (Rondahl, 1985). When the plant shut down in 1907, the company announced

[17] The party was in government between 1932 and 1976 without interruption.

[18] The 8 hour workday and universal suffrage were adopted in 1919 by a Liberal-Social Democratic coalition government. In the 1938, LO become a key player in the Swedish Labor Market Model, representing employees in collective bargaining over wages and benefits without intervention of the national government.

that it would be paying pensions to older workers in gratitude for their service. The Ljusne case illustrates the use of emigration among labor activists, and indicates that it could serve as an insurance mechanism, seeming to empower citizens who would otherwise not dare to object to employers' demands.

3 Data

Emigration Data We compile local emigration histories using two distinct, individual level data sets encompassing the universe of registered emigrants during the Age of Mass Migration. The final data set contains 1.1 million emigrants from 1867 to 1920. To our knowledge, this is the first study to make use of any of these two data sources for disaggregated statistical analysis. They are described in detail below.

The State Church in Sweden was historically tasked with tracking demographic statistics in their parishes. Births, deaths, marriages as well as migration information were recorded year by year at the individual level and stored in parish records. These were later incorporated by the central statistical agency. We obtain emigration data from these parish records that were digitized by family researchers and through various municipal and county efforts.[19] Individual migrants are matched to an origin municipality and year using information on the date of exit and home parish available in the data set.

The second source of individual level emigration data is from archived passenger lists kept by shipping companies. Starting in 1869, at the peak of the first emigration wave, ships with foreign destinations were required by law to compile lists of all their passengers (Clemensson, 1996). The lists were controlled for authenticity by the police who checked off travelers as they boarded their ships. The passenger manifests were later stored in various city archives and were digitized by the Gothenburg Provincial Archives. The same matching procedure as the parish level data is used to match emigrants to origin municipalities. The data provide a less precise "home town" location rather than the exact parish, leading to lower match rates.

Since our two data sets are independently collected and record emigrants at different points in time, they afford us a rare opportunity to measure the accuracy of our data. Appendix Section A.2 shows that there is a high degree of within-year similarity between the data sources. This indicates a high reliability of the emigration numbers and that there is no important lag between leaving the home parish and boarding a ship to the United States.

In the remainder of the paper, we use a single emigration variable defined as the maximum of either the church book or passenger list data each year. The primary concern is in

[19]The data are obtained from The Swedish Migration Center in Karlstad, Sweden.

undercounting emigrants, and since the passenger list data are imperfectly matched, using the maximum value each year yields our best estimate of emigration.[20]

Election and Labor Movement Data Municipal level voting data for all national elections between 1911 and 1921 are taken from Berglund (1988).[21] The data set includes the number of eligible voters and votes cast as well as the distribution of votes across political parties.[22] Precinct-level data from municipal and national elections 1998 to 2014 are taken from the Swedish Election Authority and are geographically matched to 1865 municipality borders.

Local organization membership 1881-1945 comes from the Social Movement Archive.[23] The Social Movement Archive lists the number of members by municipality as of December 31 each year, for the following organizations: free churches, temperance lodges, labor unions and the Social Democratic Party. We group labor unions and the Social Democratic Party into one variable that we label *labor movement membership*.

Participation numbers for the 1909 general strike, divided by union and non-union members, are digitized from the original government report following the strike (Kommerskollegii, 1910).

Weather Data Daily temperature data are obtained from the historical records of the Swedish Meterological and Hydrological Institute. We complement this with daily data for Norwegian weather stations near the Swedish border, provided by the Norwegian Meteorological Institute. The Swedish data contain temperature readings three times per day: 6 am, 12 pm and 8 pm. In addition, most observations have daily minimum and maximum temperatures. The Norwegian data contain daily average temperatures only. Appendix A.1 describes how daily minimum temperatures are predicted from existing data in cases when the minimum temperature is not available.

In total, the data contain 32 unique temperature stations between 1864 and 1867, with a median distance from municipality centroids to the nearest station of 36 kilometers.[24] The relatively small number of stations could be a problem for our ability to find enough

[20]Note that after 1895, all data are necessarily from passenger lists since church books have not been digitized after that year.

[21]Provided through the Swedish National Data Service (SND).

[22]The data begin in 1911 as it was the first year when party denominations were formally required of all members of parliament. Before then, the parliament consisted of a mix of partisans and independents and partisanship was not systematically recorded. In the absence of roll-call data from the period, this makes it hard to determine the political identification of MPs before 1911. Roll call data from the Parliament were not recorded until 1927.

[23]The data were collected by historians at Uppsala University (Andrae and Lundqvist, 1998). Provided through the Swedish National Data Service (SND).

[24]The mean distance is 39 kilometers.

variation in weather conditions to precisely predict emigration. However, temperature is known to be evenly distributed over large areas, especially in the northern hemisphere. Rain is, by comparison, more idiosyncratic (Dell et al., 2014). Climatologists have also established that temperature deviations from long-run means are more similar over large distances as compared to levels (Hansen and Lebedeff, 1987). Intuitively, the reason for this is that even if two neighboring locations have different temperature levels, e.g. due to differences in altitude, they are likely to experience similar deviations from their long-run means within a given window of time due to common weather shocks. As our identification strategy relies on estimating shocks to weather, we are precisely interested in using deviations, allowing us to exploit this feature of the data. Section 4 describes how we define frost shocks in detail.

Additional Data In the final data set, all variables are aggregated to the municipality level using 1865 boundaries. Georeferenced data on administrative borders in 1865 are taken from the National Archives of Sweden. Distance to an emigration port is defined as the shortest distance to either Gothenburg or Malmö, the two main emigration ports during the Age of Mass Migration.[25] Population data were kindly shared by Lennart Palm (Palm, 2000). Soil suitability data (for barley, oats, wheat, livestock and forestry), used as control variables, are from the FAO GAEZ database. County-level harvest grades 1860 to 1870 are from Hellstenius (1871). The data set grades harvests yearly on a scale from 0 to 6, with higher values indicating larger yields.

Municipal level welfare expenditures and type of political institutions (direct or representative democracy) are taken from Hinnerich and Pettersson-Lidbom (2014). Mortality data for infants, children and mothers, averaged over the 1850–1859 period, are from the The Demographic Data Base, CEDAR, Umeå University. Complete decennial censuses for 1880-1920 were obtained from the National Archives of Sweden and the North Atlantic Population Project. The census gives population-wide data on demographic variables including gender, civil status, family structure, and occupation. Summary statistics are presented in Table 1.

4 Empirical Framework

Our goal is to estimate the effect of emigration over the course of the Age of Mass Migration on long-run political outcomes in origin municipalities. The cross-sectional equation of interest is

[25] All distances are calculated using the great circle haversine formula. The results are robust to excluding lakes and waterways between municipalities and Gothenburg or Malmö. Figure 5 shows that the distance to Gothenburg and Malmö is well approximated by a straight line for most locations in Sweden.

$$y_{mct} = \beta Emigration_{mct} + \phi_c + \mathbf{X}'_{mc}\beta_X + \varepsilon_{mct}, \tag{1}$$

where y_{mct} is a political outcome in municipality m, county c and year t, $Emigration_{mct}$ is the log of cumulated emigration from 1867 to year t, ϕ_c is a fixed effect for the 24 counties and $\mathbf{X}_{mc}$ is a vector of municipality characteristics determined before the start of mass emigration. The specification focuses on the stock of emigrants as a determinant of political outcomes. Relating back to the above discussion, this variable captures the extent of overseas social networks present in a municipality and hence, the ease of future migration for current citizens. Throughout the paper, we estimate (1) by OLS as a baseline and reference for comparing other estimates, always including the log of population in 1865 in $\mathbf{X}_{mc}$ in order to scale the level of emigrants to the initial municipality size.

For several reasons, long-run emigration histories can be expected to correlate with important characteristics of the origin municipality, either observable or unobservable, that can have a direct impact on the outcomes of interest. A strong concern in estimating (1) by OLS is hence that it may yield biased estimates of the effects of emigration. In particular, the risk of picking up reverse causation is high. Locations with favorable initial institutions may induce more emigration because of better access to information or higher incomes. By contrast, places with more repressive leaders might actively prohibit emigration, thus leading to a positive bias in the OLS estimate of β. In the abstract, the reverse situation is, however, equally likely: fewer people may want to leave locations with good institutions and bad institutions could act as a push factor for emigrating. Without the ability to quantify the relative importance of these effects, OLS estimates yield limited information about the causal effect of emigration on local politics.

To overcome the issues related to omitted variables and to consistently estimate parameters, we propose an identification strategy exploiting only migration-related push factors prior to the first wave of mass emigration: the interaction between growing-season frost shocks 1864–1867 and the distance from a municipality to the nearest of the two main emigration ports. The remainder of the section describes how we construct frost shocks and presents the instrumental variables strategy in more detail.

Frost shocks The empirical economics literature often uses rainfall as source of exogenous variation in income for developing countries, motivated by the idea that rainfall has a direct effect on crop yields. Somewhat less attention has been given to the importance of temperature variation. However, low temperatures and frost in particular are closely linked to agricultural outcomes in non-tropical climates (Snyder and Melo-Abreu, 2005). Frost has severe effects on crop growth and the likelihood of plant death. In the United States, more economic losses are caused by freezing of crops than by any other weather hazard (White

and Haas, 1975). The perniciousness of frost is linked to its non-linear effects once temperatures fall below zero degrees Celsius. One night of freezing temperatures can lead to a complete crop loss (Snyder and Melo-Abreu, 2005). As mentioned in Section 2, the poor harvests in Sweden in the 1860s occurred during years with unusually cold temperatures in the growing season. Throughout Sweden, frost was observed as late as in June, in the middle of the growing season for most municipalities in our data. Estimating the incidence of frost is difficult, however, as it does not only require daily data but also estimates of the *minimum* temperature at a daily resolution.

Our measure of frost shocks follows the approach of Harari and La Ferrara (2013). It defines a binary shock indicator by month, and expresses shocks relative to the local long-run weather in that particular month. Shocks are constructed as follows. For each month r, we calculate the total number of frost days, defined as days with a minimum daily temperature below zero degrees Celsius. At the weather station level, we compute a series of monthly deviations from the mean,

$$deviation(Frost\ Days)_{srt} = Frost\ Days_{srt} - \overline{Frost\ Days}_{sr},$$

where $\overline{Frost\ Days}_{sr}$ is the long-term mean of frost days per calendar month r at station s. Each municipality is then matched with the nearest station available in each month.[26] This is used to compute the municipality's long-term standard deviation of frost days in each month, $sd(Frost\ Days)_{mr}$. A monthly frost shock at the municipality level is then defined as a binary variable:

$$Shock_{msrt} \equiv I[deviation(Frost\ Days)_{srt} > sd(Frost\ Days_{mr})], \tag{2}$$

where $Shock_{msrt}$ is an indicator equal to one if municipality m, whose nearest station is s, experienced a positive frost shock in month r of year t. Note that we compute the deviation from the long-term mean at the *weather station* level rather than the municipality level. This exploits the fact that weather variables are more precisely interpolated in deviations from long-term means than in levels, as discussed in Section 3 (Hansen and Lebedeff, 1987). Given that we are exactly interested in anomalous temperature variation, this feature increases accuracy of our measures. Finally, we sum the number of shocks over the growing season for each municipality over the 1864-1867 period. A growing season month is defined as a month with a long-term mean temperature above 3 degrees Celsius, following guidelines of the Swedish Meteorological and Hydrological Institute. The frequency distribution of frost

[26]Enough variation is captured by the nearest station that using more weather stations (e.g. the second and third nearest ones) does not contribute any additional information. In our data, the adjusted R^2 from regressing monthly frost days at weather station s on frost in the nearest neighboring weather station is slightly lower when we add the frost of the second nearest weather station.

shocks 1864–1867 is displayed in Figure 4. As evidenced by the figure, this period saw a high incidence of cold temperatures in the growing season, with the median municipality experiencing three frost shocks. Figure 5 displays the spatial distribution of growing season frost shocks 1864–1867, indicating considerable variation in shocks across Sweden.

Identification strategy In order to consistently estimate β in (1), we instrument for emigration using the number of growing season frost shocks 1864–1867 interacted with the log distance to the nearest emigration port. We only exploit shocks occurring during this four-year period as it was bookended by a particularly high incidence of cold temperatures, with shocks rarely occurring in other years of the decade. The direct effects of frost shocks and port distance are used as controls. The first-stage equation is

$$Emigration_{mct} = \gamma_1 Shocks_{mc} + \gamma_2 DistPort_{mc} + \gamma_3 Shock_{mc} \times DistPort_{mc} + \theta_c + \mathbf{X}'_{mc}\gamma_X + v_{mct},$$

where $Emigration_{mct}$ is the log of cumulated emigration from 1867 to year t in municipality m, $Shocks_{mc}$ is the number of frost shocks experienced prior to the first wave of emigration, $DistPort_{mc}$ is the log distance to the nearest emigration port and θ_c is a county fixed effect. Because frost shocks are constructed to capture random variation with respect to fixed municipality characteristics, the coefficient of interest, γ_3, is estimated without bias.[27]

Distance to emigration port is defined as the shortest distance to Gothenburg or Malmö, the two main emigration ports.[28] Likely due to economies of scale, the points of exit were very concentrated, and between them the cities handled more than 95 percent of all emigration before 1920. Their importance is confirmed by comparing yearly emigration shares across ports.[29] Figure 3 displays the share of emigrants exiting through four ports over the period 1869 to 1920. Gothenburg was the biggest port by far throughout the period, with 79 percent of all traffic on average and about 82 percent during the first wave of emigration. Malmö was the second largest emigration port with 18 percent of emigrants on average and 14 percent during the first wave.[30] Stockholm, the capital and Sweden's largest city by far, was less suited for emigration because of its location on the eastern coast of Sweden. Its port averaged 2 percent of total emigrants. Similarly, Norrköping, the third largest city and an important trade port, was minor in terms of emigration.[31] In our data set, 75 percent of municipalities have Gothenburg as their closest emigration port, while the rest are closer to

[27] This implication is tested below.

[28] All results are robust to using levels of distance instead of logs, see Appendix Table A.8.

[29] Shares are computed using the passenger list data, which includes the port of exit for all emigrants.

[30] The data distinguish between emigrants from Malmö and Copenhagen. Due to their geographical proximity and because most emigrants likely transited via Malmö before being registered in Copenhagen, we count the two exit ports as one unit.

[31] Gothenburg and Malmö were the second and fourth largest cities in 1865, respectively.

Malmö.

Exclusion restriction The identification strategy only relies on the *interaction term* of frost shocks and port distance. This has two main advantages. First, a basic cost-benefit analysis would suggest that potential migrants let the cost of traveling to the emigration port factor into their decision. By implication, including it in the empirical model should improve the explanatory value of the model.[32] Second, and more importantly, it allows us to control for the direct effects of distance to the port as well as the frost shocks themselves. A typical complication in studies that use weather shocks as instruments for some endogenous variable is that weather may simultaneously affect many variables, including citizens' values and attitudes (Giuliano and Spilimbergo, 2014).[33] Hence, there are potential direct effects of the shocks on our variables of interest, which would violate the exclusion restriction and invalidate the use of the shocks as an instrument. By using only the interaction term, we are able to isolate exogenous variation in shocks that is solely related to migration push factors.

For the identifying assumptions to hold, it is nevertheless required that no variables other than emigration correlate with the instrument. We test this by performing a simple randomization test of the instrument on a number of observable characteristics of municipalities. Table 2 displays the outcome of these tests. The instrument is uncorrelated with all variables but one, log population in 1865. By random chance, we should expect some variable to be correlated with the instrument. Yet, it is reassuring that the correlation predicts that high-emigration municipalities have *lower* baseline population levels, while it is expected that larger municipalities are more politically organized.[34] Nevertheless, the 1865 population is included in all regressions as a control. We include the following additional control variables: log area, latitude, longitude, the share of arable land in 1810, an urban indicator, as well as indicators for high soil suitability for the production of barley, oats, wheat, dairy and lumber. We also include the following distance measures, all in logarithms: to the nearest weather station, to the nearest railway, to Stockholm, to the nearest town and to the nearest of the ten most important trade ports in 1890. The three mortality variables at the bottom of Table 2 are not included as control variables in our regressions due to a lower number of observations. They provide relevant tests of the instrument, however, as they directly relate to municipal policy and wellbeing.[35] Reassuringly, the instrument is not correlated with any

[32]Beijbom (1995) highlights the importance of travel possibilities, noting that the northern regions of Sweden were hit hard by the bad weather in the famine years, while most emigrants came from southern Sweden.

[33]For example, Sarsons (2015) shows that rainfall might have effects on conflict through other channels than agricultural yields, invalidating its use as an instrument for income.

[34]Indeed, OLS regressions show that the population in 1865 is weakly positively correlated with labor organization rates and welfare spending, while it is unrelated to support for left-wing parties.

[35]Maternal mortality was partially a function of local policies, as midwives were employed by parishes

measure of mortality, for infants, children or mothers.

5 Frost shocks, travel cost and emigration

Before investigating the link between the instrument and emigration, we verify the effect of frost shocks on agriculture using a panel of county-level harvest grades from 1860 to 1870. Column 1 of Table 3 shows that frost shocks in the growing season indeed cause worse harvests in the same year. A standard deviation increase in frost shocks causes a 17 percentage point higher probability of crop failure, an increase of about 0.8 standard deviations. The result is robust to the inclusion of fixed effects for counties as well as county-specific linear trends. Column 2 provides evidence that the distinction between growing and non-growing seasons is crucial, as shocks that occur in the non-growing season have a near-zero and insignificant effect on harvests. If emigration was indeed caused by poor agricultural yields, we should expect to find the same pattern when emigration is the dependent variable. Columns 3 and 4 re-estimate the specifications in the first two columns using the full 0–6 scale of harvest grades, with results displaying the same pattern.

Path dependency in migration patterns has been well established in the migration literature.[36] This would hold in our setting if emigration in the first wave positively affected later emigration flows. Figure 6 gives a first indication that this might be the case. Panel A plots the spatial distribution of emigration rates during the first wave of emigration 1867–1879, while Panel B displays emigration in the whole 1867–1920 period. Comparing the raw data between the two maps reveals a substantial correlation in the propensity to emigrate over time. This is consistent with the fact that up 50 percent of emigrants traveled on pre-paid tickets sent by network members in the US (Runblom and Norman, 1976; Beijbom, 1995). Figure 7 also displays the relationship between early and later emigration in a scatter plot, also displaying a strong positive correlation.

Table 4 estimates how emigration over the full sample period is related to growing season frost shocks 1864–1867, log distance to the nearest emigration port and our instrument: the interaction between the two. The results in Column 1 are in line with our expectations — over the 1867–1920 period, municipalities further away from a port emigrate less in response to an additional frost shock. As distance and shock variables are demeaned, the estimates show that municipalities one standard deviation closer to ports emigrate by an additional 6.5 percent given a frost shock, while the effect at the mean distance is weakly positive but insignificant.[37] This result is robust to the inclusion of pre-determined control variables in

(Pettersson-Lidbom, 2009)

[36] See e.g. Bryan et al. (2014); Giulietti et al. (2014); McKenzie and Rapoport (2007); Munshi (2003).

[37] This is consistent with the theory that individuals take the internal migration cost into account in their decision to emigrate. E.g. Morten and Oliveira (2014) find that individuals with a shorter road distance to

Column 2.

If the distance to an emigration port proxies for the market access of a municipality, a potential concern could be that the instrument is associated with the differential impact of experiencing shocks in more isolated areas. This could lead to violations of the exclusion restriction in instrumental variables estimations below. To control for this possibility, we include our two measures of market access, the log distance to nearest town and the log distance to nearest major trade port, interacted with the shocks in Column 3. The coefficient on the instrument is not sensitive to this inclusion. The interaction terms themselves are also not significantly different from zero. The instrument is therefore only related the travel cost of emigration in using the distance to the nearest emigration port.

To provide support for the claim that frost shocks affect emigration through their impact on the agricultural sector, Column 4 additionally includes non-growing season frost shocks and their interaction with port distance.[38] The coefficient of the interaction term is substantially smaller and statistically indistinguishable from zero, thus mirroring the null effect found for agricultural outcomes in Table 3. The variation picked up by the growing season shocks therefore identifies economically meaningful events and not spurious correlations with underlying variables at the municipality, as captured by the distance to emigration ports.

Figure 8 displays the first-stage relationship non-parametrically. In Panel A, residuals of log emigration 1867–1920 and the instrument are plotted after controlling for the full set of covariates. Municipalities are collected in 50 groups of equal size, with dots representing the mean value in each group. The figure shows that across the whole range of the instrument, observations are clustered near the regression line.[39] The even distribution of group means indicates that there is compliance with the instrument at all values and indicates that the linear specification is an appropriate model. In Panel B, we instead use non-growing season frost shocks to form the placebo instrument. As expected, the figure shows that emigration has no apparent relationship with the placebo instrument, whether linear or non-linear.

Having established the importance of the initial frost shocks for emigration over the whole mass migration period, we next investigate two different ways in which early emigration affected future mobility and migration patterns. First, we divide the data into first and later waves and estimate the elasticity of later emigration with respect to early migration. Panel A of Table 5 estimates the effect of the instrument on first-wave emigration, 1867–1879. The results in Columns 1 to 3 indicate the same pattern as that found in Table 4, namely that locations who experienced frost shocks further away from a port were less likely to emigrate. In Panel B of Table 5, we use the relationship in Panel A as the first

the new city of Brasilia were more likely to migrate and take advantage of the comparatively high wages offered there.

[38] Non-growing season frost shocks over the period are defined analogously to growing season frost shocks.

[39] The slope of the regression line corresponds to the estimate in Column 4 of Table 4.

stage for estimating the causal effect of early emigration on later waves. The coefficients in Columns 1 to 3 show that there is a strong, causal pattern of path dependency, with an intertemporal elasticity of emigration near unity. These results confirm the canonical finding in the migration literature of strong path dependence in migration patterns (see e.g. Massey et al., 1993; Hatton and Williamson, 2002; Bryan et al., 2014; Giulietti et al., 2014; McKenzie and Rapoport, 2007; Munshi, 2003). Interestingly, the IV coefficients are greater in magnitude than their OLS counterparts.[40] This may be due to measurement error in emigration levels, since unregistered emigration was more common before 1884, when a new law made it harder to emigrate without proper documentation. The larger coefficients may also reflect the estimation of a different parameter between OLS and IV, if the instrument causes different types of individuals to emigrate.[41]

Second, we test if municipalities with more early emigration were differentially able to take advantage of favorable economic conditions in the United States. In Table 6, we use panel variation in emigration 1880–1920 and interact the instrument with a measure of the relative prosperity of the United States compared to Sweden. Exploiting panel variation, we are able to include municipality and year fixed effects. Column 1 shows that locations with more early emigration (low values of the instrument) are more likely to emigrate when the difference between US and Swedish GDP is larger. Columns 2 and 3 include additional controls for linear trends in the three major regions of Sweden as well as for a number of baseline municipal characteristics. The estimates remain strongly significant and show the same effect. Table 6 then highlights a different channel through which early emigration lead to more future labor mobility. The results suggest that overseas networks, as proxied by early emigration from the municipality, provided an option to emigrate and that this option was specifically exercised when the conditions were most favorable to do so. It also suggests that living standards were an important concern for potential emigrants, which would have to be taken into account by local elites.

To get a fuller picture of the relationship between emigration, frost shocks and within-country travel costs, Appendix Table A.2 presents additional estimates using the panel variation in emigration and frost shocks. Again, with municipality and year fixed effects we cancel out any potential biases related to fixed municipality characteristics that could potentially influence the cross-sectional relationship that we observe. The results follow the same pattern as in the cross-section models, with yearly growing season frost shocks leading to more emigration and a smaller marginal effect of shocks as the distance to an emigration port

[40] This is similar to estimates in McKenzie and Rapoport (2007).

[41] For example, liquidity constrained individuals should be more likely to emigrate as a response to the reduced migration cost of having a relative already in the US. If our instrument causes a higher fraction of poor people to emigrate than would otherwise have been the case, the incidence of chain migration could also be higher.

increases. This holds true only during the first wave of emigration, however. For the later period of emigration, 1880 to 1920, neither frost shocks nor the interaction with distance to emigration port matter. The importance of both frost shocks and port distance hence diminishes over time, perhaps as infrastructure improves and the economy shifts away from agriculture towards manufacturing. Both variables that compose our instrument can thus be thought of as only capturing variation that was relevant during the first wave of mass emigration.[42]

6 Emigration and citizens' demand for political change

This section estimates the effect of emigration on citizens' demand for political change across Swedish municipalities. The main variable of interest is membership in the labor movement, given that unions and the Social Democratic Party were the strongest proponents of political change during our period of study and were directly involved in conflicts with economic and political elites at local and national levels.

Labor movement membership The Social Democratic Party was founded in 1889. In the preceding decade, modern labor unions became more widespread, ultimately leading to the formation of the Swedish Confederation of Trade Unions in 1898. In Figure 9, we trace out the impact of emigration on municipal labor movement membership rates starting in this period and ranging until 1920. The figure displays IV coefficients from separate regressions in five-year intervals 1890–1920, including the full set of controls, with bars representing 95 percent confidence intervals. In the earliest years of the labor movement, 1890 to 1895, the IV estimates are insignificant and close to zero, albeit with a positive sign. Starting in 1900, however, emigration has a clear positive and statistically significant effect on labor organization rates. The effect sizes show an increasing trend, which mirrors the general positive trend in membership rates in the period. The figure provides the first evidence of a positive causal effect of emigration on local labor organization.

To get an aggregate picture of the relationship, Table 7 reports full regression specifications using the average labor movement membership rate between 1900 and 1920 as the dependent variable. Panel A shows first-stage and reduced-form estimates, while Panel B displays OLS and IV results. The first-stage results are similar to those explored in the previous section, indicating a negative relationship between the instrument and emigration in the period 1867–1900. In the two specifications in Columns 4 and 5 of Panel B, the estimated IV coefficients are strongly significant and stable at approximately 0.022, including

[42] In Section 9, we test the interaction between the distance to emigration ports and frost shocks occurring during all four-year periods other than 1864-1867.

when we control for pre-determined municipal characteristics. Column 6 includes the two market access interactions, using distance to the nearest trade port and town to control for potential violations of the exclusion restriction. The point estimate is robust to this inclusion and remains significant at the 1 percent level, indicating that frost shocks did not have any important differential effects between locations that were more or less connected to markets.

The point estimates are large. The preferred estimate in Column 6 suggests that a municipality which doubles its emigration over a 30-year period increases local labor movement membership by 2.3 percentage points. The effect size corresponds to moving a municipality from the mean to the 90^{th} percentile of the distribution of membership rates, or approximately 0.6 standard deviations. The IV coefficients are also just over twice as large as the corresponding OLS estimates. The difference implies a downward bias in OLS and that, if anything, OLS estimates provide a lower bound on the effect of emigration. A possible reason for this is that emigration was more common in regions that were also less likely to develop labor organizations, perhaps where landlords and employers were particularly powerful. This would be consistent with bad institutions acting as a push factor for emigrants. Measurement error in emigration may in addition be contributing to the difference in estimates.

These results provide evidence of a strong positive effect of emigration on membership the Swedish labor movement. The fact that the effect is present in its initial growth phase, during which it established itself as an important political player, provides a novel explanation for the well-known strong position of labor unions in Sweden.[43] These findings thus contrast with the hypothesis that vocal political dissidents would emigrate and decrease the level of activism in origin communities. It instead lends support to the hypothesis that emigration increased the pool of activists over time, possibly by acting as an insurance device and lowering the cost to citizens of organizing under the risk of repression by elites. Another, complementary, interpretation is that higher labor mobility increased the responsiveness of elites to citizens' demands, which raised the incentives for collective action and by consequence also organization rates.

To verify the robustness of our results, we also graphically display nonparametric estimates of the first stage and reduced-form relationships. Figure 10 plots the instrument against emigration 1867-1900 (the first stage) and labor movement membership 1900-1920 (the reduced-form). All variables are residualized using the full set of covariates. We see that both outcomes are negatively correlated with the instrument across the entire range of its values. Taken together, these results imply a positive relationship between emigration and labor movement membership, summarizing the main result of this section.

To further probe the effects on demand for political change, we study the effect of em-

[43] In 2000, Sweden had the second highest trade union density among OECD countries, behind Iceland.

igration on a direct, costly action directed towards employers. In response to a downturn in the business cycle in 1909, the Swedish Employers Association sought to lower workers wages. Anticipating opposition by labor organizations, it enacted a lockout of thousands of workers in order to force acceptance from the unions. The Swedish Confederation of Trade Unions instead responded by calling a general strike, affecting 300,000 laborers who halted work for three months. Using data on strike participation by municipality, we estimate the effect of emigration on mobilization of workers in Table 8. If our estimated effect on labor movement size indeed captures a greater ability to organize and mobilize citizens, we should expect high-emigration municipalities to display greater participation in the strike. This is confirmed by the IV results in Columns 2 and 3, which show a positive and significant effect. Similar to the case with labor movement membership, the estimated effect is large and indicates that a doubling of emigration increases emigration by 0.75 standard deviations. Membership in the labor movement was not only ceremonial then, but also resulted in effective collective action.

Separating strikers by union membership, we can define the the share of *unionized* strikers as more direct indicator for the extent to which the labor movement was the mechanism behind strike participation. This variable is constructed to equal zero for locations with no strikers, while it takes on negative values where nonunionized strikers were more common and positive values where union members were a larger fraction of strikers. As a result, a statistically significant estimate indicates that emigration causes more strike participation, while the sign of the coefficient shows which group that was most common. Column 4 of Table 8 indicates that emigration indeed causes a greater share of union members among strikers. Approximately 9 percent of municipalities that participated in the strike had more nonunionized than unionized strikers. As the so-called "striking weapon" was the most common tool available for the political and economic protest, this finding suggests that emigration developed a stronger bargaining position of citizens, through its effect on the labor movement.[44]

The labor unions that we observe in our data were almost exclusively organized in non-agricultural sectors. If emigration was concentrated among agricultural workers, who were more directly affected by the shocks used in constructing the instrument, a worry is that the effects on labor movement membership and strike participation in Tables 7 and 8 could potentially be a mechanical result from agriculture-skewed emigration. To test for this possibility, we use employment data from the 1910 census and rerun regressions for the effect of emigration on labor movement and strike participation expressed *per industrial*

[44]While the 1909 general strike was not considered a victory for the labor movement, strikes often resulted in favorable outcomes for workers. Summary evidence on 748 strikes 1863-1902 found that strikes resulted in concessions to workers' demands in 47 percent of the cases, while 32 percent of the cases ended in a compromise and only 20 percent sided with the employers (Kommerskollegii Arbetsstatistik E:1, 1909).

worker rather than per capita. This specification will net out any changes in the sectoral composition of employment. Columns 5 and 6 of Table 8 show the results of this test. Both variables are positive and significant, indicating that the main conclusions are robust to this variation. Hence, holding the number of industrial workers constant, labor organization in 1910 as well as strike participation in 1909 were still more intense in high-emigration areas.

Electoral effects The Social Democratic Party had strong ties with labor unions, and the central Confederation of Trade Unions in particular, each side making up one leg of the Swedish labor movement. The greater local membership of labor unions made them interesting for the Social Democrats, who saw a way of expanding the local penetration of socialist ideas. Unions indeed participated in election campaigns for the Social Democrats and a large fraction of voters for the left are thought to have come from labor union members (Westerståhl, 1945). Having established that emigration increased labor organization and striking, we therefore proceed to test if the relationship also extends to electoral mobilization. For this purpose, we look at turnout rates and support for left-wing parties in national elections between 1911 and 1921. This period ranges from the first election with mandatory party affiliations to the first election with universal suffrage.[45]

Panel A of Figure 11 displays the IV coefficients of emigration on the vote share of the Social Democrats across these elections. Emigration did indeed lead to significantly greater support for the Social Democrats in 1911 as well as in the two elections in 1914. After 1917, the coefficients fall, however, with confidence intervals including null effects. A reason for the relative decline in support for the Social Democratic Party can be found in Panel B of Figure 11. Here, we aggregate vote shares for the Social Democrats and the Socialist Party, which was formed in 1917 as an off-shoot of the former. The first three estimates are identical, but starting in 1917 – the first election in which the Socialists were active – the coefficients are larger and remain significantly different from zero at the 5 or 10-percent level through 1921. High-emigration municipalities thus have higher support for the more leftist party as well, which was formed in response to a demand for a more radical policy. Rather than displaying a regression to the mean, estimates suggest that municipalities with more emigration are associated with more radical ideas.

Aggregating the two left-wing parties, Table 9 reports regression results for the effect of emigration on the average vote share of the Social Democratic and Socialist parties between 1911 and 1921. The IV estimates in Columns 2 to 4 range from 0.122 to 0.138, implying that an increase in emigration by 10 percent increased the vote share of the left by approximately

[45]In addition, data on municipal elections would have been informative, because the weighted voting scheme present in local elections until 1919 would have given an indication of how elites' preferences were affected by emigration. It would also have been directly relevant for municipal policy. Unfortunately, such data are unavailable to us.

1.2–1.4 percentage points. Similar to the case of labor movement membership, the IV estimates are larger than OLS. Even after adding controls for baseline characteristics, the OLS estimates are smaller, implying the presence of unobserved factors that yield a downward bias in OLS.

Until 1921, voting eligibility was reserved for men who had payed their taxes, who were not in poverty care or bankruptcy and who had performed their military service. These restrictions disenfranchised one fifth of otherwise voting eligible men (Grenholm et al., 1985). Nevertheless, even during a time when only relatively well-off men could vote, there is a shift in party preferences toward left-wing parties. It is difficult to distinguish how much of this change is due to an increased popularity among working-class voters and how much is due to elites shifting their voting towards parties that would be more popular among average citizens. Given that one-man, one-vote was used in national elections, however, the effects are unlikely to be driven mostly by the voting of elites.

It is also relevant to note that these results take into account any potential changes in the voter turnout due to emigration, as the left-wing vote share is computed using the total number of votes as the denominator. The increase in the vote share of left-wing parties is therefore not simply explained by an increased mobilization of poor voters, but is due to a differential voting behavior among voters.

Finally, Columns 5 to 8 of Table 9 display regression results for the effect of emigration on average voter turnout during the period. We find positive effects on turnout, with the IV results in Panel B ranging from 0.068 to 0.075 in the preferred specification with market access interactions. The effect sizes indicate that a doubling of emigration increases voter turnout by approximately 7 percentage points, from an average of 60 percent during the period. This result suggests a complementary role of labor organization and voting, in line with the goal of the Social Democratic Party of using local organizations to mobilize citizens for larger, national-wide political change.

Selection Selection effects are a first-order concern when studying migration. If those who chose to emigrate are very different than those who stay behind, migration may change the composition of the origin community population substantially over time. This could itself have direct, mechanical effects on our outcomes of interest. It thus constitutes a competing explanation for our results, one which does not imply any changes in behavior.

Data on the share of eligible voter allow us to test for a certain type of selection effect that could explain our results. Given that voting eligibility was based on economic status and gender, it can serve as an indicator of changes to the composition of the population that has direct bearing on electoral outcomes. However, Columns 9 to 12 of Table 9 indicates that emigration had no significant effect on the share of eligible voters. Moreover, the sign of

the estimated changes from positive without additional control variables, to negative when we include controls. As a result, the effects on voting patterns are thus cannot be explained by this form of selection.

Selection may also be active along other dimensions than voting eligibility. Using 1910 census data, we test for a wider range of demographic differences across high and low emigration municipalities. Table 10 tests the effect of emigration on a number of indicators of demographic change. Column 1 shows that there is no differential in-migration from other regions of Sweden in municipalities with more emigration, ruling out e.g. welfare migration and selective in-migration of more leftist individuals. Columns 2 and 3 test for evidence of a fertility transition related to emigration. However, we find no effect on the average family size, nor on the incidence of unmarried adults. Finally, Column 4 shows that the ratio of women to men is not affected by emigration. Estimated coefficients are all relatively small compared to their mean. Hence, low power should not be the reason for failing to reject the null hypothesis. Overall, these results are consistent with the view represented in Runblom and Norman (1976) that the mass migration became "general and popular", and hence that individuals who chose to emigrate to the US were not substantially different to the general population.

Another possibility is that emigrants were ideologically selected. If more right-leaning individuals chose to emigrate, for example because of the pull factors of more freedom or because of a more risk-taking or entrepreneurial preferences, the pool of voters would mechanically change in favor of the left. To deal with this concern, we perform a simple bounding exercise. First, we count the number of emigrants 1867–1910 and assume that they would have voted in all elections 1911–1921. We then consider three scenarios for the ideological selection of emigrants. Appendix Table A.3 displays the sensitivity of our estimates for the left's vote share when assuming that 75 percent, 90 percent or 100 percent of emigrants would have voted for a non-left party if they had stayed in Sweden. As to be expected, point estimates becomes successively smaller as we assume a more skewed ideological selection, reducing the baseline result by up to about half. Nevertheless, all results remain sizable and statistically significant, indicating that such selection cannot explain the entire effect of emigration on left-wing voting.

7 Emigration and elites' supply of political change

The results in the previous section show that emigration increased the political organization, mobilization and, arguably, the bargaining power of citizens during the Age of Mass Migration. In this section, we turn to analyzing whether these changes were also reflected in the supply of political change, by looking at local policy and political institutions.

Welfare expenditures We use welfare expenditures as a measure of redistributive, pro-citizen actions taken by local governments. The choice of expenditures can also be seen as an equilibrium outcome of bargaining between elites, who hold political power, and citizens. We study per capita expenditures on welfare in 1918, one year before democratization, and in 1919, immediately afterwards. In 1918, municipal voting was restricted by wealth, income and property ownership. Votes were also weighted by a factor of up to 40 in favor of richer voters (Nilsson, 2008).[46] As a result, formal authority over spending levels was heavily biased in favor of economic elites in 1918. Changes in policy at this time are thus reflective of their choices rather than those of common citizens. This can be quantified by comparing average welfare spending before and after democracy: in 1918 it was 2.42 SEK per capita while it rose by 13 percent to 2.74 SEK in 1919 as ordinary citizens could vote.[47] Nevertheless, it is possible that citizens could wield some influence on the welfare spending decisions of elites before democracy.

Table 11 displays our results. Column 2 shows that emigration leads to significantly higher per capita expenditure in 1918, one year before democracy was introduced. The estimate remains stable as we include baseline controls (Column 3) as well as the market access interactions to control for potential violations of the exclusion restriction (Column 4).

How could welfare expenditures rise even before ordinary citizens could vote? A potential mechanism is that Social Democrats and labor representatives were allowed positions in municipal governance. By 1917, several municipalities had representatives from the labor movement present in local administration (Östberg, 1995). Social Democrats could also be voted into formal political power by being given a place on the election lists of other, more popular parties which sought to increase their representativeness (Lundkvist, 1977).

The effects on welfare spending remain in 1919, as voting rights were extended on an equal basis. The estimates are higher than in 1918, and are also robust to the inclusion of controls in Columns 7 and 8. The estimate in Column 8 indicates that a doubling of emigration leads to approximately 1.1 SEK higher expenditures per capita in 1919, an increase of 36 percent over the mean. Overall, both before and after citizens had the formal power of affecting welfare policy in municipalities, emigration thus lead to higher levels of redistribution.

Form of democracy in local governments During this period, rural local governments could adopt two different institutions for decision making, direct or representative democracy. In direct democratic municipalities, public town meetings would be held at least three times a year to decide on economic matters. Deliberations were open, as well as many votes. By contrast, in municipalities of the representative type, eligible citizens voted for their party

[46] In 1905, 1 percent of the rural population held as many votes as the remaining 99 percent (Nilsson, 2008). The cap on votes was 1000 at that time, however, rather than 40.

[47] Expenditure data are deflated by CPI.

of choice in closed elections. Starting in 1919, there was an assignment rule dictating that municipalities with more than 1500 inhabitants adopt the representative form of government, whereas those below the threshold were free to choose between the two. In practice, however, only a small fraction of municipalities chose the representative form voluntarily. Hinnerich and Pettersson-Lidbom (2014) study the effects of these institutions in detail. They find that direct democracies implement substantially lower levels of welfare spending per capita, potentially due to direct democracies being more easily captured by elites. This is partly seen by the low attendance rate at town meetings, 12 percent, whereas voter turnout in national elections which was rutinely above 50 percent. The choice of institution was then to a large extent a choice about its inclusiveness, the relative power of elites and the amount of redistribution. This may, in turn, explain the low rate of voluntary transitions from direct democracy, which was the default, to representative democracy.

We use data on the local form of democracy to test for the effect of emigration on institutional change. This is done by coding a dummy variable taking the value of 1 if the municipality was a representative democracy by 1919, 1920 or 1938 (the final year of our data), and had a population of 1500 or less in the preceding year. The last condition is included to take into account only voluntary transitions from direct to representative government. We take this measure to be an indicator of the inclusivenes of local political institutions. In addition, we include dummies for a municipality having ever crossed the population threshold in the preceding years.

Panel B of Table 12 shows that high emigration municipalities were indeed more likely to adopt the more inclusive form of democracy in their local governments. The effects are statistically significant in 1919 and 1920, with coefficients implying between a 4.7 and 5.8 percentage point increase in the likelihood of a representative democracy from a doubling of emigration. In 1938, when a larger share of municipalities had transitioned voluntarily, the effect is larger. Transitions were hence more common in the longer run, possibly reflecting that organized citizens gained more influence over time.

How should we interpret these effects? An important institutional feature was that municipalities that switched to representative democracy were required to keep the institution for at least five years. Reversions back to direct democracy were rare, however. An interesting question for interpreting these results is to what extent these institutional changes represent elites' concessions to citizens, versus citizens' own enforcement of their preferences. While we only observe transitions between political regimes after the introduction of one man, one vote, it is not necessarily the case that ordinary citizens held complete *de facto* political power in rural municipalities. Some elites were able to maintain important positions of power even after 1919. Moreover, electoral competition was generally limited, with 30 percent of the elections only having one party in 1919 (see Hinnerich and Pettersson-Lidbom, 2014, and

references therein). While the preferences of citizens should more directly affect outcomes after the introduciton of local democracy in 1919, our results may therefore still reflect the outcome of bargaining between elites and citizens. This may especially be the case as the default institution was direct democracy, which had been restricted to wealthy citizens for decades. In this regard, observing that emigration leads to the adoption of long-lasting and more inclusive institutions may reflect a strategy of elites to commit to more pro-citizen policies by reforming the basic rules of the game, as suggested by Acemoglu and Robinson (2000, 2006).

8 Alternative mechanisms

So far, we have emphasized the role of labor mobility in promoting the organization of labor, in turn leading to political change. Above, robustness checks have shown that labor movement membership was not solely driven by changes in the employment structure, nor by changes to the electorate as measured by the share of voting eligible citizens. We also found small and insignificant effects on demographic indicators, as well as rates of in-migration to municipalities. A bounding exercise moreover suggested that ideologically skewed emigration of right-leaning individuals could not account for the estimated effects on left-wing vote shares. This section considers additional alternative explanations for our results.

Existing studies linking migration and political outcomes have emphasized the potential of host country attitudes being transmitted to origin countries, and thereby potentially affecting political outcomes. This raises the question of whether American attitudes could have inspired the Swedish labor movement, whether it be via return migration or information transmission through social networks. We provide an indirect test of the hypothesis of such a cultural transmission effect by estimating the impact of emigration on two other voluntary associations that we observe in the data: non-state free churches and temperance lodges. Both types of organizations had strong influences in the United States. Methodists were common among the free churches and the temperance movement did largely consist of Swedish chapters of an American organization, the International Order of Good Templars (IOGT).

Table 13 displays our results for per capita membership in both types of organizations. Similar to the specification for the labor movement, we consider the average membership between 1900 and 1920. If there was transmission of information or attitudes to Swedes through their overseas networks, one would expect to see increased participation in these types of organizations. The results in Table 13 show no positive effect, however. Free churches do not see any significant change in membership with more emigration and temperance lodges experience a negative effect. These results do not rule out that the labor movement was in

some way influenced by the United States, but they nevertheless provide evidence for a cultural transmission effect not being a significant factor in explaining our findings.

Our identification strategy isolates exogenous variation in the amount of emigration, but those who choose to emigrate may nevertheless be selected. Such self-selection may change the composition of municipalities and affect our outcomes through a separate channel. Research on Norwegian migrants during the Age of Mass Migration has found that migrant self-selection in terms of earnings potential was negative from urban areas, but ambiguous from rural areas (Abramitzky et al., 2012). Given that most of the variation used in this study is from rural areas, as well as the fact that the majority of emigrants were rural, our results should not be substantially affected by this form of self-selection. To verify this, we replicate the effect of emigration on labor movement participation only using the rural sample. Table 14 shows that our findings on labor movement membership, striking and voting are robust to the exclusion of urban areas.[48] All point estimates remain highly significant, and point estimates are roughly similar, with one being higher and three being lower than the main estimates. However, while not statistically significant, the overall somewhat smaller estimates within the rural sample could be an indication that emigration of the poor had a more positive impact on the labor movement.

9 Placebo and robustness tests

The available time-series data suggest a natural placebo test for our identification assumption. Since we only rely on frost shocks occurring in the 1864–1867 period, we run placebo reduced-form regressions for the main outcomes (including emigration) using frost shocks during all other four-year periods from 1859 to 1900, interacted with port distance.[49] As the variation in frost shocks is random, placebo coefficients are expected to be distributed around zero. However, a potential worry with this prediction is that frost shocks in other periods could also have affected emigration and yield large point estimates. Nevertheless, we believe placebo treatments provide a meaningful test, for two reasons. First, no weather events other than those associated with the 1860s famine have been identified by historians as causes of emigration. Second, our instrument interacts shocks with the distance to emigration ports, which ceased to be an important barrier to emigration after the first wave of emigration (see last paragraph of Section 5). Hence, even if later shocks affected emigration, they should not be doing so through the interaction with port distance.

We should therefore expect coefficients associated with the treatment period to be in the

[48]Note that regressions on welfare expenditures and representative democracy are already restricted to the rural sample.

[49]Shocks 1864–1869 are excluded to avoid the treatment period.

extremes of the distribution. To make frost shocks comparable across periods with very few or very many shocks, and avoid the influence of outliers, they are categorized in quintiles of the shock distribution over the period. Appendix Figure A.1 displays probability density functions of all placebo point estimates. The black bars represent the treatment reduced-form effect, while white bars represent placebo periods. As expected, placebo estimates are scattered across the range of values while the treatment coefficients are consistently at the ends of the distribution for all outcomes.

In Table 4, we found that constructing the instrument using *non-growing season* frost shocks could not predict emigration. This was the expected result, given that frost shocks have no effect on agricultural outcomes outside of the growing season. Appendix Table A.4 further shows that such shocks do not have any reduced form effects on our outcomes either. The estimates in all columns are insignificant and close to zero. Hence, the main effects that we find are do not appear to be driven by unobserved fixed characteristics of municipalities.

Different cutoffs for defining frost shocks are examined in Appendix Table A.5. Panel A displays the reduced-form estimates of our main outcome variables using shocks defined at the baseline of 1 standard deviation, while Panels B and C display estimates from letting shocks count at 0.75 or 1.25 standard deviations. Finally, in Panel D, we define growing season using months with a mean temperature of above 5 degrees Celsuis, as this is the upper bound for counting a month as being in the growing season following the recommendations of the Swedish Meterological and Hydrological Institute. The baseline case uses 3 degrees. The signs, magnitudes and statistical significance of these results are similar to the main results.

We next evaluate the robustness of our analysis to large (absolute) values of our key variables. In particular, we want to control for the possibility of certain locations that are very distant from ports driving our results.[50] To do so, we censor variables at the 5^{th} and 95^{th} percentiles, assigning observations outside of that interval the variable value at the nearest bound. This compresses the range of values that variables take on and reduces the potential for a small number of observations with extreme values to affect estimates. We also display the results after tightening variable distributions further, by censoring at the 10^{th} and 90^{th} percentiles. Panels A and B of Appendix Table A.6 do this for two variables: growing-season frost shocks 1864–1867 and distance to the nearest emigration port. The resulting variables are then used to redefine the instrument, i.e. the interaction between shocks and port distance. All results are robust to this change. Panels C and D then extend this procedure to *all* non-binary variables that are included in our models. Our results are robust to this modification as well.

[50]The inclusion of 24 county fixed effects in all our specifications should to a large extent already take this possibility into account.

To test for the robustness of our inference, Appendix Table A.7 provides estimates of the reduced-form regressions using two different types of standard errors. In Panel A, we cluster standard errors at the county level rather than at the weather station, as political organization and policy may be more correlated within counties, which are established political boundaries. Panels B and C instead estimate spatial-correlation robust standard errors which allow linearly declining correlations across municipalities of up to 100 or 200 kilometers, using the method of Conley (1999). This method has the advantage of not relying on a fixed number of clusters and allows residuals to be correlated within a given radius of each unit of observation. Panel D generates standard errors using the wild cluster-t bootstrap method, which may improve tests when there are few clusters (Cameron et al., 2008). The estimates in Panels A to D display the same pattern as our main regressions, with few changes to significance levels. The estimates on transitioning to representative democracy lose precision with the wild cluster-t bootstrap method but are nevertheless robust to oth levels of spatial dependence using Conley-type standard errors.

10 Are the effects persistent?

Political preferences have been found to exhibit path-dependence within individuals and to be shaped by pivotal events (Kaplan and Mukand, 2011; Madestam et al., 2011). Political preferences and behaviors are also known to be correlated between parents and children (Alford et al., 2005; Jennings et al., 2009). Institutions may also have long-lasting effects on individual beliefs and values (Nunn and Wantchekon, 2011). There is thus reason to believe that the effects found in previous sections may persist in the long run. Below, we conduct such tests for long-run persistence of the effect of emigration on labor movement participation and support for left-wing parties.

Figure 12 plots the IV coefficients from regressing labor movement participation between 1925 and 1945 on emigration. The point estimates are in line with the earlier results, but larger, albeit less precisely estimated for two years. At least until 1940, labor movement membership is significantly higher in locations with more emigration. Table 15 displays the full regression output for the average per capita membership rate 1921–1945. The results are significant and larger than during the 1900–1920 period, indicating that as the labor movement grew larger, the differences in participation widened further.

Next, we extend the analysis to contemporary elections to investigate the effect of emigration in the very long run. We use data on both national and municipal elections from the five most recent election rounds, 1998 to 2014, to estimate the persistence of emigration on left-wing voting. Table 16 displays reduced-form effects of the instrument as well as IV estimates using emigration from 1867 until 1945, after which emigration was uncommon.

Strikingly, the results show that the frost shocks occurring 1864–1867 have significant effects on voting up to the five latest Swedish elections. The results are stronger in municipal elections than at the national level, possibly due to issues at the national level having a stronger sway over voters as compared to tradition. The estimate in Column 6 of 0.074 is roughly half as large as the corresponding estimate for the 1911–1921 elections. The mean vote share of the left is also higher in the later period, i.e. 38 percent rather than 24 percent.

11 Discussion and conclusion

During the Age of Mass Migration, 30 million Europeans settled in the United States. Among them were more than one million Swedish citizens, making Sweden one of the major origin countries in per capita terms. This paper uses detailed Swedish data from the period 1860–1920 to shed light on a big question: can large-scale emigration lead to political development in undemocratic origin countries? Our results indicate that it may indeed be the case. Using an instrument based on travel costs and the severe agricultural shocks that sparked the initial wave of migration to the United States, we predict total emigration flows over 50 years. We show that emigration caused significantly higher rates of labor organization, strike participation, voter turnout and left-wing voting in the long run. The findings are consistent with the hypothesis that emigration possibilities provided insurance for potential labor activists, who faced repression from local elites for organizing. Since the labor movement had strong ties to the political left, our findings on turnout and political party preferences are likely driven by that mechanism.

Emigration also lead to real political change. Welfare expenditures per capita rose in high emigration municipalities, as did the likelihood of adopting more inclusive institutions by transitioning from direct to representative democracy. These results are consistent with the mechanism proposed by Acemoglu and Robinson (2000, 2006), in which elites implement institutional change in order to commit to better outcomes for citizens.

Overall, the Age of Mass Migration brought positive effects on citizens' bargaining power, support for redistribution and actual redistribution during a time when Sweden was still undemocratic. Migration arguably played a role in the country's transition to a full democracy in the early 20th century.

How externally valid are these results? We briefly discuss this question, in two parts. First, do our findings have any bearing on current emigration waves and record global refugee numbers? The mechanism that we propose, that greater outside options may encourage risky activism, is general and potentially applies to many other settings, including contemporary ones. However, the question of how responsive political elites will be to such activism is less straightforward. Agricultural and early industrial economies, such as Sweden in our

period of study, are heavily reliant on labor for production. Hence, the upper echolons of society necessarily depend on the ready supply of workers for their livelihood. This may explain the urgent political response of Swedish elites as emigration took on greater proportions. However, elites in many developing countries today are quite independent of labor for their income, relying instead on rents from natural resources to finance their regimes. Such external resource dependence may decrease the impact of emigration political institutions. Nevertheless, to the extent that activists are able to reach a significant mass, institutional change may occur as economies experiences critical junctures (Acemoglu and Robinson, 2006).

The second type of external validity concerns the extent to which our findings generalize to the Age of Mass Migration as a whole, during which several countries experienced high rates of emigration to the United States. At the same time, there was a considerable variation in sending rates, even among European countries. Below, we conduct a simple exercise to investigate the possibility that the positive relation that we find between emigration and labor movement participation in Sweden could apply more generally across countries.

Data on trade union density in 25 OECD countries are matched with the size of the population residing in the United States in 1920. We then make the (admittedly stark) assumptions that our estimate of the causal effect in Table 7 is generalizable across countries to predict trade union densities for OECD countries.[51] Figure 13 displays a scatter plot between actual and predicted trade union density in 2000, roughly 80 years after the United States closed its borders. Interestingly, there is a very clear positive relationship between the two variables, even though the estimated relationship naturally does not saying anything about whether the correlation reflects a causal effect. Countries with well-known diasporas in the United States, such as Ireland, Italy, Germany and Sweden, are also those countries that exhibit high levels of trade union density, while countries with less emigration, such as France and Spain, have lower unionization rates.

In Table 17, we show the regression output of emigration on trade union density in both 1960 and 2000. The relationship is stable across the two periods, and is robust to the inclusion of basic control variables, including GDP per capita, life expectancy, share of urban population and length of schooling. The results indicate that the simple model in Columns 1 and 4 with only emigration and initial population size can account for 43 to 52 percent of the variation in the sample.

Extrapolating from our main setting to a cross-country analysis with a small number of observations is certainly very speculative. Nevertheless, the robust correlation raises the possibility that the free immigration policy maintained by the United States in the 19th century and until World War I may have had significant unintended consequences for political

[51] Predicted values are *log(US population in 1920)* $\times$ *IV coefficient*.

development in the rest of the world.

References

Abramitzky, R., Boustan, L. P., and Eriksson, K. (2012). Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the Age of Mass Migration. *American Economic Review*, 102(5):1832–1856.

Acemoglu, D., Johnson, S., and Robinson, J. (2005). The rise of europe: Atlantic trade, institutional change, and economic growth. *The American Economic Review*, 95(3):546–579.

Acemoglu, D. and Robinson, J. A. (2000). Why did the west extend the franchise? Democracy, inequality, and growth in historical perspective. *Quarterly Journal of Economics*, pages 1167–1199.

Acemoglu, D. and Robinson, J. A. (2006). *Economic origins of dictatorship and democracy*. Cambridge University Press.

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

Alford, J. R., Funk, C. L., and Hibbing, J. R. (2005). Are political orientations genetically transmitted? *American Political Science Review*, 99(02):153–167.

Andrae, C. G. and Lundqvist, S. (1998). *Folkrörelsearkivet 1881-1950 (The Social Movement Archive)*. Uppsala universitet, Historiska institutionen, Svensk Nationell Datatjänst.

Barton, H. A. (1994). *A Folk Divided : Homeland Swedes and Swedish Americans, 1840-1940*. Southern Illinois University Press.

Batista, C. and Vicente, P. C. (2011). Do migrants improve governance at home? Evidence from a voting experiment. *World Bank Economic Review*, 25(1):77–104.

Beijbom, U. (1995). *Mot löftets land: Den svenska utvandringen*. LT.

Beine, M., Docquier, F., and Rapoport, H. (2001). Brain drain and economic growth: theory and evidence. *Journal of Development Economics*, 64(1):275–289.

Bengston, H. and Brook, M. (1999). *On the left in America: memoirs of the Scandinavian-American labor movement*. SIU Press.

Berglund, S. (1988). *Svenska Valdata 1911-1944*. Umeå universitet, Statsvetenskapliga institutionen, Svensk Nationell Datatjänst.

Besley, T., Persson, T., and Reynal-Querol, M. (2015). Resilient leaders and institutional reform: Theory and evidence. *Working paper*.

Bohlin, J. and Eurenius, A.-M. (2010). Why they moved—emigration from the Swedish countryside to the United States, 1881–1910. *Explorations in Economic History*, 47(4):533–551.

Boix, C. (2003). *Democracy and redistribution*. Cambridge University Press.

Bryan, G., Chowdhury, S., and Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica*, 82(5):1671–1748.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3):414–427.

Clemensson, P. (1996). Göteborgska källor till den stora utvandringen. *Dokumentet*, (3).

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

Dell, M., Jones, B. F., and Olken, B. A. (2014). What do we learn from the weather? The new climate–economy literature. *Journal of Economic Literature*, 52(3):740–798.

Dippel, C., Greif, A., and Trefler, D. (2015). The rents from trade and coercive institutions: Removing the sugar coating. *NBER Working paper No. 20958*.

Docquier, F., Lodigiani, E., Rapoport, H., and Schiff, M. (2014). Emigration and democracy. *FERDI Working Paper 90*.

Edvinsson, R. (2005). *Growth, Accumulation, Crisis : With New Macroeconomic Data for Sweden 1800-2000*. PhD thesis, Stockholm University, Department of Economic History.

Giuliano, P. and Spilimbergo, A. (2014). Growing up in a recession. *Review of Economic Studies*, 81(2):787–817.

Giulietti, C., Wahba, J., and Zenou, Y. (2014). Strong versus weak ties in migration. *IZA Discussion Paper No. 8089*.

Grenholm, G., Rosén, J., Carlsson, S., and Cornell, J. (1985). *Den svenska historien. 14, Från storstrejken 1909 till folkhemspolitik*. Bonnier, Stockholm, [ny tr.] edition.

Hansen, J. and Lebedeff, S. (1987). Global trends of measured surface air temperature. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 92(D11):13345–13372.

Harari, M. and La Ferrara, E. (2013). Conflict, climate and cells: A disaggregated analysis.

Hatton, T. J. (1995). A model of Scandinavian emigration, 1870–1913. *European Economic Review*, 39(3):557–564.

Hatton, T. J. and Williamson, J. G. (1993). *After the famine: Emigration from Ireland 1850-1913*. Cambridge Univ Press.

Hatton, T. J. and Williamson, J. G. (2002). What fundamentals drive world migration? Technical report, National Bureau of Economic Research.

Hellstenius, J. (1871). Skördarna i Sverige och deras verkningar. *Statistisk Tidskrift*, (29).

Hinnerich, B. T. and Pettersson-Lidbom, P. (2014). Democracy, redistribution, and political participation: Evidence from Sweden 1919–1938. *Econometrica*, 82(3):961–993.

Hirschman, A. O. (1970). *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*, volume 25. Harvard university press.

Hirschman, A. O. (1978). Exit, voice, and the state. *World Politics*, 31(01):90–107.

Hvidt, K. (1975). *Flight to America: The social background of 300,000 Danish emigrants*. Academic Press New York.

Jantunen, J. and Ruosteenoja, K. (2000). Weather conditions in northern Europe in the exceptionally cold spring season of the famine year 1867. *Geophysica*, 36(1-2):69–84.

Jennings, M. K., Stoker, L., and Bowers, J. (2009). Politics across generations: Family transmission reexamined. *Journal of Politics*, 71(03):782–799.

Kälvemark, A.-S. (1972). *Reaktionen mot utvandringen: emigrationsfrågan i svensk debatt och politik 1901-1904*. Acta Universitatis Upsaliensis.

Kaplan, E. and Mukand, S. (2011). The persistence of political partisanship: Evidence from 9/11. Technical report, Mimeo, University of Maryland.

Kommerskollegii (1910). *Redogörelse för Lockouterna och Storstrejken i Sverige år 1909 III*. Kungliga Kommerskollegii Afdelning för arbetsstatistik, Stockholm.

Lundkvist, S. (1977). Folkrörelserna i det svenska samhället 1850-1920.

Madestam, A., Yanagizawa-Drott, D., et al. (2011). Shaping the nation: The effect of Fourth of July on political preferences and behavior in the United States. *Preprint*.

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., and Taylor, J. E. (1993). Theories of international migration: A review and appraisal. *Population and development review*, pages 431–466.

McKenzie, D. and Rapoport, H. (2007). Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico. *Journal of Development Economics*, 84(1):1–24.

Morten, M. and Oliveira, J. (2014). Migration, roads and labor market integration: Evidence from a planned capital city. *Working paper*.

Morton, R. B. (1991). Groups in rational turnout models. *American Journal of Political Science*, pages 758–776.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. *Quarterly Journal of Economics*, pages 549–599.

Myrdal, G. (1944). *An American Dilemma, Volume 2: The Negro Problem and Modern Democracy*, volume 2. Transaction Publishers.

Nilsson, G. B. (2008). Folkval och fyrkval 1863-1909. två rösträttssystem i teori och praktik. *Scandia: Tidskrift för historisk forskning*, 30(1).

Nordahl, P. (1994). Weaving the ethnic fabric: social networks among Swedish-American radicals in Chicago 1890-1940.

Nunn, N. and Wantchekon, L. (2011). The slave trade and the origins of mistrust in Africa. *American Economic Review*, 101:3221–3252.

Omar Mahmoud, T., Rapoport, H., Steinmayr, A., and Trebesch, C. (2015). The effect of labor migration on the diffusion of democracy: Evidence from a former Soviet Republic. *Working Paper*.

Östberg, K. (1995). Att institutionalisera demokratin: Socialdemokratin och kommunalpolitiken under mellankrigstiden. *Arkiv för studier i arbetarrörelsens historia*, 1995 (Nr 63/64), s. 1-32.

Persson, T. and Tabellini, G. (2009). Democratic capital: the nexus of political and economic change. *American Economic Journal: Macroeconomics*, 1(2):88–126.

Pettersson-Lidbom, P. (2009). Midwives and maternal mortality: Evidence from a midwifery policy experiment in Sweden in the 19th century. *Mimeo*.

Pfutze, T. (2012). Does migration promote democratization? Evidence from the Mexican transition. *Journal of Comparative Economics*, 40(2):159–175.

Puga, D. and Trefler, D. (2014). International trade and institutional change: Medieval Venice's response to globalization. *Quarterly Journal of Economics*, 753:821.

Quigley, J. M. (1972). An economic model of Swedish emigration. *Quarterly Journal of Economics*, pages 111–126.

Rondahl, B. (1985). Massutvandringen från Ljusne 1906. *Oknytt*, 6(3-4):16–21.

Runblom, H. and Norman, H. (1976). *From Sweden to America: A history of the migration*. Acta Universitatis Upsaliensis.

Sarsons, H. (2015). Rainfall and conflict: A cautionary tale. *Journal of Development Economics*, 115:62–72.

SMHI (2013). Hungersnödåret 1867 – kallaste majmånaden vi känner. *Swedish Meterological and Hydrological Institute*.

Sánchez de la Sierra, R. (2015). On the origin of states: Stationary bandits and taxation in eastern Congo. *Working Paper*.

Snyder, R. L. and Melo-Abreu, J. P. (2005). *Frost protection: fundamentals, practice and economics. Volume 1*. FAO.

Spilimbergo, A. (2009). Democracy and foreign education. *American Economic Review*, pages 528–543.

Sundbärg, G., editor (1913). *Emigrationsutredningen: Betänkande*. Norstedt & Söner.

Taylor, A. M. and Williamson, J. G. (1997). Convergence in the Age of Mass Migration. *European Review of Economic History*, 1(1):27–63.

Tedebrand, L.-G. (1983). Strikes and political radicalism in Sweden and emigration to the United States. In Hoerder, D., editor, *American Labor and Immigration History, 1877-1920s: Recent European Research*. University of Illinois Press.

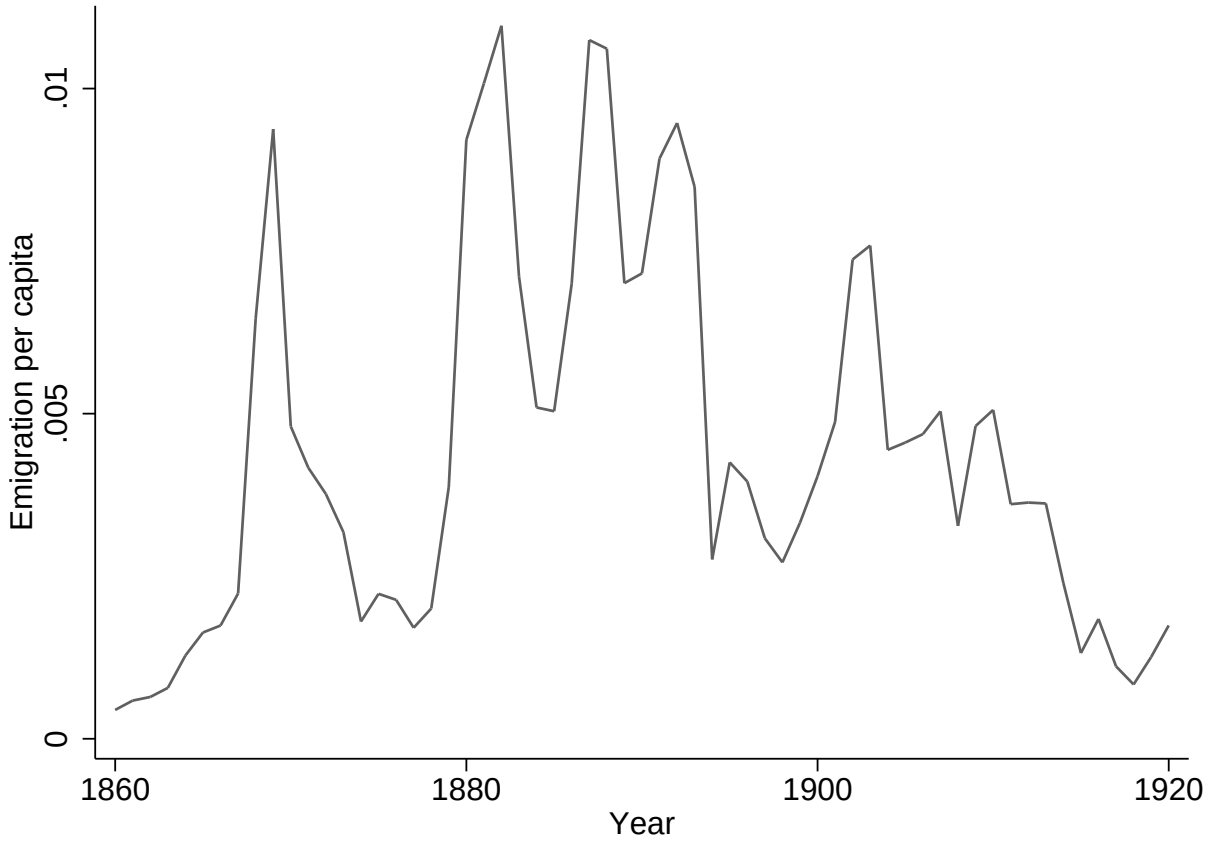
The Maddison Project (2013). *Available at http://www.ggdc.net/maddison/maddison-project/home.htm*.

Westerståhl, J. (1945). Svensk fackföreningsrörelse: organisationsproblem, verksamhetsformer, förhållande till staten.

White, G. F. and Haas, J. E. (1975). Assessment of research on natural hazards. Technical report, MIT Press.

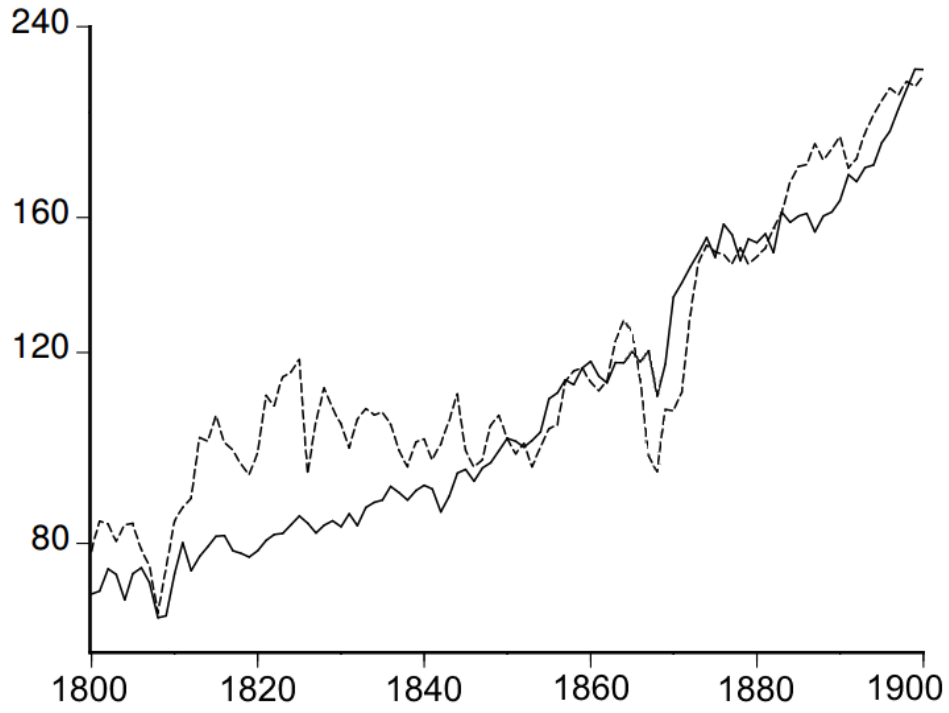
12 Graphs and Tables

Figure 1: Emigration flows 1860–1920



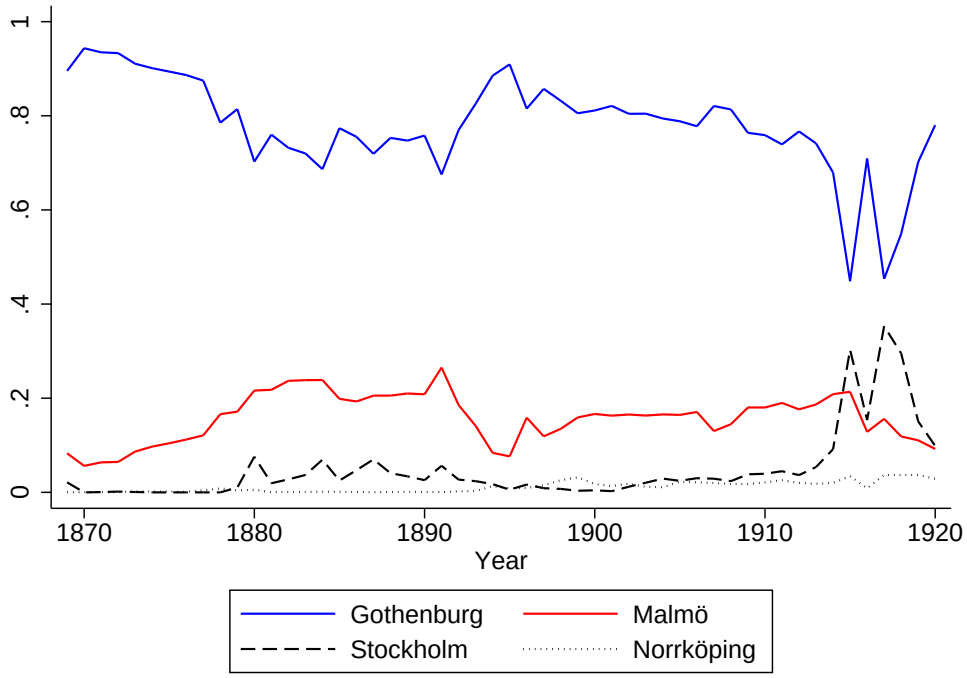
Notes: This figure displays aggregate emigration flows per year between 1860 and 1920. We label emigration during the 1867–1879 period *the first wave of mass emigration*. Later waves, during the 1880s and early 1900s, are also visible. Mass migration from Sweden ended in the 1920s, as the United States enacted immigration quotas.

FIGURE 2: GDP PER CAPITA AND REAL WAGES 1800–1900



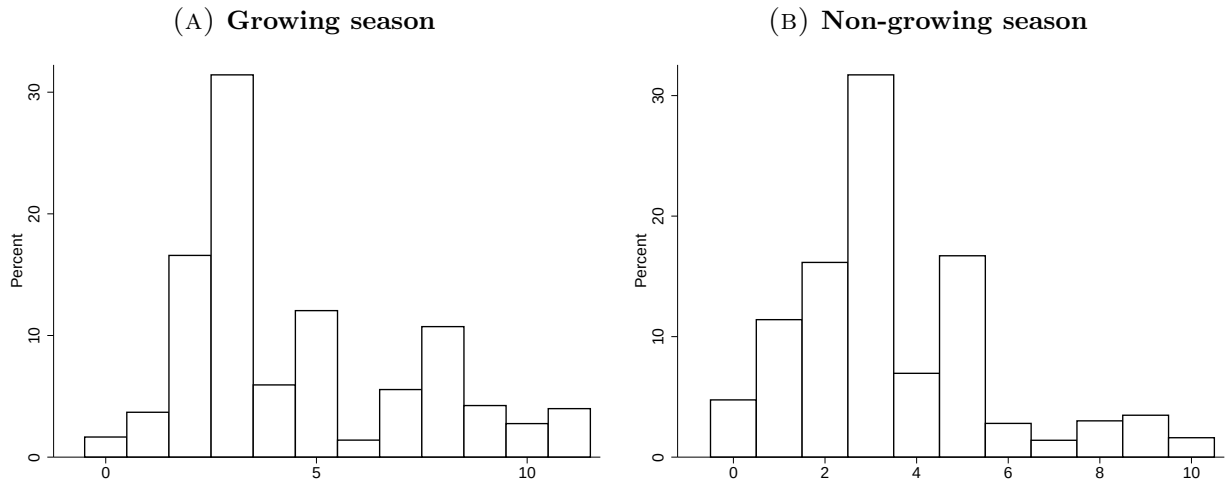
Notes: This figure displays the evolution GDP per capita (full line) and real wages (dotted line) in Sweden 1800–1900. The economic downturn during the famine years is visible around 1867. Index 1850=100.

FIGURE 3: MAIN EMIGRATION PORTS 1869–1920: GOTHENBURG AND MALMÖ



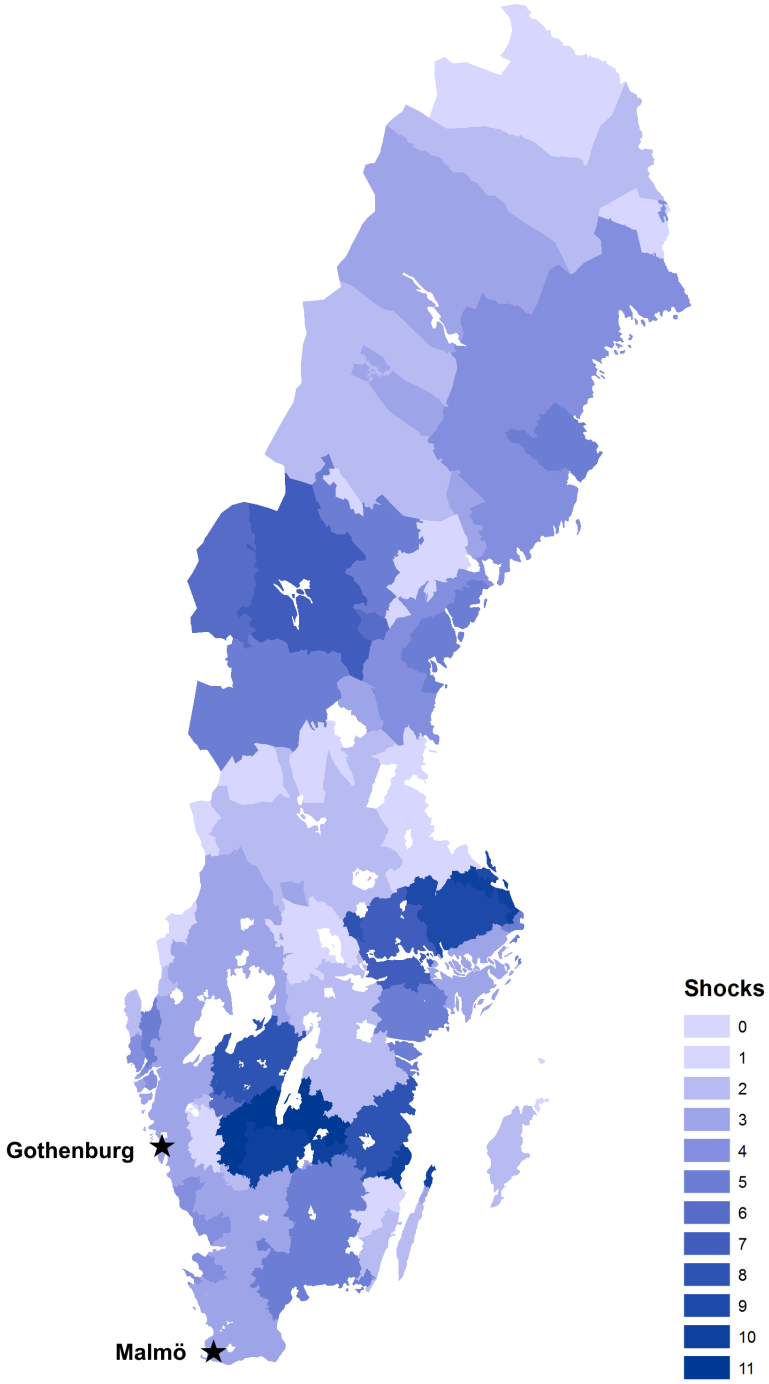
Notes: Share of total emigrants per year 1869–1920 by port of emigration. The figure shows that Gothenburg and Malmö were the main emigration ports in the mass migration period. This motivates our use of closeness to these cities to define the instrumental variable. Stockholm and Norrköping, the first and third largest cities at the time, had minor shares of emigration. Source: passenger list data set.

FIGURE 4: FREQUENCY DISTRIBUTION OF FROST SHOCKS 1864–1867



Notes: Distribution of frost shocks during 1864-1867 by growing and non-growing season. Shocks are defined at a monthly resolution. For example, a value of 5 in Panel A indicates that a municipality experienced 5 growing-season months with above-average frost between 1864 and 1867.

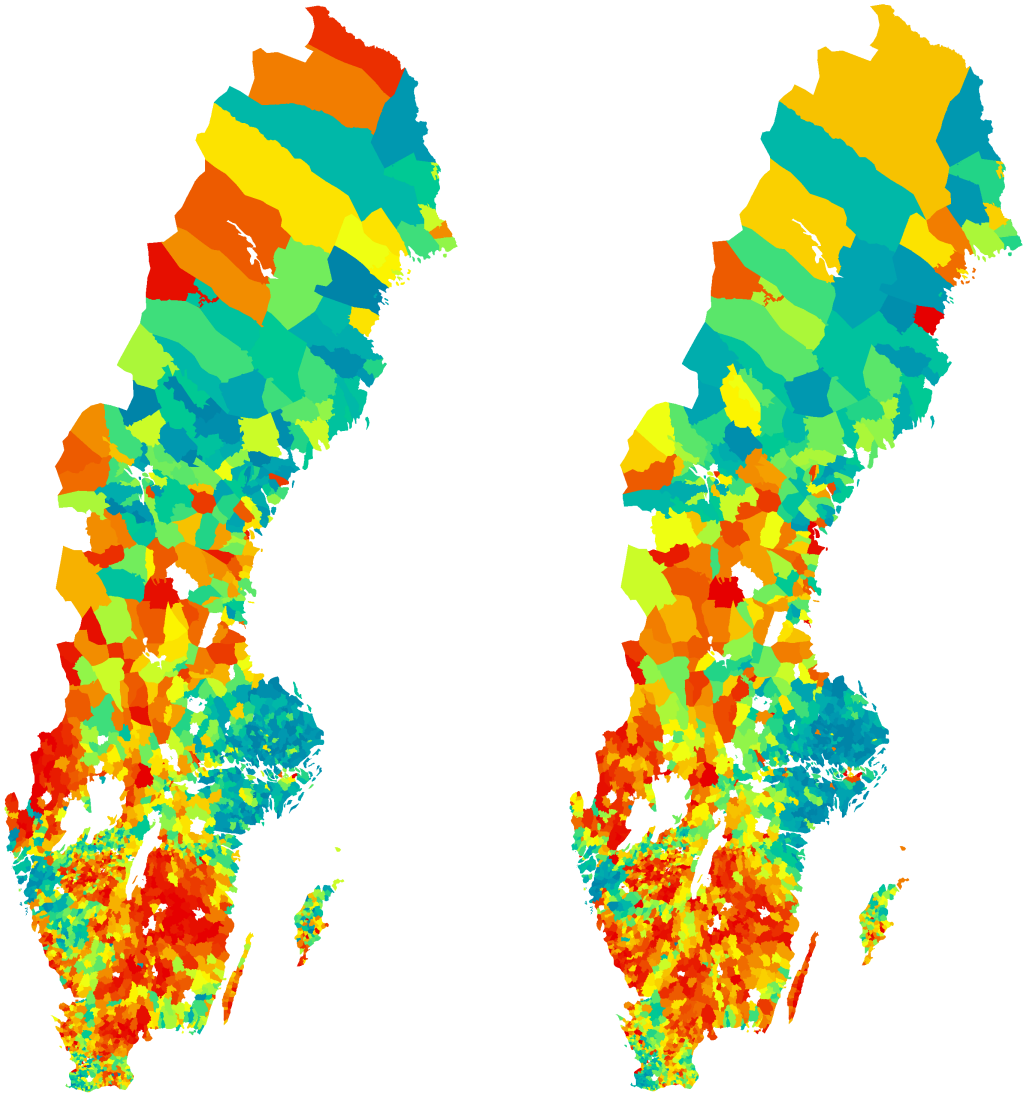
FIGURE 5: SPATIAL DISTRIBUTION OF GROWING-SEASON FROST SHOCKS 1864–1867



Notes: This figure displays the spatial distribution of growing-season frost shocks 1864–1867, used to define the instrumental variable. Darker areas indicate a higher number of shocks. Frost shocks are defined by month, relative to the local long-term mean and standard deviation of frost in that month. Gothenburg and Malmö are the two main emigration ports. In our data, 75 percent of municipalities are closest to Gothenburg, while 25 percent are closer to Malmö.

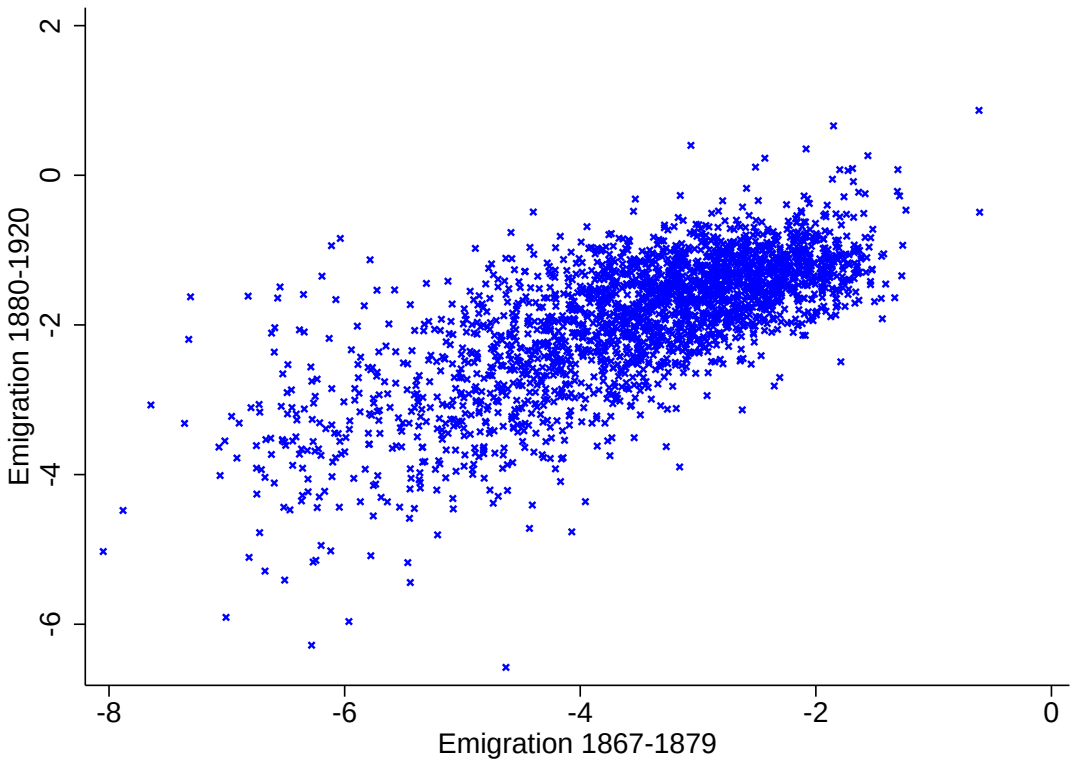
FIGURE 6: SPATIAL DISTRIBUTION OF EMIGRATION

(A) **First wave emigration, 1867–1879** (B) **Total emigration, 1867–1920**



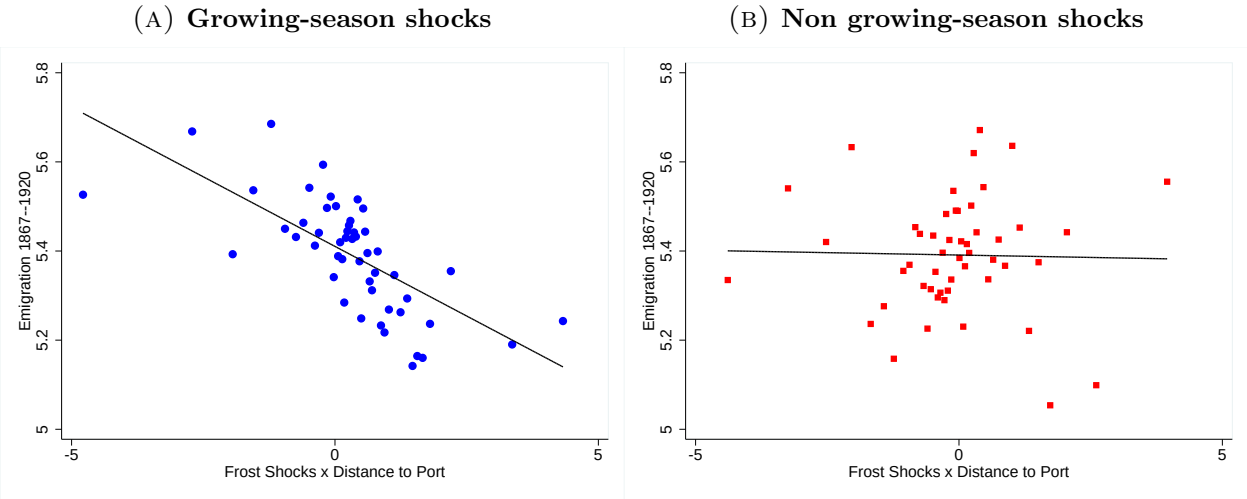
Notes: This figure displays the spatial distribution of emigration during the first wave of emigration (1867–1879) and in total (1867–1920). Each geographical unit represents one municipality. Emigration values are divided by the population in 1865. More red values indicate that a larger fraction of the 1865 population emigrated. Color scales are relative to the distribution in the period in question, hence color comparisons between Panels A and B indicate difference in relative importance across periods.

FIGURE 7: CORRELATION BETWEEN EARLY AND LATE EMIGRATION



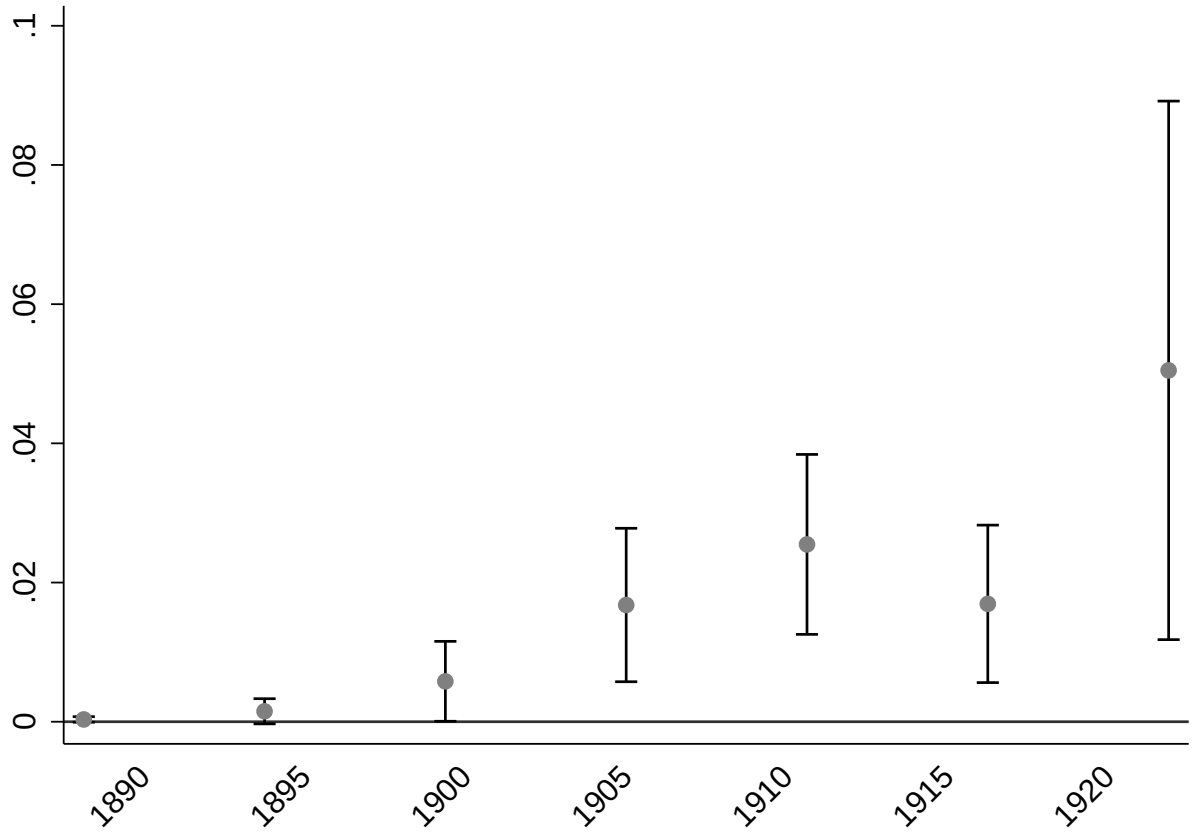
Notes: This figure displays a scatter plot of total emigration during the first wave of emigration (1867–1879) against later emigration (1880-1920). Each dot represents one municipality. Emigration values are in logarithms and divided by the population in 1865.

FIGURE 8: FIRST STAGE: RELATION BETWEEN EMIGRATION AND THE INSTRUMENT



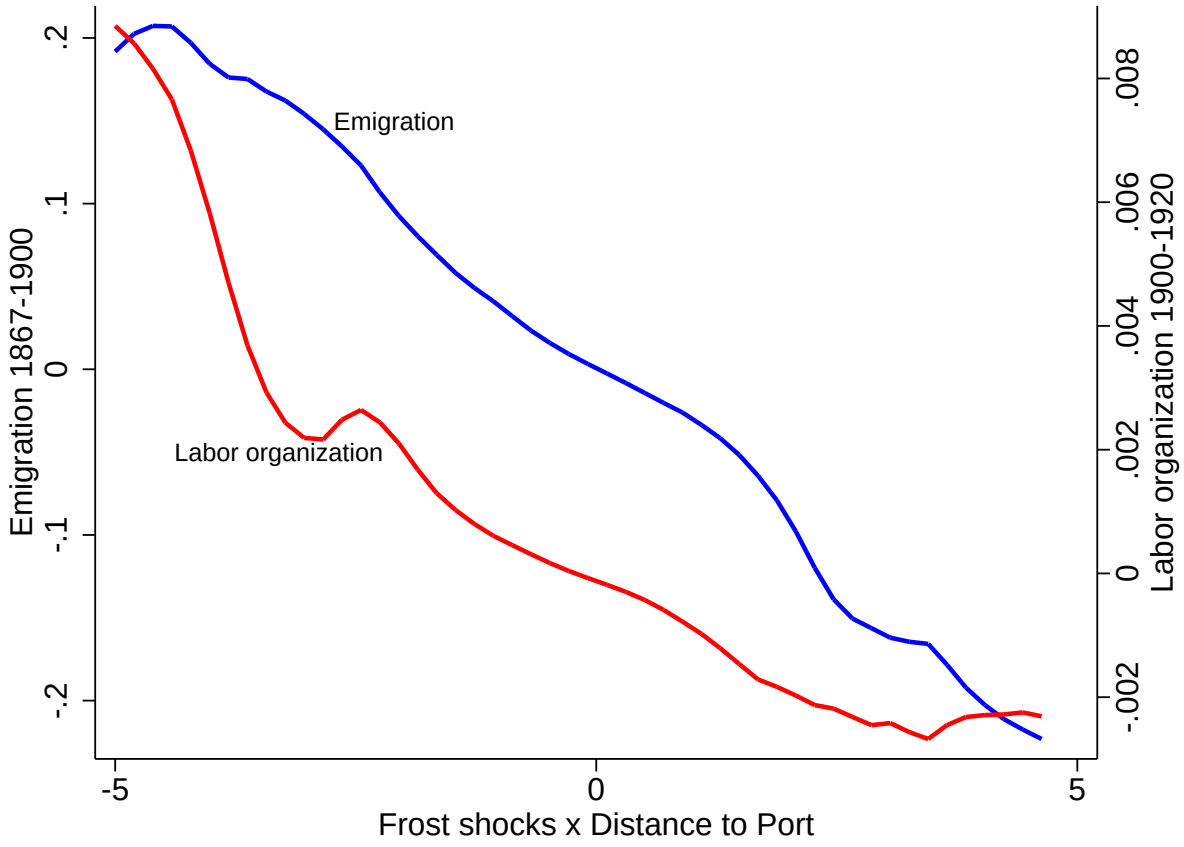
Notes: This figure shows the first stage relationship non-parametrically. Panel A plots log total emigration 1867–1920 against the instrument, defined as the interaction between the number of growing-season frost shocks 1864–1867 and the log distance to the nearest emigration port. Panel B instead displays a placebo instrument using shocks occurring in the non growing-season during the same period. Municipalities are sorted into 50 groups of equal size. Dots indicate the mean value in each group. A linear regression line based on the underlying (ungrouped) data is also shown. Included controls are county fixed effects, frost shocks 1864–1867, log distance to nearest emigration port, log distance to nearest trade port, log distance to weather station, log distance to nearest town, log distance to Stockholm, log population in 1865, log area, latitude, longitude, arable land share in 1810 and indicators for urban municipalities and high soil suitability for the production of barley, oats, wheat, dairy and timber. The total number of observations is 2358.

FIGURE 9: EMIGRATION AND LABOR ORGANIZATION RATES 1890–1920



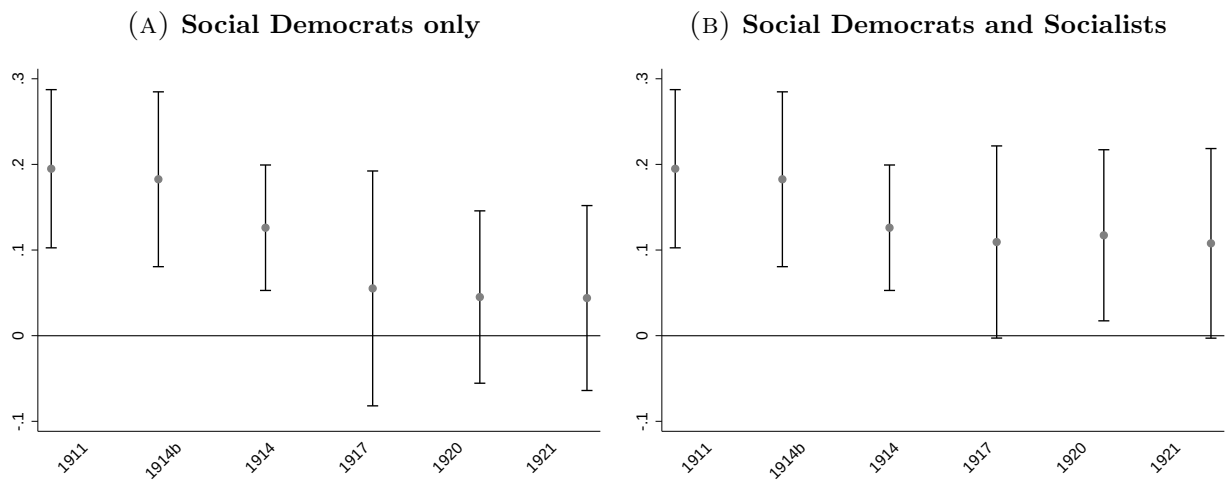
Notes: This figure displays the IV coefficients on the log of total emigration from 1867 to year t on the labor organization rate, defined as the number of members of labor unions and the Social Democratic Party over total population. All regressions include county fixed effects, the log of the population at baseline, log area, latitude, longitude, distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, the number of growing season frost shocks in 1864–1867, the interaction between growing season frost shocks and the log distance to the nearest town and trade port, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber.

FIGURE 10: NONPARAMETRIC EFFECT OF THE INSTRUMENT ON LABOR MOVEMENT AND EMIGRATION



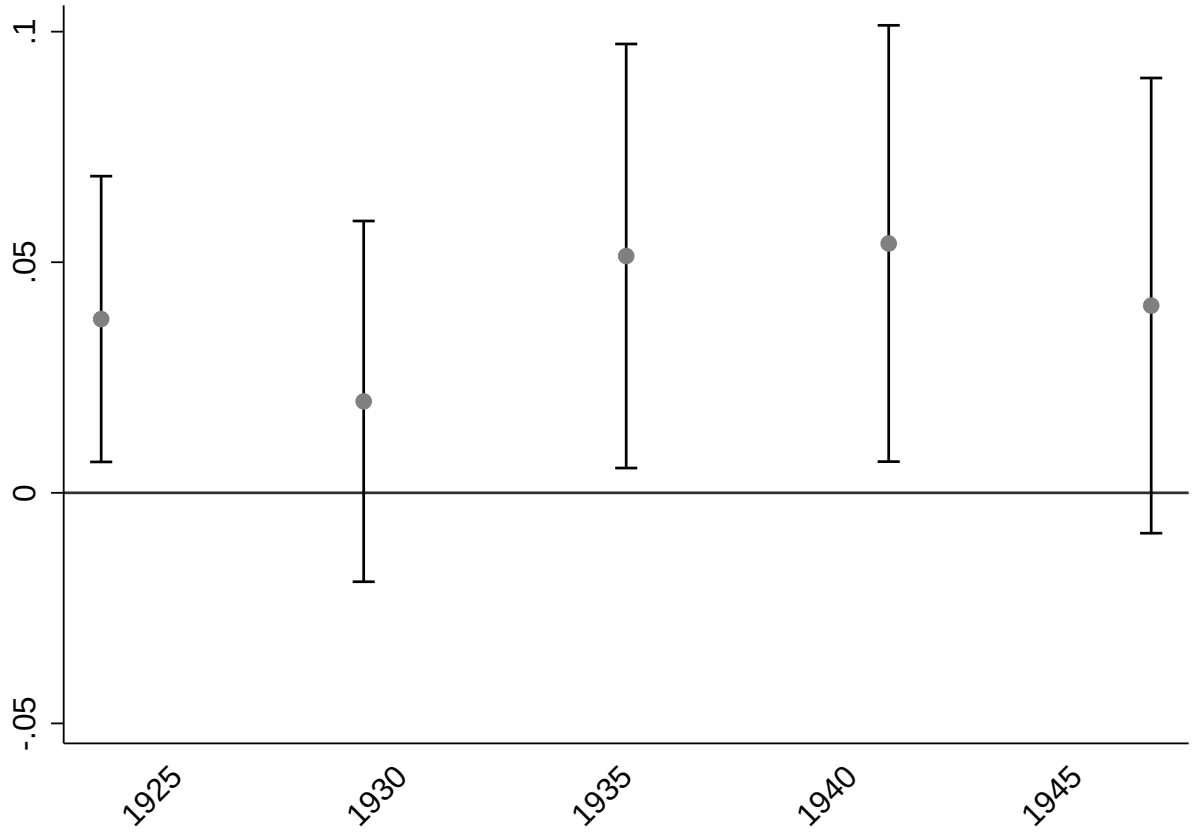
Notes: Local mean smooth. Bandwidth: 1. This figure nonparametrically displays the first stage relationship, as well as the reduced-form effect of the instrument on the average labor organization rate 1900–1920. The instrument is the interaction between the number of growing season frost shocks 1864–1867 and the log distance to the nearest emigration port. All variables have been residualized using the following covariates: county fixed effects, the log of the population at baseline, log area, latitude, longitude, distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, the number of growing season frost shocks in 1864–1867, the interaction between growing season frost shocks and the log distance to the nearest town and trade port, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. 16 observations that have residuals values below -5 have been bottom coded at -5 to reduce noise.

FIGURE 11: EMIGRATION AND LEFT-WING VOTE SHARES IN NATIONAL ELECTIONS 1911-1921



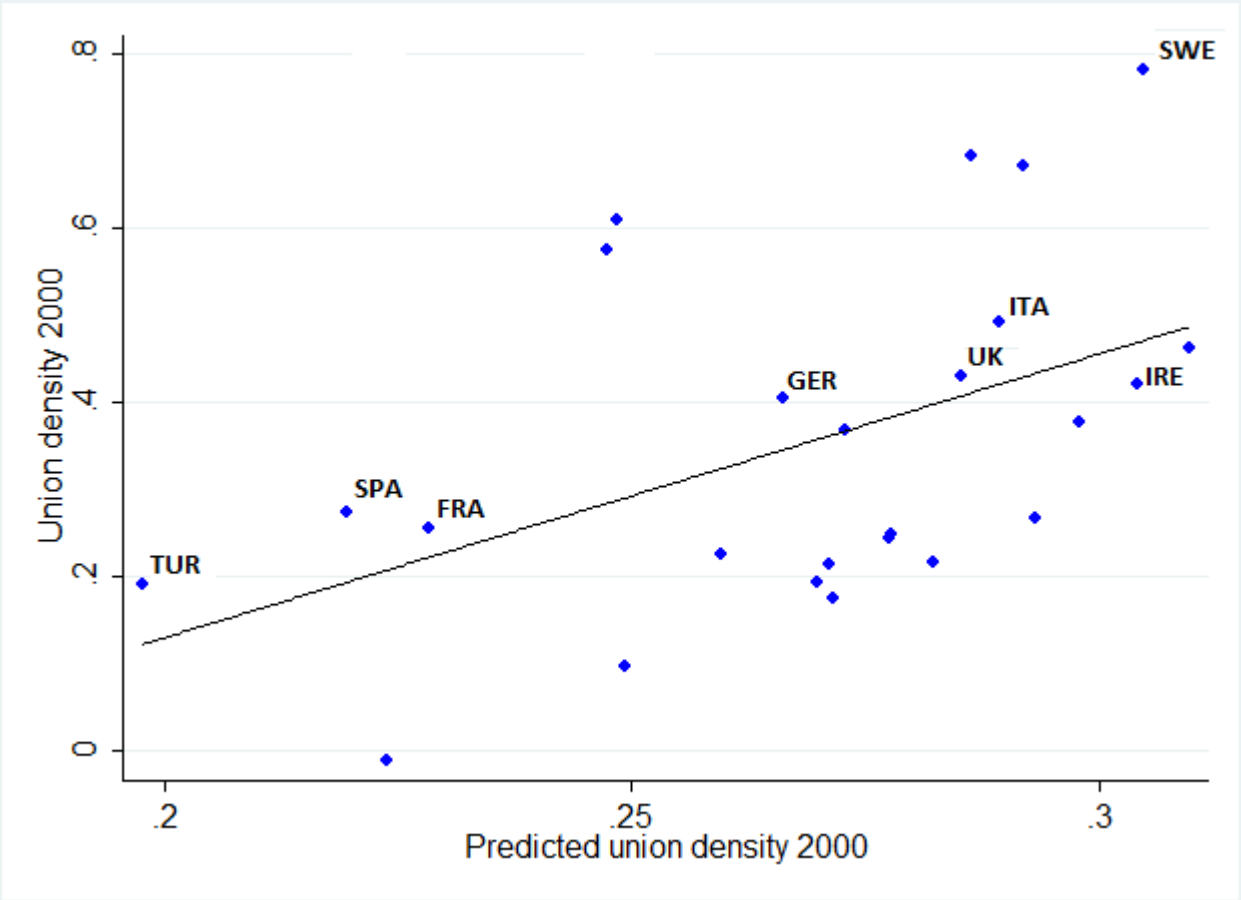
Notes: This figure displays the IV coefficients on the log of total emigration from 1867 to year t on vote shares for the Social Democratic Party (Panel A) and both the Social Democratic and Socialist parties (Panel B). All regressions include county fixed effects, the log of the population at the baseline, log area, latitude, longitude, distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, the number of growing season frost shocks in 1864–1867, the interaction between growing season frost shocks and the log distance to the nearest town and trade port, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. The excluded instrument is the interaction between the number of growing season frost shocks 1864–1867 and the log distance to the nearest emigration port. Standard errors are clustered at the weather station level. Bars around point estimates represent 95 percent confidence intervals.

FIGURE 12: EMIGRATION AND LABOR ORGANIZATION RATES 1925-1945



Notes: This figure displays the IV coefficients on the log of total emigration from 1867 to year t on the labor organization rate, defined as the number of members of labor unions and the Social Democratic Party over total population. All regressions include county fixed effects, the log of the population at baseline, log area, latitude, longitude, distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, the number of growing season frost shocks in 1864–1867, the interaction between growing season frost shocks and the log distance to the nearest town and trade port, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. The excluded instrument is the interaction between the number of growing season frost shocks 1864–1867 and the log distance to the nearest emigration port. Standard errors are clustered at the weather station level. Bars around point estimates represent 95 percent confidence intervals.

FIGURE 13: EMIGRATION AND TRADE UNION DENSITY ACROSS OECD COUNTRIES



Notes: This figure displays a scatter plot of trade union density in 2000 across 25 OECD countries on the y-axis, with predicted union density on the x-axis. Predicted union density is estimated using the estimates in Column 6, Panel B of Table 7 multiplied by the log population residing in the United States in 1920 for each country. The data control for log population size in 1820. A regression line based on the underlying data is drawn, also controlling for initial population.

Table 1: Summary statistics

	Mean	SD	P10	P50	P90
Frost shocks 1864–1867	4.578	2.764	2.000	3.000	9.000
Distance to emigration port	5.024	1.012	3.745	5.137	6.084
Emigration 1867–1920	5.391	1.271	3.664	5.485	6.922
Emigration 1867–1879	3.590	1.482	1.609	3.714	5.394
Emigration 1880–1920	5.172	1.259	3.526	5.247	6.682
Labor organization 1900–1920	0.012	0.037	0.000	0.001	0.031
Strike participants 1909	0.011	0.041	0.000	0.000	0.027
Left vote share 1911–1921	0.246	0.188	0.030	0.210	0.519
Turnout 1911–1921	0.603	0.099	0.476	0.606	0.725
Welfare exp. per capita 1918	2.413	2.116	0.841	2.064	4.183
Welfare exp. per capita 1919	2.756	1.942	0.972	2.381	4.794
Direct democracy 1919	0.634	0.482	0.000	1.000	1.000
Population 1865	7.079	0.782	6.094	7.047	8.076
Urban	0.048	0.214	0.000	0.000	0.000
Area	8.633	1.263	7.209	8.482	10.145
Arable land share	0.702	0.220	0.500	0.667	1.000
Distance to trade port	4.390	0.928	3.215	4.518	5.294
Distance to town	2.879	0.842	1.933	2.920	3.833
Distance to Stockholm	5.531	0.768	4.452	5.746	6.193
Distance to railway	3.152	1.429	1.317	3.264	4.907
Distance to station	3.482	0.681	2.620	3.594	4.174
Latitude	58.336	2.022	55.881	58.170	60.417
Longitude	14.823	2.064	12.594	14.217	17.859
Barley suitability	0.239	0.426	0.000	0.000	1.000
Oat suitability	0.136	0.343	0.000	0.000	1.000
Wheat suitability	0.177	0.382	0.000	0.000	1.000
Livestock suitability	0.223	0.417	0.000	0.000	1.000
Lumber suitability	0.179	0.384	0.000	0.000	1.000

Notes: This table provides summary statistics for emigration as well as outcome and control variables. Emigration, population, area and distance variables are in logs.

Table 2: Balance tests

	Frost shocks × Distance to emigration port
Population 1865	0.056**
	(0.021)
Urban	0.004
	(0.003)
Area	0.017
	(0.032)
Arable land share	-0.003
	(0.006)
Latitude	0.029
	(0.029)
Longitude	0.021
	(0.025)
Distance to trade port	0.016
	(0.027)
Distance to town	-0.034
	(0.026)
Distance to Stockholm	0.028
	(0.031)
Distance to railway	0.017
	(0.049)
Distance to weather station	0.001
	(0.023)
Barley suitability	-0.008
	(0.010)
Oat suitability	0.002
	(0.007)
Wheat suitability	-0.004
	(0.004)
Livestock suitability	-0.006
	(0.011)
Timber suitability	0.013
	(0.009)
Infant mort. per 1000 births	0.169
	(0.114)
Child mort. per 1000 births	0.065
	(0.117)
Maternal mort. per 1000 births	0.013
	(0.059)

Notes: OLS regressions. Each row represents a separate regression. This table displays the results of a randomization test on the instrument, defined as the number of growing season frost shocks 1864-1867 interacted with the log distance to the nearest emigration port. Population, area and distance variables are in logs. All columns include county fixed effects and control for the main effects of growing season frost shocks and log distance to nearest port. The final three variables have fewer observations (1784, 1778 and 1268, respectively) and are therefore not used as controls in the main analysis. Standard errors clustered at the weather station level in parentheses. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 3: Frost shocks and agricultural outcomes in a panel 1860–1870

Dependent variable:	Crop failure		Harvest Grade	
	(1)	(2)	(3)	(4)
Frost Shocks	0.166**	0.167**	-1.025**	-1.022**
	(0.060)	(0.062)	(0.424)	(0.423)
Frost Shocks NGS		-0.010		-0.044
		(0.042)		(0.095)
County fixed effects	Yes	Yes	Yes	Yes
County linear trend	Yes	Yes	Yes	Yes
Observations	264	264	264	264

Notes: Columns 1-2: OLS regressions. Columns 3-4: Ordered probit regressions. This table displays the effect of frost shocks on county level agricultural outcomes in a panel 1860–1870. The dependent variable in Columns 1 and 2 is a yearly indicator of crop failure, defined as a harvest grade below 3 on a scale from 0 to 6. The dependent variable in Columns 3 and 4 is the full harvest grade index. *Frost Shocks* is the mean number of growing season frost shocks among a county's municipalities. *Frost Shocks NGS* is defined analogously but for the non-growing season. Both variables are normalized by their standard deviations. Regressions are weighted by arable land area. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 4: First stage: Frost shocks and emigration 1867-1920

Dependent variable:	Emigration 1867-1920			
	(1)	(2)	(3)	(4)
Shocks×Distance to port	-0.064***	-0.062***	-0.062***	-0.063***
	(0.016)	(0.013)	(0.013)	(0.013)
Shocks	0.005	0.015**	0.012	0.013
	(0.007)	(0.007)	(0.009)	(0.009)
Shocks×Distance to trade port			0.014	0.011
			(0.024)	(0.024)
Shocks×Distance to town			-0.004	-0.005
			(0.008)	(0.008)
Shocks NGS×Distance to port				0.010
				(0.015)
Shocks NGS				0.002
				(0.011)
County fixed effects	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes
Observations	2358	2358	2358	2358

Notes: OLS regressions. This table displays the effects on log emigration of frost shocks 1864-1867 and their interaction with log distance to the nearest emigration port. *Shocks NGS* indicate frost shocks occurring in the non-growing season. All regressions include county fixed effects and control for the log of the population at baseline. Additional control variables are log area, latitude, longitude, distance to nearest town, trade port, Stockholm and weather station, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. *Shocks* × *Market Access* includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 5: Intertemporal Elasticity of Emigration

A. Dependent variable:	Emigration 1867-1879				
	(1) OLS	(2) OLS	(3) OLS		
Shocks×Distance to port	-0.067*	-0.064**	-0.065**		
	(0.034)	(0.030)	(0.030)		
Shocks	0.039**	0.051***	0.048**		
	(0.015)	(0.015)	(0.020)		
County fixed effects	Yes	Yes	Yes		
Controls	No	Yes	Yes		
Shock×Market Access	No	No	Yes		
B. Dependent variable:	Emigration 1880-1920				
	(1) IV	(2) IV	(3) IV	(4) OLS	(5) OLS
Emigration 1867-1879	0.949**	0.965**	0.944**	0.391***	0.381***
	(0.378)	(0.384)	(0.377)	(0.022)	(0.022)
County fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes
Shock×Market Access	No	No	Yes	No	No
Observations	2358	2358	2358	2358	2358
F-statistic	3.87	4.58	4.58		

Notes: OLS and IV regressions. Panel A displays the effects of frost shocks, log distance to nearest emigration port and their interaction on log emigration 1867-1879. Panel B displays the relationship between early and late emigration. The excluded instrument in Columns 1 to 3 of Panel B is the number of frost shocks interacted with the log distance to the nearest emigration port. Controls for main effects of shocks and distance to emigration port are included. All regressions include county fixed effects and control for the log of the population at baseline. Additional control variables are log area, latitude, longitude, distance to nearest town, trade port, Stockholm and weather station, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. The F-statistic refers to the excluded instrument. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 6: Elasticity of emigration with respect to US-Sweden business cycle differences

Dependent variable:	Yearly emigration 1880-1920		
	(1)	(2)	(3)
US-SWE GDP × Shock × Dist. to port	-0.210***	-0.168***	-0.164***
	(0.052)	(0.053)	(0.046)
Municipality FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Regional trends	No	Yes	Yes
Covariate trends	No	No	Yes
Observations	96637	96637	96637

Notes: OLS regressions. This table displays how locations with different initial early emigration – as measured by the *Shock* × *Dist. to port* interaction – emigrate as the US business cycle is improved relative to the Swedish one. Yearly emigration is measured as the log of emigrants over population one year prior. $US-SWE\ GDP$ is defined as the difference in log real GDP per capita from The Maddison Project (2013), where the two series have been detrended using a linear trend and controls for three lags before differencing. All specifications control for interaction terms between $US-SWE\ GDP$, *Shock* and *Dist. to port*. *Regional trends* indicates additional controls for linear trends across three major regions of Sweden. *Covariate trends* indicates additional controls for linear trends interacted with the following baseline control variables: log population in 1865, an urban dummy, distance to nearest emigration port and latitude. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 7: The Effect of Emigration on Labor Organization

A. Dependent variable:	Emigration 1867–1900			Labor organization 1900–1920		
	First stage			Reduced form		
	(1)	(2)	(3)	(4)	(5)	(6)
Shocks×Distance to port	-0.066***	-0.064***	-0.064***	-0.001***	-0.001***	-0.001***
	(0.018)	(0.015)	(0.015)	(0.000)	(0.000)	(0.000)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Shocks×Market Access	No	No	Yes	No	No	Yes
B. Dependent variable:	**Labor organization rate 1900–1920**					
	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Emigration 1867–1900	0.009***	0.009***	0.009***	0.022***	0.021***	0.023***
	(0.002)	(0.002)	(0.002)	(0.008)	(0.007)	(0.007)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Shocks×Market Access	No	No	Yes	No	No	Yes
Observations	2357	2357	2357	2357	2357	2357
F-statistic	.	.	.	13.30	18.49	17.89

Notes: OLS and IV regressions. This table displays the effects of log emigration 1967-1900 on the average labor organization rate 1900-1920. All regressions include county fixed effects and control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. The excluded instrument is the interaction between the number of growing season frost shocks 1864-1867 and the log distance to the nearest emigration port. The F-statistic refers to the excluded instrument. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 8: Mobilization of workers

Dependent variable:	Strikers 1909 per capita			Share unionized strikers 1909	Labor org. 1910 per industrial worker	Strikers 1909 per industrial worker
	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV	(6) IV
Emigration 1867–	0.008***	0.031***	0.031***	0.045**	0.222***	0.207**
	(0.002)	(0.011)	(0.011)	(0.018)	(0.068)	(0.096)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Shocks×Market Access	No	No	Yes	Yes	Yes	Yes
Observations	2357	2357	2357	2358	2298	2300
F-statistic		20.79	20.48	20.45	21.63	20.77

Notes: OLS and IV regressions. Columns 1 to 3 report the effects of log emigration 1967-1908 on the participation in the 1909 general strike. The dependent variable in Column 4 is the share of strikers who were union members. It is normalized so that zero indicates equal fractions of union and nonunion strikers. Municipalities without any strikes are also assigned value zero. Coumns 5 and 6 report labor organization and strike participation per number of industrial worker in 1910, rather than per capita, to account for changes in employment structure over time. All regressions include county fixed effects and control for the log of the population at the baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. The excluded instrument is the interaction between the number of growing season frost shocks 1864-1867 and the log distance to the nearest emigration port. The F-statistic refers to the excluded instrument. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 9: Electoral Effects, National Elections 1911-1921

Dependent variable:	Left-wing vote share				Voter turnout				Eligible voter share			
	OLS	IV			OLS	IV			OLS	IV		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Emigration 1867–1910	0.031***	0.122**	0.124**	0.138***	0.016***	0.072*	0.068*	0.075**	0.039	0.044	-0.066	-0.066
	(0.009)	(0.060)	(0.054)	(0.051)	(0.003)	(0.040)	(0.038)	(0.037)	(0.056)	(0.076)	(0.063)	(0.063)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes
Shocks×Market Access	No	No	No	Yes	No	No	No	Yes	No	No	Yes	Yes
Observations	2358	2358	2358	2358	2358	2358	2358	2358	2357	2357	2357	2357
F-statistic	.	15.66	21.39	21.22	.	15.66	21.39	21.22	.	15.64	21.22	21.22

Notes: OLS and IV regressions. This table displays the effects of log emigration 1967-1910 on three outcomes in national elections between 1911 and 1921. The dependent variable in Columns 1 to 4 is the average vote share of the Social Democratic and Socialist parties. The dependent variable in Columns 5 to 8 is voter turnout, defined as the number of cast votes over total eligible voters. The dependent variable in Columns 9 to 12 is the share of eligible voters, defined as eligible voters per capita. All regressions include county fixed effects and control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. The excluded instrument is the interaction between the number of growing season frost shocks 1864-1867 and the log distance to the nearest emigration port. The F-statistic refers to the excluded instrument. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 10: Emigration and census outcomes 1910

	(1) In-migration	(2) Family size	(3) Unmarried	(4) Female ratio
Emigration 1867–1910	3.388	-0.137	0.334	0.143
	(5.160)	(0.297)	(1.275)	(0.884)
County fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Shocks×Market Access	Yes	Yes	Yes	Yes
Observations	2300	2285	2300	2300
F-statistic	21.65	21.25	21.65	21.65
Mean dep. var.	44.84	6.53	35.80	51.25

Notes: IV regressions. This table displays the effects of log emigration 1967-1890 on demographic variables 1890-1910. *In-migration* is the share of inhabitants born in another municipality. *Family size* is the average number of individuals per family. *Unmarried* is the number of unmarried individuals, in percent. *Female ratio* is the share of women, in percent. All regressions include county fixed effects and control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. The excluded instrument is the interaction between the number of growing season frost shocks 1864-1867 and the log distance to the nearest emigration port. The F-statistic refers to the excluded instrument. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 11: Welfare expenditures per capita in 1918 and 1919

Dependent variable:	Expenditures per capita 1918				Expenditures per capita 1919			
	(1) OLS	(2) IV	(3) IV	(4) IV	(5) OLS	(6) IV	(7) IV	(8) IV
Emigration 1867-	0.227**	1.049**	1.080**	1.038***	0.121	1.108**	1.121***	1.126***
	(0.098)	(0.506)	(0.456)	(0.389)	(0.077)	(0.526)	(0.353)	(0.339)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	No	Yes	Yes	Yes	No	Yes	Yes
Shock×Market Access	No	No	No	Yes	No	No	No	Yes
Observations	2218	2218	2218	2218	2202	2202	2202	2202
F-statistic	.	13.57	15.02	15.76	.	12.93	14.79	15.08

Notes: OLS and IV regressions. This table displays the effects of log emigration 1967-1918 on per capita welfare expenditures in 1918 and 1919. All regressions include county fixed effects and control for the log of the population at the baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. The excluded instrument is the interaction between the number of growing season frost shocks 1864-1867 and the log distance to the nearest emigration port. The F-statistic refers to the excluded instrument. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 12: Effect of Emigration on Choice of Political Institutions

A. Dependent variable:	Emigration 1867-			Representative democracy		
	First stage			Reduced form		
	(1)	(2)	(3)	(4)	(5)	(6)
	1919	1920	1938	1919	1920	1938
Shocks×Distance to port	-0.060***	-0.060***	-0.058***	-0.003*	-0.004*	-0.009**
	(0.016)	(0.016)	(0.014)	(0.001)	(0.002)	(0.004)
Threshold dummies	Yes	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Shock×Market Access	Yes	Yes	Yes	Yes	Yes	Yes
B. Dependent variable:	Voluntary transition to representative democracy					
	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
	1919	1920	1938	1919	1920	1938
Emigration 1867-	0.010**	0.019**	0.028**	0.048**	0.059**	0.156**
	(0.004)	(0.007)	(0.012)	(0.022)	(0.026)	(0.068)
Threshold dummies	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Shocks×Market Access	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2220	2220	2207	2220	2220	2207
F-statistic				15.05	14.97	17.38

Notes: IV regressions. This table displays the effects of log emigration 1867 to 1918 on the form of government and on having a population above 1500. The dependent variable in Column 8 is an indicator for having representative democracy while being under the mandatory population threshold of 1500 inhabitants. All regressions include county fixed effects and control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. The excluded instrument is the interaction between the number of growing season frost shocks 1864-1867 and the log distance to the nearest emigration port. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 13: Emigration and membership in non-labor organizations

Dependent variable:	Free church members			Temperance lodge members		
	(1) OLS	(2) IV	(3) IV	(4) OLS	(5) IV	(6) IV
Emigration 1867–1900	0.005***	-0.007	-0.008	0.007***	-0.032**	-0.030*
	(0.001)	(0.008)	(0.008)	(0.002)	(0.015)	(0.016)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Shocks×Market Access	No	No	Yes	No	No	Yes
Observations	2357	2357	2357	2357	2357	2357
F-statistic		18.49	17.89		18.49	17.89

Notes: OLS and IV regressions. This table displays the effects of log emigration 1967-1918 on per capita membership in non-labor organizations 1900-1920. All regressions include county fixed effects and control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. The excluded instrument is the interaction between the number of growing season frost shocks 1864-1867 and the log distance to the nearest emigration port. The F-statistic refers to the excluded instrument. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 14: Excluding urban municipalities

Dependent variable:	Labor org.		Striking		Left vote		Turnout	
	(1) All	(2) Rural	(3) All	(4) Rural	(5) All	(6) Rural	(7) All	(8) Rural
Emigration 1867–	0.023***	0.018***	0.031***	0.025**	0.139***	0.132**	0.075**	0.092***
	(0.007)	(0.007)	(0.011)	(0.010)	(0.051)	(0.059)	(0.037)	(0.035)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Shocks×Market Access	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2357	2244	2357	2244	2358	2245	2358	2245
F-statistic	17.91	13.49	20.45	15.53	21.19	16.16	21.19	16.16

Notes: IV regressions. This table displays the effects of log emigration 1867 to year t on four outcome variables, with and without urban municipalities. *Labor org.* denotes the average per capita membership in labor unions and the Social Democratic Party 1900–1920. *Striking* denotes per capita strike participation in the 1909 general strike. *Left Vote* denotes the average vote share of the Social Democratic and Socialist parties in national elections 1911-1921. *Turnout* denotes the average turnout rate in those same election. All regressions include county fixed effects and control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. The excluded instrument is the interaction between the number of growing season frost shocks 1864-1867 and the log distance to the nearest emigration port. The F-statistic refers to the excluded instrument. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 15: Long-run Effects of Emigration on Labor Organization

A. Dependent variable:	Emigration 1867–1920			Labor organization 1921–1945		
	First stage			Reduced form		
	(1)	(2)	(3)	(4)	(5)	(6)
Shocks×Distance to port	-0.064***	-0.062***	-0.062***	-0.003**	-0.002**	-0.003**
	(0.016)	(0.013)	(0.013)	(0.001)	(0.001)	(0.001)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Shocks×Market Access	No	No	Yes	No	No	Yes
B. Dependent variable:	Labor organization rate 1921–1945					
	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Emigration 1867–1920	0.029***	0.028***	0.028***	0.042*	0.038*	0.042**
	(0.005)	(0.004)	(0.004)	(0.022)	(0.021)	(0.020)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Shocks×Market Access	No	No	Yes	No	No	Yes
Observations	2357	2357	2357	2357	2357	2357
F-statistic	.	.	.	16.46	22.44	22.60

Notes: OLS and IV regressions. This table displays the effects of log emigration 1967-1920 on the average labor organization rate 1921-1945. All regressions include county fixed effects and control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. The excluded instrument is the interaction between the number of growing season frost shocks 1864-1867 and the log distance to the nearest emigration port. The F-statistic refers to the excluded instrument. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 16: The Effect of Emigration on Contemporary Left-wing Voting

Dependent variable:	Left vote share 1998–2014					
	Reduced-form					
	Municipal elections			National elections		
	(1)	(2)	(3)	(4)	(5)	(6)
Shocks×Distance to port	-0.003	-0.005**	-0.005**	-0.003	-0.004*	-0.005**
	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Shocks×Market Access	No	No	Yes	No	No	Yes
Dependent variable:	Left vote share 1998–2014					
	IV					
	Municipal elections			National elections		
	(1)	(2)	(3)	(4)	(5)	(6)
Emigration 1867–1945	0.050	0.075**	0.083**	0.043	0.067*	0.074**
	(0.047)	(0.038)	(0.038)	(0.048)	(0.036)	(0.036)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Shocks×Market Access	No	No	Yes	No	No	Yes
Observations	2357	2357	2357	2353	2353	2353
F-statistic	18.64	24.69	25.43	18.68	24.75	25.53

Notes: OLS and IV regressions. This table displays the effects of log emigration 1967-1945 on the average vote share of the Social Democratic and Socialist parties 1998-2014. All regressions include county fixed effects and control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. The excluded instrument is the interaction between the number of growing season frost shocks 1864-1867 and the log distance to the nearest emigration port. The F-statistic refers to the excluded instrument. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table 17: Emigration and trade union density across OECD countries

Dependent variable:	Union density 1960			Union density 2000		
	(1)	(2)	(3)	(4)	(5)	(6)
Log population in US 1920	0.073**	0.127**		0.074***	0.099**	
	(0.024)	(0.037)		(0.026)	(0.035)	
Predicted Union Density			5.625**			4.390**
			(1.619)			(1.544)
Log population 1820	-0.107***	-0.100***	-0.100***	-0.132***	-0.109***	-0.109***
	(0.022)	(0.026)	(0.026)	(0.028)	(0.035)	(0.035)
Controls	No	Yes	Yes	No	Yes	Yes
Observations	14	14	14	25	21	21
R-squared	0.52	0.83	0.83	0.43	0.59	0.59

Notes: OLS regressions. This table displays the relationship between US emigration and trade union density across OECD countries. Predicted Union Density is computed using the estimate from Column 6, Panel B of Table 7 and the log population in the US 1920. Controls include: real GDP per capita in 1960, rural share of the population in 1960, life expectancy at birth in 1960 and length of primary and secondary schooling in 1970. Robust standard errors in parentheses. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

A Appendix

A.1 Estimating minimum daily temperatures

In order to fill in the missing values on minimum temperatures, we use the remaining variables to predict daily minimum temperatures. Observations containing minimum daily temperatures are used to fit a model relating minimum recorded daily temperature to the minimum of temperatures at 6 am, 12 pm and 8 pm, respectively, as follows:

$$min(Temp)_{sdmt} = \alpha_0 + \alpha_m min(Temp^{6am}, Temp^{12pm}, Temp^{8pm})_{sdmt} + \nu_{sdmt},$$

where $min(Temp)_{sdmt}$ is the minimum temperature on day d at station s, month m and year t and $min(Temp^{6am}, Temp^{12pm}, Temp^{8pm})_{sdmt}$ is the minimum of the three daily readings. The coefficients are allowed to vary by month to capture seasonal variation in the relationship. We then use this model to predict daily minimum temperatures for observations with missing values.

A.2 Correlation between emigration data sets

Appendix Table A.1 quantifies the correlation by regressing the passenger list data on the parish data using the 1869-1895 period when both data sets are available. As in the remainder of the analysis, we aggregate all data to the municipality level using 1865 borders. Using county fixed effects, the estimated relationship is 1.3 passenger-data emigrants for every church-data emigrant. This reflects the fact that some parishes are not fully matched in the passenger data, leading to underestimation in the latter. Controlling for municipality fixed effects, however, the point estimate becomes statistically indistinguishable from one. This indicates that for those parishes from the passenger data that we are able to match, the two data sets report the same number of emigrants on average. The R-squared value of 0.84 indicates a high degree of similarity. For comparison, in Columns 3 and 4, we use one-year lagged values of emigration from both data sets to predict parish emigration. Both models return lower and similar point estimates of 0.70 and 0.63. Taken together, these results suggest a high reliability of the emigrant data sets and that there is no important lag between leaving the home parish and boarding a ship to the United States.

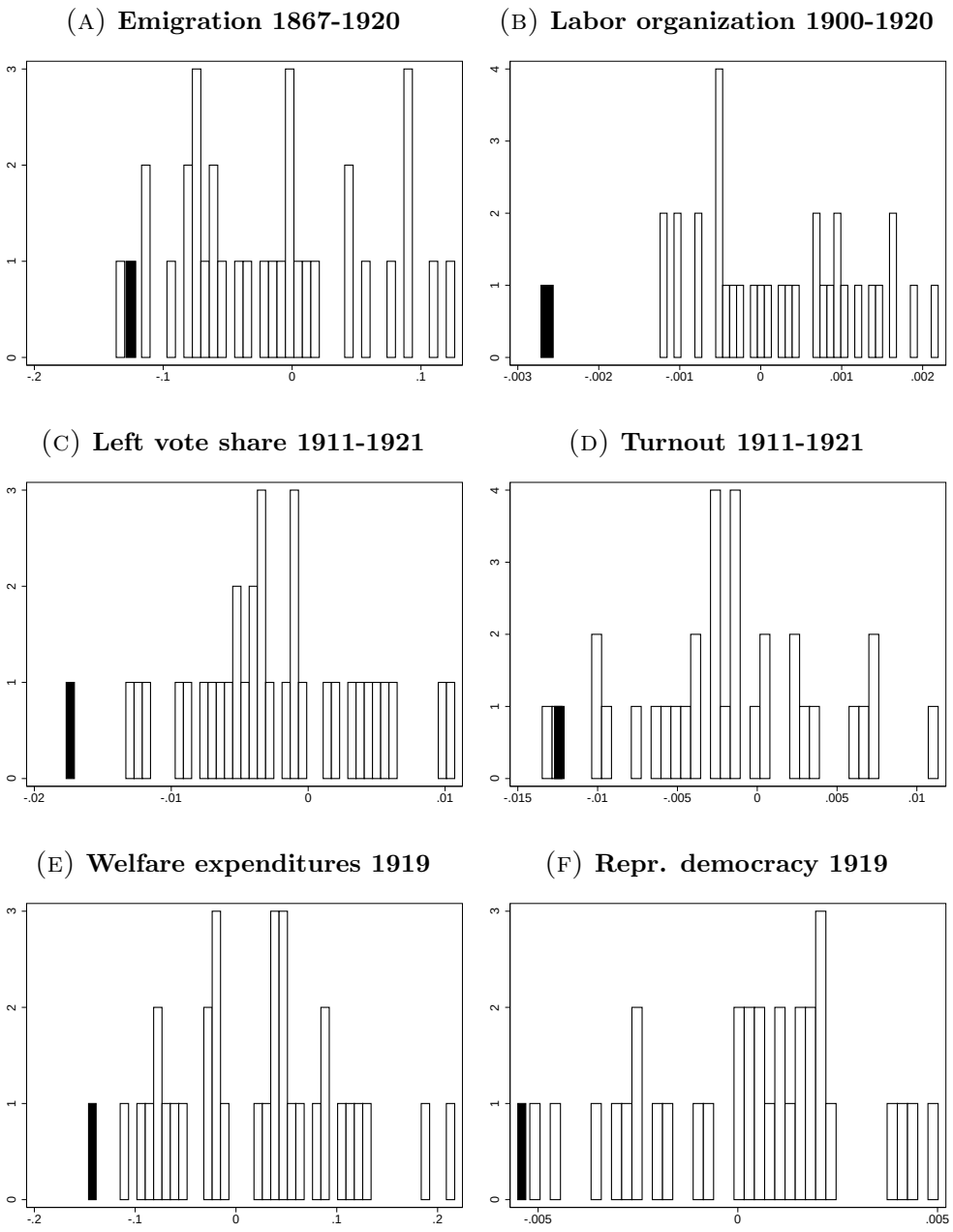
TABLE A.1: COMPARISON BETWEEN EMIGRATION DATA SETS

	Church book emigrants			
	(1)	(2)	(3)	(4)
Passenger emigrants	1.288***	0.987***		
	(0.068)	(0.062)		
Lag church book emigrants			0.703***	
			(0.041)	
Lag passenger emigrants				0.639***
				(0.081)
County FE	Yes	No	No	No
Municipality FE	No	Yes	Yes	Yes
Observations	64530	64557	64557	64557
R-squared	0.66	0.84	0.83	0.73

Notes: OLS regressions. This table displays the relationship between the church book and passenger list data 1869-1895. Lag variables are lagged one year.

A.3 Additional material

FIGURE A.1: TREATMENT AND PLACEBO SHOCKS 1859-1900



Notes: Probability density functions of reduced form coefficients from regressing the outcome variables on growing season frost shocks times distance to emigration port during all consecutive four-year periods between 1859 and 1900. The coefficient associated with the treatment period of 1864–1867 is highlighted in black. Frost shocks are categorized into quintiles of the distribution before interacting with port distance. All regressions include county fixed effects and control for the log of the population at baseline. All regressions control for growing season frost shocks in the relevant four-year period, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. Regressions also include the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively.

Table A.2: Frost shocks and emigration in the panel 1860-1920

Dependent variable:	Yearly emigration per capita			
	1860-1879		1880-1920	
	(1)	(2)	(3)	(4)
Shocks×Distance to port	-0.014***	-0.014***	-0.003	-0.003
	(0.003)	(0.003)	(0.004)	(0.004)
Shocks	0.127***	0.127***	0.018	0.018
	(0.017)	(0.017)	(0.018)	(0.018)
Municipality FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Linear trend	No	Yes	No	Yes
Observations	47792	47792	74091	74091

Notes: OLS regressions. This table displays the relationship between growing season frost shocks, log distance to nearest emigration port and their interaction using panel data 1860-1920.

TABLE A.3: BOUNDING IDEOLOGICAL SELECTION OF EMIGRANTS

Dependent variable:	Left-wing vote share 1911–1921			
	(1) Main result	(2) 75 percent	(3) 90 percent	(4) 100 percent
Emigration 1867–1910	0.138***	0.111***	0.091**	0.077**
	(0.051)	(0.034)	(0.036)	(0.037)
County fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Shocks×Market Access	Yes	Yes	Yes	Yes
Observations	2358	2358	2358	2358
F-statistic	21.22	21.22	21.22	21.22

Notes: IV regressions. This table puts lower bounds on the size of our estimated effect of emigration on left-wing party voting, taking into account the possibility of ideological selection of emigrants. The basic assumption is that all emigrants would have been eligible to vote and would have voted in all elections 1911-1921. Columns 2 to 4 then consider 3 different scenarios for how emigrants would have voted if they had stayed. Column 2 assumes that 75 percent of all emigrants would have voted for the non-left. Columns 3 and 4 make this assumption 90 and 100 percent, respectively. All regressions include county fixed effects and control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table A.4: Non-growing season shocks as placebo instrument

	Placebo test using non-growing season shocks							
	(1) Emi. 1867–1920	(2) Labor	(3) Strike	(4) Left	(5) Turnout	(6) Welfare 1918	(7) Welf. 1919	(8) Repr. dem. 1919
Shocks NGS×Distance to port	0.006	0.000	-0.000	0.000	0.000	-0.055	-0.012	-0.001
	(0.021)	(0.000)	(0.001)	(0.004)	(0.002)	(0.041)	(0.030)	(0.002)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Shocks×Market Access	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2358	2357	2358	2358	2357	2218	2206	2220

Notes: OLS regressions. This table displays the effects of non-growing season frost shocks 1864–1867 interacted with distance to emigration port on our outcome variables. *Emi. 1867-1920* denotes log total emigration 1867–1920 in a municipality. *Labor* denotes average per capita membership in the labor movement 1900–1920. *Strike* denotes per capita participation in the 1909 general strike. *Left* and *Turnout* denote average vote share of left-wing parties and average turnout in national elections 1911–1921. *Welfare 1918* and *Welf. 1919* denote per capita expenditures on welfare in 1918 and 1919, respectively. *Repr. Dem. 1919* denotes whether a municipality had voluntarily adopted representative rather than direct democracy in 1919. All regressions control for the log of the population at baseline. Additional control variables are log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between frost shocks and the log distance to the nearest town and trade port, respectively. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table A.5: Different cutoffs for frost shocks

	(1) Emi. 1867–1920	(2) Labor	(3) Strike	(4) Turnout	(5) Left	(6) Welfare 1918	(7) Welf. 1919	(8) Repr. dem. 1919
A.	**Shocks defined with 1 standard deviations as threshold**							
Shocks×Distance to port	-0.062***	-0.001***	-0.002***	-0.005**	-0.009***	-0.065***	-0.067***	-0.003*
	(0.013)	(0.000)	(0.001)	(0.002)	(0.002)	(0.023)	(0.022)	(0.001)
B.	**Shocks defined with 0.75 standard deviations as threshold**							
Shocks×Distance to port	-0.047***	-0.001***	-0.002***	-0.006***	-0.008***	-0.065***	-0.050*	-0.003**
	(0.013)	(0.000)	(0.001)	(0.002)	(0.002)	(0.024)	(0.025)	(0.001)
C.	**Shocks defined with 1.25 standard deviations as threshold**							
Shocks×Distance to port	-0.060***	-0.002***	-0.002***	-0.003	-0.008***	-0.063**	-0.063**	-0.003*
	(0.013)	(0.000)	(0.000)	(0.002)	(0.002)	(0.025)	(0.023)	(0.002)
D.	**Growing season cutoff at 5 degrees C instead of 3**							
Shocks×Distance to port	-0.049***	-0.001***	-0.001*	-0.003	-0.006**	-0.059**	-0.060***	-0.003**
	(0.011)	(0.000)	(0.001)	(0.002)	(0.003)	(0.022)	(0.020)	(0.001)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Shocks×Market Access	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2358	2357	2357	2358	2358	2218	2206	2220

Notes: OLS regressions. This table displays the sensitivity of the main results to changing the definition of frost shocks. Panel A displays the baseline reduced-form results. Panels B and C displays results from counting frost shocks with a 0.75 or 1.25 deviation cutoff. Panel D maintains the baseline cutoff of 1 standard deviation but counts as growing season months with a long-run mean temperature above 5 degrees Celsius, rather than 3 as in the baseline. *Emi. 1867-1920* denotes log total emigration 1867–1920 in a municipality. *Labor* denotes average per capita membership in the labor movement 1900–1920. *Strike* denotes per capita participation in the 1909 general strike. *Left* and *Turnout* denote average vote share of left-wing parties and average turnout in national elections 1911–1921. *Welfare 1918* and *Welf. 1919* denote per capita expenditures on welfare in 1918 and 1919, respectively. *Repr. Dem. 1919* denotes whether a municipality had voluntarily adopted representative rather than direct democracy in 1919. All regressions control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. $Shocks \times Market\ Access$ includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table A.6: Sensitivity to large values: Winsorizing variables

	(1) Emi. 1867–1920	(2) Labor	(3) Strike	(4) Turnout	(5) Left	(6) Welfare 1918	(7) Welf. 1919	(8) Repr. dem. 1919
A.	**Censor shocks and port distance at 5^{th} and 95^{th} percentiles**							
Shocks × Distance to Port	-0.062***	-0.001***	-0.002***	-0.006**	-0.007***	-0.058**	-0.064**	-0.003*
	(0.015)	(0.000)	(0.001)	(0.002)	(0.002)	(0.024)	(0.023)	(0.002)
B.	**Censor shocks and port distance at 10^{th} and 90^{th} percentiles**							
Shocks × Distance to Port	-0.083***	-0.002***	-0.002***	-0.009***	-0.007***	-0.059**	-0.067***	-0.003*
	(0.018)	(0.000)	(0.001)	(0.003)	(0.002)	(0.027)	(0.023)	(0.002)
C.	**Censor all variables at 5^{th} and 95^{th} percentiles**							
Shocks × Distance to Port	-0.058***	-0.001***	-0.002**	-0.006***	-0.008***	-0.068**	-0.075***	-0.004**
	(0.017)	(0.000)	(0.001)	(0.002)	(0.002)	(0.025)	(0.024)	(0.002)
D.	**Censor all variables at 10^{th} and 90^{th} percentiles**							
Shocks × Distance to Port	-0.079***	-0.001***	-0.002**	-0.008***	-0.010***	-0.062**	-0.068***	-0.004**
	(0.020)	(0.000)	(0.001)	(0.003)	(0.003)	(0.025)	(0.020)	(0.002)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Shocks×Market Access	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2358	2357	2357	2358	2358	2218	2206	2220

Notes: OLS regressions. This table displays the sensitivity of the main results to large values in the model's variables. It does so by censoring variables at the bottom and top 5^{th} (10^{th}) percentile of the distribution. That is, observations below the 5^{th} percentile of values are assigned the value at the 5^{th} percentile, and so on. The first two panels censor two variables: growing-season frost shocks 1864–1867 and distance to port, and re-define the instrument using these new variables. The bottom two panels repeat the exercise for *all* non-binary variables that are in the model. *Emi. 1867-1920* denotes log total emigration 1867–1920 in a municipality. *Labor* denotes average per capita membership in the labor movement 1900–1920. *Strike* denotes per capita participation in the 1909 general strike. *Left* and *Turnout* denote average vote share of left-wing parties and average turnout in national elections 1911–1921. *Welfare 1918* and *Welf. 1919* denote per capita expenditures on welfare in 1918 and 1919, respectively. *Repr. Dem. 1919* denotes whether a municipality had voluntarily adopted representative rather than direct democracy in 1919. All regressions control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. *Shocks × Market Access* includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. Standard errors are clustered at the weather station level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table A.7: Varying methods for estimating standard errors

	(1) Emi. 1867–1920	(2) Labor	(3) Strike	(4) Turnout	(5) Left	(6) Welfare 1918	(7) Welf. 1919	(8) Repr. dem. 1919
A.	**Standard errors clustered at county level**							
Shocks×Distance to port	-0.062***	-0.001***	-0.002**	-0.005**	-0.009***	-0.065**	-0.067**	-0.003*
	(0.013)	(0.000)	(0.001)	(0.002)	(0.001)	(0.026)	(0.025)	(0.002)
B.	**Standard errors robust to spatial correlation up to 100 kilometers**							
Shocks×Distance to port	-0.062***	-0.001***	-0.002***	-0.005**	-0.009***	-0.065***	-0.067***	-0.003*
	(0.011)	(0.000)	(0.001)	(0.002)	(0.002)	(0.024)	(0.025)	(0.002)
C.	**Standard errors robust to spatial correlation up to 200 kilometers**							
Shocks×Distance to port	-0.062***	-0.001***	-0.002***	-0.005***	-0.009***	-0.065***	-0.067***	-0.003**
	(0.015)	(0.000)	(0.001)	(0.001)	(0.002)	(0.021)	(0.020)	(0.001)
D.	**Wild cluster-t bootstraped errors at weather station level**							
Shocks×Distance to port	-0.062***	-0.001***	-0.002**	-0.005*	-0.009***	-0.065**	-0.065**	-0.003
	(0.020)	(0.000)	(0.001)	(0.003)	(0.003)	(0.026)	(0.026)	(0.002)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Shocks×Market Access	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2358	2357	2357	2358	2358	2218	2218	2220

Notes: OLS regressions. This table estimates reduced-form regressions on the main outcomes, including emigration, using different standard errors. Panel A uses cluster-robust standard errors at the county-level (24 clusters). Panel B and C estimate spatial correlation-robust standard errors (Conley, 1999), with spatial dependencies allowed up to 100 and 200 kilometers from the center of a municipality, respectively. Panel D estimates wild cluster-t bootstrapped standard errors (Cameron et al., 2008). *Emi. 1867-1920* denotes log total emigration 1867–1920 in a municipality. *Labor* denotes average per capita membership in the labor movement 1900–1920. *Strike* denotes per capita participation in the 1909 general strike. *Left* and *Turnout* denote average vote share of left-wing parties and average turnout in national elections 1911–1921. *Welfare 1918* and *Welf. 1919* denote per capita expenditures on welfare in 1918 and 1919, respectively. *Repr. Dem. 1919* denotes whether a municipality had voluntarily adopted representative rather than direct democracy in 1919. All regressions control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. *Shocks* × *Market Access* includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table A.8: Distance in levels instead of logarithms

	Distance in levels instead of logs							
	(1) Emi. 1867–1920	(2) Labor	(3) Strike	(4) Turnout	(5) Left	(6) Welfare 1918	(7) Welf. 1919	(8) Repr. dem. 1919
Shocks × Distance to Port	-0.025***	-0.001***	-0.001***	-0.001	-0.004***	-0.034***	-0.037***	-0.001*
	(0.005)	(0.000)	(0.000)	(0.001)	(0.001)	(0.010)	(0.010)	(0.001)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Shocks×Market Access	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2358	2357	2357	2358	2358	2218	2206	2220

Notes: OLS regressions. This table estimates reduced form regression on the main outcomes, including emigration, using the distance to nearest port in levels rather than in logarithms. Because point estimates become very small when using the distance in levels, all outcomes have been multiplied by 100. *Emi. 1867-1920* denotes log total emigration 1867–1920 in a municipality. *Labor* denotes average per capita membership in the labor movement 1900–1920. *Strike* denotes per capita participation in the 1909 general strike. *Left* and *Turnout* denote average vote share of left-wing parties and average turnout in national elections 1911–1921. *Welfare 1918* and *Welf. 1919* denote per capita expenditures on welfare in 1918 and 1919, respectively. *Repr. Dem. 1919* denotes whether a municipality had voluntarily adopted representative rather than direct democracy in 1919. All regressions control for the log of the population at baseline. Additional control variables are growing season frost shocks 1864-1867, log distance to nearest emigration port, nearest town, nearest trade port, nearest weather station and Stockholm, log area, latitude, longitude, as well as an urban indicator and a set of indicators for high soil quality for the production of barley, oats, wheat, dairy and timber. *Shocks* × *Market Access* includes the interaction between growing season frost shocks and the log distance to the nearest town and trade port, respectively. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.