

Mass Emigration and Political Change: Evidence from Historical Swedish Elections

NEKN01: Master Essay I*

Aron Strandberg
Department of Economics
Lund University

January, 2018

Abstract

Migrants do not only affect the societies in which they arrive. When people leave in large numbers, their absence will also have indirect consequences for the societies from which they left. Between 1860 and 1930, 1.4 million Swedes emigrated abroad, most of them settling in the United States. In this essay I look at how this historic migration episode, in which a quarter of the population left the country, affected the political outcomes in Sweden.

I link emigration records, election data and population censuses for 2363 municipalities observed over 8 general elections between 1911 to 1928. I show that municipalities with more emigration saw larger relative gains for left-wing parties in subsequent elections.

Looking at migrant selection, I find evidence that municipalities with more emigration turned relatively more collectivistic, lending some support to the hypothesis that part of the left-wing gain can be explained as a consequence of ideological selection of emigrants.

Supervisor: Kaveh Majlesi

Keywords: Emigration, Political Change, Individualism

*Presented on January 26, 2018. Last revised: January 31, 2018.

Contents

1	Introduction	2
2	Background and theory	3
2.1	Emigration history	3
2.2	Political development	6
2.3	Literature and theoretical approach	6
3	Data	9
3.1	Election data	9
3.2	Emigration data	10
3.3	Other data	12
4	Empirical specification	13
5	Results	16
5.1	Emigration and voting	16
5.2	Emigration and names	20
5.3	The missing link	20
6	Concluding remarks	21
	Appendix	25

1 Introduction

The Age of Mass Migration was a historically unprecedented episode of long-distance migration. Between 1850 to 1920, around 55 million Europeans left their home countries for the New World. Of these, 30 million went to the United States (Abramitzky 2015). Some European countries saw a particularly dramatic outflow of people. Around 1.4 million people Swedes, or a quarter of the population, emigrated between 1860 and 1930. The overwhelming majority headed to the United States, with a smaller fraction going to places such as Canada, Australia and South America. On a per-capita basis, the Swedish emigration rate was the third highest in Europe (Barton 1994).

The arrival of immigrants shapes the societies in which they settle, both economically and culturally. Many locations in the United States today have an identity rooted in the diasporas that arrived during the mass migration era, connecting them to places such as Scandinavia, Italy and Ireland. However, this is not the only effect of their journey. When people leave in large numbers, their absence will indirectly also have consequences for the societies from which they came.

In this essay, I focus on such one aspect in particular: the effect of migration on the election outcomes in the sending country. I construct a dataset linking emigration records, election data and population censuses for 2,362 municipalities or municipality groups observed over the 8 general elections taking place between 1911 to 1928, the final years of the Swedish emigration period. By setting the 1911 election outcome as the initial benchmark, my regression analysis estimates whether the changes in left-right vote shares in the subsequent elections can be linked to the cumulative emigration outflow preceding them. I show that municipalities with a larger emigration rate saw larger relative gains for left-wing parties.

After establishing this relationship, I take a closer look at the mechanisms that can explain it. I look in particular at migrant selection, and investigate its role in bringing about the observed political change. Based on a strategy from a study of individualism at the

United States frontier and its translation into support for the Republican Party (Bazzi et al. 2017), I apply a name-based proxy for individualism and find that Swedish municipalities with more emigration saw a relative fall in this measure. This lends tentative support to the hypothesis that part of the electoral success of the Social Democrat-headed left in the early 20th century can be attributed to the ideological selection of emigrants.

The rest of the essay proceeds as follows. Section 2 provides a historical background of Swedish emigration and the political development of the era. It also provides a discussion of the literature and sets up the theoretical framework. Section 3 describes the data sources, the variable construction, and the matching process used to link the different sources together. Section 4 describes the empirical strategy. Section 5 presents the results. Section 6 concludes the essay.

2 Background and theory

2.1 Emigration history

When explaining the causes that set the Swedish mass emigration in motion, the historical literature usually frames it in terms of the different *push* factors that encouraged people to leave, and the *pull* factors that attracted them to the United States.

The first large emigration wave took place in the 1868-1873 period after a series of bad harvests had led to famine in Sweden. In addition to this, important underlying push factors for the early waves was the religious intolerance at home, dissatisfaction with the rigid class system, overpopulation, low wages, and the lack of new arable farmland (Clemensson 1996).

The major pull factor for the early waves of emigrants was the promise of cheap land at the frontier in the upper Midwest, made available for settlers through Lincoln's 1862 Homestead Act (Barton 1994). The next emigration wave between 1879-1993 was similar to the first one. As the time and cost of traversing the Atlantic Ocean fell rapidly with steamships replacing sailing as the main mode of transportation, emigration became a viable option

for an increasing amount of people. After the ball was set rolling by the initial pioneers, the flow of emigrants became self-perpetuating as the stories of success in newspapers and letters home from friends and families provided a significant pull from the United States and encouraged more people to follow in their footsteps.

In the mid-1890s, the farmers were followed by waves of workers who moved to the American cities that grew rapidly during the Second Industrial Revolution. For Swedes, Chicago became a major destination. In 1900, the estimated Swedish population of Chicago was 100,000, making it more numerous than Malmö and Gothenburg, and second only to Stockholm, in terms of its Swedish population size (Pehrson 2014).

Although emigrating to the United States brought benefit to the migrants themselves, its scale raised concern at home about the implications for the future of Sweden. In the contemporary public debate as well as in academia, views were split on whether it was harmful or benign. At an economics conference in 1881, Knut Wicksell tells about one side of economists describing the exodus as “an aberration and a national disease that must be extinguished”¹ (Wicksell 1882). Wicksell himself aligned with the other side of the debate. He argued that the surge in emigration was a symptom of the rapid population growth and found it most likely, although with some reservation, that it would eventually balance itself out through the balancing of labor supply and demand.

When the large-scale emigration showed no sign of stopping at the turn of the century, a government commission was created with the purpose of figuring out solutions on how to stem the flow. When the results were first published in 1908, around 20 percent of all Swedes lived in the United States. The commission rejected the calls for clamping down on emigration with restrictive laws against it. It argued instead in favor of “bringing the best sides of America to Sweden”, which generally meant further economic and social reform that would remove the incentives to migrate (Sundbärg 1913, Barton 1994).

The era of mass migration was made possible by the United States policy of near-open

¹My translation

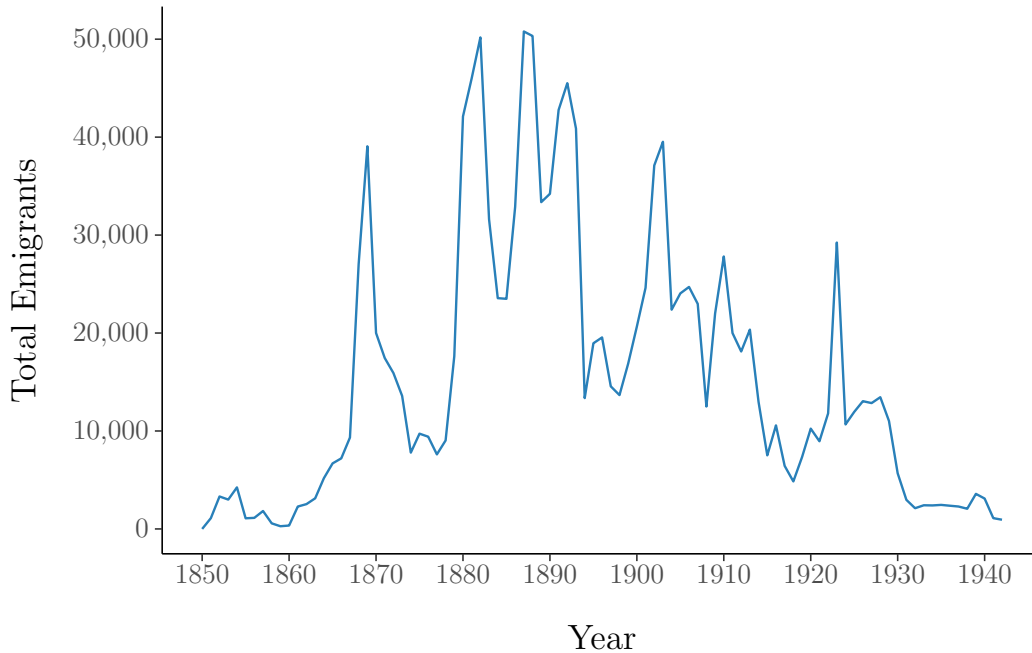


Figure 1: Annual emigration flow from Sweden during the age of mass migration. *Source: Befolkningsstatistik, SCB (Statistics Sweden)*

borders that allowed most Europeans to freely move into the country and naturalize as citizens². The first major restriction was a literacy test for entry in 1917, followed by the Immigration Act of 1921 that set quotas on the maximum number of people permitted to immigrate annually (Abramitzky & Boustan 2017). With this change of American policy, the Age of Mass Migration came to an end. However, Swedes were generally literate and the quotas were calculated in proportion to the size of the groups already living in the country. This allowed Swedish migration into the United States to continue, reaching a new peak in the 1920s before finally waning towards the end of the decade. After 1930 the migration flow in Sweden reversed, and it has nearly every year since been a country of net *immigration* (Statistics Sweden n.d.).

²For other groups, the rules were not as permissive. In 1882 and 1908 respectively, Chinese and Japanese immigration was been prohibited with targeted exclusion acts

2.2 Political development

Simultaneous to the mass migration era, Sweden also underwent political reform, starting in 1866 with the introduction of a modern bicameral parliament with a directly elected lower chamber. The income-restricted franchise was gradually expanded over the following decades as incomes grew and a larger share of men qualified. After major constitutional reform, the 1911 election was the first one with a proportional voting system, mandated party affiliations for candidates, and abolished income-related voting restrictions. Now, a majority of men were given the right to vote.³ From the 1921 election, voting rights were finally expanded to women, and thus a majority of the population (Esaiasson 1990).

During the same period, Sweden saw the rise of social democracy and the labor union movement. After its founding in 1889, the Social Democratic party went on a path of steadily increasing political influence, and would soon establish itself as Sweden's key political actor. It entered government for the first time in 1917 as the junior coalition partner. After heading several governments and serving as the main opposition throughout the 1920s, it got a definitive hold on power in 1932, beginning a period of uninterrupted Social Democratic governments that would last for 44 years.

2.3 Literature and theoretical approach

In one early study of Swedish emigration, Hatton (1995) estimates the emigration rate as a function of a number of macroeconomic variables. Hatton finds that important explanatory factors are the relative wage ratio between home and the destination, the relative employment rates, and the stock of previous emigrants.

The earliest micro-level studies on the period are done by Wegge (1998, 2002). Analysing a sample of German emigrants in the 1850s, Wegge finds that the establishment of emigrant networks was negatively correlated to the wealth of emigrants, suggesting that chain migra-

³The only major remaining qualification requirement for men—the completion of military service—was abolished in 1922.

tion networks served as an efficient substitute for self-financing, allowing poor but connected individuals to emigrate. Looking at migrant selection, she finds that middle-skilled German workers disproportionately emigrated to the United States compared to low- and high-skilled ones, suggesting that the former did not have the means to go and the latter were affluent enough to not finding it worthwhile.

In more recent years, the digitalization of censuses and historical records has opened up a new field research where individual and local decisions can be studied at a fine-grained level that simply was not possible before. Most noteworthy in this field is the work by Abramitzky, Boustan, and Erickson that has resulted in a series of papers covering different aspects of the age of mass migration. Abramitzky et al. (2012) link Norwegians between United States and Norwegian censuses and estimate the return to migration by comparing Norwegian emigrants with their brothers who remained in Norway. They find that the return to migration was relatively low compared to today's standard, and that urban emigrants were negatively selected from the skill distribution.

Considering its historical significance, the effect of the mass migration era on the sending countries has received little attention in research (Abramitzky & Boustan 2017). My essay is most closely related to Karadja & Prawitz (2016), which is the first to study the effects of emigration on political outcomes in Sweden. After instrumenting emigration with the series of frost shocks that triggered the initial emigration wave, they go on to show that instrumented emigration can explain a multitude of later political outcomes, including the increasing share of left-wing votes, higher rates of labor organization, strike participation, and voter turnout. Karadja and Prawitz consider the entire migration period before democracy and its effect on the political outcome that followed, measured as the average left-wing share in the elections between 1911 to 1921, and find a significant effect of the former on the latter.

To measure the relationship between the two trends, I choose a different approach, wherein I use the first available election outcome in 1911 as the pre-treatment starting point. From there, I look at how the emigration that followed affected voting, measured at

the deviation from the initial 1911 level, while controlling for observable and unobservable heterogeneity.

After establishing the link between emigration and left-wing electoral gains, there are mainly two groups of explanations for the underlying mechanism that causes it, focusing on whether the left-wing gain is the result of *individual* change or *demographic* change.

Representative of the first explanation is the hypothesis that the outside option of emigration strengthened the bargaining power of the left-wing labor unions, vis-a-vis the elite. With the reduced labor supply or the threat of further emigration making the elite more susceptible to labor demands, this would then translate into stronger popular support for the Social Democratic Party (and other left-wing parties) which was closely connected to the labor union movement.

The alternative hypothesis is that emigrants disproportionately consisted of people that were more likely to vote for non left-wing parties. It is possible that these people were more attracted by American pull factors such as the promise of individual liberty or the higher return to entrepreneurialism. When these people left Sweden for the United States, the proportion of voter groups would then shift in favor of the left.

Karadja & Prawitz believe that the first explanation is the most likely, and that emigrants were not substantially different from the rest of the population on relevant characteristics. In the second part of my analysis, I give the alternative theory a new chance. For this part, I draw inspiration from Bazzi et al. (2017), who study the impact of the United States frontier experience on the political and social outcomes of the counties that historically was a part of it. Based on social psychology findings, the authors start by noting that individualistic-minded people are more prone to give their children infrequent names, whereas collectivists more often pick frequent names, reflecting their respective desire to stand out or fit in. This attitude is transmitted intergenerationally as it is reflected in the upbringing and family identity.

Using the frequencies of different names as a proxy for individualism in society, the

measurement has been shown to be strongly correlated with other proxies such as the relative use of collective pronouns (we, us) versus singular ones (I, me), and the well-used Hofstede index. Individualism in this context, as contrasted to collectivism, should be thought of as traits related to individual independence, self-reliance and self-interest.

Making use of this relationship, Bazzi et al. analyze the people settling at the United States frontier in the 19th century. The authors show that migrant selection disproportionately drew individualists to the frontier. The time length of a region’s frontier exposure in the past translates to individualism that persists even to this day. Counties with longer frontier experience are more strongly Republican, more individualist, and more opposed to redistribution and regulation. It is easy to see how such an attraction would have the inverse consequences for places with a high sending rate. In the sense that the United States itself constituted a “frontier” for Swedes with the characteristics described above, what we would observe in municipalities with large emigration is a shift *away* from individualism. After establishing the relationship between emigration and voting, this is what I investigate in the second part of the analysis.

3 Data

The data used in the empirical analysis comes from a variety of sources. This section describes the data sources, the matching process that combines them together, and the transformations used to bring forth the final variables used in the analysis.

3.1 Election data

Historical Swedish election data has been digitized by Berglund (1988) and is provided for research through the Swedish National Data Service. The election data is available at the municipality level and consists of the number eligible voters and votes for each party in the general elections to the lower chamber of the Swedish Riksdag. Although Riksdag

elections have been taking place since 1866, the 1911 election was the first one in Sweden with proportional voting, mandatory party affiliation for candidates, and a generally franchised male population. This makes it the first election with results that can be directly comparable to subsequent ones, and the natural starting point for analysis.

Eight elections took place between 1911 and 1928, consisting of both regularly scheduled ones and snap elections⁴. Of the 2,576 different municipalities in the election data, not every municipality existed every year, as many were the results of splits and mergers that took place throughout the period. To create time-consistent units of observation, I use a codebook by Öhngren (1977) that document all such administrative changes, and merge all municipalities with any shared border changes until I have groups of municipalities whose outer borders are unchanged throughout the entire period. I also merge neighboring municipalities that have identical names except from a separating suffix (such as *city* and *rural*) to avoid ambiguity when matching emigrants to municipalities. 7 percent of municipalities are merged as a consequence and the result is 2,362 municipality groupings which constitute the units of observation in the empirical analysis. For brevity, I refer to these as just municipalities in the rest of this essay.

The dependent variable *LeftDiff*. is calculated by first dividing the number of votes for left-wing parties by the total number of valid votes in each election, and then subtracting the initial 1911 result. The initial outcome is added separately as a control variable. I define left-wing parties as the large Social Democratic Party plus the smaller socialist and communist parties (whom varied in name and number between elections). The voter turnout is calculated by dividing the number of eligible voters with the total number of votes.

3.2 Emigration data

The Emiganten Populär dataset (Clemensson 2006) consists of the passenger list records of 1.4 million people embarking to North America from Swedish ports during the mass migra-

⁴Election years: 1911, 1914 (#1), 1914 (#2), 1917, 1920, 1921, 1924, 1928. All elections took place in September except the first 1914 election which took place in March.

tion era, including Swedes who traveled over Copenhagen and Hamburg. When emigrants were about to embark, they were registered by the police. These emigration lists were then confirmed by the passenger lines at departure. Each person provided a range of personal information, as well as their departure date, destination, home county, and place of origin (most often their parish). I make use the latter two to match as many individuals as possible to each of the 2,363 municipalities in the election dataset. This is done with an iterative matching process. First starting with matching the two records on county and exact municipality name⁵, the latter is gradually made more permissive by allowing misspellings, abbreviations, and other variations. Although many locations across Sweden share similar names, by constraining the search to within the provided county, most matches can be made without ambiguity.

In total, I am able to match 832,102 individuals to their home municipalities. This constitutes 94 percent of those that have provided a location string and a valid county, and 77 percent of those that are likely to be from Sweden (Those remaining after subtracting other nationalities and non-residents; mainly Swedes already living in America). By visual inspection of individuals that can not be matched, the most common reason for non-matching is that many have simply provided “no location”, or something too unspecific such as their county name, or just “Sweden”. Some undercounting of emigrants is therefore unavoidable.

Apart from outbound passenger lists, an alternative source of individual emigration statistics is the records created by the parish churches as people emigrated. Karadja & Prawitz (2016) reviews both sources and show that the numbers on average are corresponding when controlling for gaps in the latter dataset. The independence of the two sources indicates that the records have a good degree of reliability. Even still, when summing up the matched tally of each county and comparing them with the aggregate numbers from the statistical yearbooks, it is clear that some counties are more affected than others by emigrant under-

⁵In this period, non-urban municipalities and parishes were largely corresponding in name and area, making the matching straight-forward. A complimentary list of parish names is added from Andrae (1998). The municipality keys to remaining (city) parishes are inputted manually.

counting⁶. In the main presented results, county fixed effects are included to take this into account. As a part of ensuring robustness of the results, I also rerun all the regressions with counties sequentially omitted to make sure those with implied undercounting are not significantly biasing the results.

The main independent variable, *Log Cumulative Emigration*, is defined as the log cumulative emigration from the 1911 election up to the time of the observation. Past emigration highly predicts future emigration. To take this into account, a control variable is included that sums up the total emigration from 1870 up to the 1911 election. In the second analysis, concerning individualism, I measure all the emigration variables and available controls according to the same principles as above, but instead starting in 1880 and accumulating up to the three years of observation, the last of which is 1910. To deal with the log transformation problem in observations without any observed emigration, I study the outcome both when omitting them and when adding one extra emigrated person per municipality. As the estimates remain in line with either method, I use the latter in the presented results.

3.3 Other data

The North Atlantic Population Project (Minnesota Population Center 2017) maintains a database with four decennial Swedish censuses between 1880 and 1910. Each census is full-count and contains about 5 million individuals, the total population at the time. I summarize them to calculate the mean age, sex ratio, urbanization, and total population of each municipality. Almost all parishes can be immediately name-matched to their corresponding municipalities in the election dataset. I link the remaining few manually.

The main measures used in the individualism analysis come from this census data. Within each decadal cohort and gender, I calculate whether each person has a frequent name according to different thresholds such as *top 5* and *top 20*, taking differences in spelling into account. Bazzi et al. (2017) use even larger thresholds for American names, up to *top 100*.

⁶See Appendix table 5.

In Sweden, the name variation was much smaller. In the average municipality in 1880, 74 percent of people had one of the 20 most common first names of their gender and cohort. In one municipality, 91 percent of people had first names ranked in the *top 20*. I explore some alternative measures, ranking names within regions and counties instead of country-wide popularity. Due to their large similarity, I decide that the country-wide measurement will suffice as a proxy for individualism. Each municipality has the share of frequent names observed in four years: 1880, 1890, 1900 and 1910. For the latter three years, the change in this variable compared to 1880 is then calculated.

Gothenburg and Malmö were the main ports of embarkation. From there, people were first usually escorted to Liverpool, where they boarded an ocean liner that took them to the United States. Of the people emigrating from Sweden during the mass migration era, 83 percent departed from Gothenburg and 13 percent departed from Malmö. The remaining fraction departed from a number of smaller ports, such as Stockholm, Norrköping and Kalmar (Clemensson 2006; my calculations). Following Karadja & Prawitz (2016), I use the distance from municipalities to ports of embarkation as a proxy for remoteness. Historical administrative shapefiles are provided by Swedish National Archives (2016). Using these, I calculate the centroid of each municipality and from them the straight line distance to either Gothenburg or Malmö, whichever is closest. One extra kilometer is added to avoid the log transformation problem for the two cities themselves.

Summary statistics of the calculated data are presented in table 1 and 2. Table 1 summarizes the data that is used in the analysis of emigration and election outcomes. The following table 2 summarizes the data that is used in the analysis of emigration and individualism.

4 Empirical specification

The goal of the first analysis is to estimate whether changes in the election outcome can be explained by the number of emigrants that left in the period preceding the election. To do

Table 1: Election Analysis - Summary Statistics

	Mean	Std.Dev.	Obs	Min	Max
Year	1919.71	4.80	16534	1914.00	1928.00
Left Diff.	0.11	0.15	16396	-0.70	0.86
Left Share	0.28	0.21	16499	0.00	0.94
1911 Left Share	0.16	0.20	16422	0.00	0.90
1910 Urban Share	0.03	0.17	16534	0.00	1.00
1910 Mean Age	30.78	2.39	16534	22.27	39.59
1910 Male Share	0.50	0.02	16534	0.41	0.60
1910 Population	2369.76	10007.08	16534	122.00	404605.00
1911 Voter Turnout	0.58	0.13	16513	0.04	0.90
1910 Union Share	0.00	0.02	16534	0.00	0.31
County	12.09	5.87	16534	1.00	25.00
1910 Top 10 Name Share	0.52	0.08	16534	0.28	0.72
1910 Top 20 Name Share	0.68	0.06	16534	0.44	0.84
Cumul. Emigration	25.94	106.23	16534	1.00	8617.00
Port Distance	223.69	191.30	16534	1.00	1236.68
1870-1911 Emigration	278.04	647.43	16534	1.00	20365.00
1890-1911 Emigration	147.06	392.95	16534	1.00	13097.00

Table 2: Name Analysis - Summary Statistics

	Mean	Std.Dev.	Obs	Min	Max
Year	1900.02	8.17	7058	1890.00	1910.00
1880 Top 3 Name Share	0.33	0.09	6984	0.08	0.59
Top 3 Diff	-0.02	0.03	6975	-0.22	0.11
1880 Top 5 Name Share	0.42	0.08	6984	0.18	0.67
Top 5 Diff	-0.03	0.03	6975	-0.23	0.10
1880 Top 10 Name Share	0.58	0.10	6984	0.30	0.79
Top 10 Diff	-0.04	0.04	6975	-0.30	0.12
1880 Top 20 Name Share	0.74	0.08	6984	0.46	0.91
Top 20 Diff	-0.04	0.04	6975	-0.37	0.12
1880 Share Male	0.49	0.02	6984	0.41	0.58
1880 Mean Age	28.57	1.64	6984	22.83	35.44
Port Distance	222.67	191.29	7058	0.00	1235.68
County	12.08	5.87	7058	1.00	25.00
Cumul. Emigration	182.57	453.89	7058	0.00	19787.00
1880 Population	1932.66	4206.86	6984	52.00	153090.00

this, I use the following specification:

$$LeftDiff_{mct} = \beta Emigration_{mct} + \phi_c + \gamma_t + \mathbf{X}'_{\mathbf{m}}\beta_{\mathbf{x}} + \varepsilon_{mct} \quad (1)$$

The outcome variable $LeftDiff_{mct}$ is the change of the share of votes for left wing parties in municipality m of county c at election t , as compared to the 1911 election. In other words, this variable measures a municipality's *relative* political change, compared to its starting point.

The main independent variable $Emigration_{mct}$ is the log cumulative number of emigrants since the 1911 election up until election t . Its estimated coefficient, β , gives us the level-log estimate of how accumulated emigration affects the vote outcome.

ϕ_c is a fixed effect for each of the 25 counties and γ_t a fixed effect for each of the 7 observed elections. $\mathbf{X}'_{\mathbf{m}}$ is a vector of municipality controls, all of which are fixed at their initial values in 1910 or 1911. The inclusion of the log 1910 population in $\mathbf{X}'_{\mathbf{m}}$ scales the level of emigrants to the initial size of the municipality. ε_{mct} is the error term.

The main threat to a causal interpretation of the result is the existence of omitted variables that biases the estimates. Although the risk of omitted variable bias can never be completely eliminated, a number of steps can be made to mitigate their potential influence, allowing for a more causal interpretation of the result. After establishing a significant relationship between emigration and the outcome, I progressively expand the regressions with a larger set of control variables to take into account the initial conditions of each municipality. To minimize the risk of unobserved spatial and temporal heterogeneity biasing the results, all regressions include both county fixed effects and election fixed effects. The inclusion of fixed effects estimators is standard practice in economic analysis and greatly reduce the risk of influential omissions, although at the expense of losing potentially useful signal.

In the second analysis, concerning changes in individualism, I uses a very similar specification to the first one. Instead of 1911-1928, the observed period is now 1880-1910, corresponding to the available census data. The year 1880 is used as the initial (and omitted)

year that the other variables are calculated relative to. To measure the effect of emigration on individualism, I use the following specification:

$$TopNameDiff_{mct} = \beta Emigration_{mct} + \phi_c + \gamma_t + \mathbf{X}'_m \beta_{\mathbf{X}} + \varepsilon_{mct} \quad (2)$$

The outcome variable $TopNameDiff_{mct}$ is the change of the share of people in municipality m at year t with a frequent name compared to 1880. As a proxy variable, it therefore measures how a municipality gets less or more individualistic, *relative* to its starting position. As mentioned earlier, the popularity of each name is measured within each decadal cohort and gender. I vary the threshold of what constitutes a common name when running the regressions. In the results, I present regressions setting it at top 20, top 10, and top 5.

Just as in equation 1, ϕ_c is a fixed effect for each of the 25 counties. Here, γ_t a fixed effect for each of the three observed years.

The only other differences of this specification compared to (1) is that $Emigration_{mct}$ accumulates from 1880 up to the observed year, and that the covariates included in \mathbf{X}'_m are fixed at their 1880 levels. The vector of controls also includes the initial proportion of frequent names for each threshold.

The main coefficient of interest, β , is then an estimation of how the accumulated emigration affects the proportion of names in the municipality.

5 Results

5.1 Emigration and voting

Five different estimates of equation 1 are presented in table 3 below. From left to right, it starts with a baseline estimation in column (1) and is progressively expanded until it includes the maximum number of covariates in column (5). All regressions include county fixed effects, election fixed effects, and use robust standard errors.

In all columns, the relationship between emigration and the vote outcome is in level-log form. The interpretation of such a relationship is that a 1 percent change in cumulative emigration results is associated with an estimated β percentage point change in the outcome.

The baseline column (1) is included as a comparison benchmark. By not controlling for the initial election outcome and past emigration—two vital factors—it is clear that it does not produce well-fitting results. In all columns (2)-(5) we see a small but significant positive effect of emigration on the left-wing vote share outcome. To take an example, the point estimate in column (2) of a ten percent increase in emigration is a predicted 0.054 percentage point increase in the left-wing vote share ⁷. A doubling of emigration in the same estimation associated with 0.394 percentage point shift in favor of the left.

As a larger set of covariates are introduced, the emigration point estimates and their significance levels falls somewhat as expected, but remain in line with the previous columns. Initial population size is positively associated with left-wing gains. Both the initial left-wing share and earlier emigration are both strongly associated with a decline in the left-wing vote. This should likely be interpreted as a consequence of regression toward the mean. If the left-wing vote share was very high to begin with, there will be more room for decline than for any further growth. Initial emigration is negatively associated with the outcome, possibly for the same reason. If a large past emigration led to a high initial left-wing share, the difference could then tend toward the mean as a result.

There is a positive relationship between left-wing voting and the initial share of men, the distance to the nearest port, and initial union membership per capita. Initial mean age is negatively associated with the outcome. I find no predictive value in the initial urban share variable, the voter turnout, or the initial share of people with a frequent first name.

⁷With the level-log model—used in all presented regressions—the effect of a 10 percent increase in emigration is obtained by multiplying the estimated coefficient by $\ln(1.1)$

Table 3: Emigration vs. Election Results

	(1)	(2)	(3)	(4)	(5)
	Left Diff.	Left Diff.	Left Diff.	Left Diff.	Left Diff.
Log Cumul. Emig.	-0.00176 (0.00121)	0.00569*** (0.00139)	0.00344** (0.00133)	0.00354** (0.00133)	0.00321* (0.00133)
1910 Log Population	0.0195*** (0.00189)	0.0388*** (0.00179)	0.0251*** (0.00181)	0.0250*** (0.00186)	0.0238*** (0.00187)
1911 Left Share		-0.353*** (0.00635)	-0.443*** (0.00701)	-0.442*** (0.00718)	-0.448*** (0.00735)
1870-1911 Log Emig.		-0.00949*** (0.00127)	-0.00386** (0.00121)	-0.00407*** (0.00122)	-0.00390** (0.00122)
1910 Male Share			0.230*** (0.0610)	0.202*** (0.0612)	0.223*** (0.0611)
1910 Urban Share			0.0106 (0.00679)	0.0105 (0.00681)	0.00512 (0.00686)
1910 Mean Age			-0.0197*** (0.000615)	-0.0197*** (0.000620)	-0.0194*** (0.000623)
1911 Voter Turnout				0.00190 (0.00900)	-0.00225 (0.00902)
Log Port Dist.				0.0176*** (0.00313)	0.0184*** (0.00316)
1910 Union Share					0.379*** (0.0662)
1910 Top 10 Name Share					0.00326 (0.0235)
Constant	-0.329*** (0.0240)	-0.368*** (0.0218)	0.292*** (0.0447)	0.201*** (0.0501)	0.183*** (0.0506)
Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	16396	16156	16156	16156	16156
Adj. R-sq	0.208	0.347	0.402	0.404	0.405

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Emigration vs. Name Frequency Results

	(1)	(2)	(3)	(4)
	Top 20 Diff.	Top 20 Diff.	Top 10 Diff.	Top 5 Diff.
Log Cumul. Emig.	0.000915** (0.000341)	0.00136*** (0.000327)	0.00123*** (0.000317)	0.00103*** (0.000281)
1880 Log Population	-0.00115 (0.000662)	-0.00156* (0.000612)	-0.000272 (0.000601)	0.000745 (0.000539)
1890 Dummy	0 (.)	0 (.)	0 (.)	0 (.)
1900 Dummy	-0.0208*** (0.000799)	-0.0210*** (0.000756)	-0.0213*** (0.000767)	-0.0138*** (0.000681)
1910 Dummy	-0.0451*** (0.000900)	-0.0455*** (0.000823)	-0.0434*** (0.000824)	-0.0318*** (0.000752)
1880 Top N Share		-0.254*** (0.00763)	-0.258*** (0.00725)	-0.285*** (0.00686)
Log Port Dist.		0.000257 (0.00126)	0.000732 (0.00116)	0.00339*** (0.000942)
1880 Share Male		0.0297 (0.0227)	0.0729** (0.0226)	0.0753*** (0.0207)
1880 Mean Age		0.0000264 (0.000238)	0.0000290 (0.000238)	0.000214 (0.000209)
Constant	0.00315 (0.00768)	0.158*** (0.0182)	0.0770*** (0.0170)	0.0140 (0.0154)
Fixed Effects	Yes	Yes	Yes	Yes
Observations	6975	6975	6975	6975
Adj. R-sq	0.394	0.499	0.571	0.455

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.2 Emigration and names

The results in table 4 shows the relationship between emigration and the change in popular name shares. In the table, I present one baseline regression in column (1) and three regression with full controls in columns (2)-(4), all with different thresholds in the dependent variable to illustrate the variation. All regressions include county fixed effects, election fixed effects, and use robust standard errors. Just like in the previous results, the relationship between emigration and the outcome is in level-log form.

When emigration increases, the result is an increase in the share of remaining residents with a frequent name. It should be noted, as made evident by the displayed dummy variables, that the general trend over the years was toward more name variation. To the extent that the relationship holds and more emigration results in less individualism, it is a *relative* one compared to other municipalities. The point estimates of the emigration effect are very small, ranging from 0.000915 to 0.00136. For example, according to the largest point estimate in column (2), a quadrupling of emigration would result in a 0.189 percentage point shift toward top 20 names in the municipality.

The initial name share has a strong negative impact on the dependent variable. This is almost certainly due to the regression to the mean effect that I also observed in the equivalent variables in table 3.

Generally, the significance of the effects increase as the threshold for the outcome variable increases. Initial population size is only weakly associated with the outcome in one of the regressions. I can find no effect of the initial mean age on the outcome. The remaining covariates that control for the share of males and distance to ports are positive and significant only with higher name frequency thresholds.

5.3 The missing link

Taken individually, the results of the regression analyses in table 3 and table 4 are in line with the hypothesis I laid out based on the findings of Bazzi et al. (2017). Emigration is

shown to both increase the left-wing vote share and the share of remaining people with a frequent name, the latter interpreted as a relative shift toward less individualism.

Before any conclusions about the nature of emigrant selection can be drawn based on these results, one missing link remains. This is to demonstrate that it is the decrease in individualism, as proxied by name shares, that is shifting the election outcomes in favor of the left. Swedish censuses after 1910 have not yet been made available for public use. Therefore, it is not yet possible to perform a similar analysis where changes in names are the treatment on election outcomes.

I perform a number of simple regressions of name frequency levels on later average election outcomes to see whether a preliminary correlative relationship can be established between the two ⁸. However, this does not produce any significant results in the relevant variables. The relevance of the second result should therefore be viewed tentatively. For now, it is mainly a suggestion of where to search for the mechanism that is causing the positive effect of emigration on the left-wing vote share.

6 Concluding remarks

The aim of this essay has been to explore the impact of emigration on election outcomes, as well as analyzing its underlying mechanism.

The main finding is that more emigration, as measured cumulatively, resulted in larger vote shares for left-wing parties in the elections that followed. The point estimates are small, with doubling or quadrupling of emigration levels only resulting in shifts that are fractions of a percentage point. Although the estimated effects are strongly significant, they are unlikely to have swayed any elections in the short timespans examined in this study, even considering the large sending rates of the period.

Secondly, I find that in municipalities with more emigration, the share of people with a frequent name increased. This could be related to migrant selection, and that more individ-

⁸See Appendix for examples

ualistic people were more likely to emigrate. A remaining step to connect these two parts of analysis is to directly establish the relationship between decreasing individualism through emigration and increasing left-wing voting. Findings in previous research have connected the two in United States, to where the overwhelming majority of Swedish emigrants went. For now, I can only show that they are indirectly connected through emigration.

I have taken a series of steps to underpin the causal inference of the results, including controlling for initial conditions and applying fixed effects estimators both temporally and at the county level. Even still, the risk of unobserved confounding factors can not be completely eliminated, meaning that the results should be taken with caution.

In recent years, there has been large increase in quantitative historical data being digitized and made available for research. In databases such as the the North Atlantic Population Project, more censuses are continuously being made available. Using distant historical data for economic analysis comes with both advantages and disadvantages, as made evident in this essay. Most interestingly, the historical setting allows us to study migration decisions in institutional settings that are currently not existing—in this case, a world of near-open borders.

The findings of my essay sheds some further light on the characteristics of Swedish emigration and its political consequences. The true nature of ideological selection and the explaining mechanism between emigration and election outcomes remains a topic for future research to discover.

References

- Abramitzky, R. (2015), 'Economics and the modern economic historian', *The Journal of Economic History* **75**(4), 1240–1251.
- Abramitzky, R. & Boustan, L. (2017), 'Immigration in american economic history', *Journal of Economic Literature* **55**(4), 1311–1345.
- Abramitzky, R., Boustan, L. P. & 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), 1832–56.
URL: <http://www.aeaweb.org/articles?id=10.1257/aer.102.5.1832>
- Andrae, C G Lundqvist, S. (1998), 'Folkrörelsearkivet 1881-1950 (the social movement archive)', *SND: Svensk Nationell Datatjänst* .
- Barton, H. (1994), *A Folk Divided: Homeland Swedes and Swedish Americans, 1840-1940*, Studia multiethnica Upsaliensia, Southern Illinois University Press.
URL: <https://books.google.se/books?id=9m1zOOpFTwwC>
- Bazzi, S., Fiszbein, M. & Gebresilasse, M. (2017), Frontier culture: The roots and persistence of "rugged individualism" in the united states, Working Paper 23997, National Bureau of Economic Research.
URL: <http://www.nber.org/papers/w23997>
- Berglund, S. (1988), 'Svenska valdata 1911–1944', *SND: Svensk Nationell Datatjänst* .
- Clemensson, P. (1996), 'Göteborgska källor till den stora utvandringen', *Dokumentet* **3**.
- Clemensson, P. (2006), 'Göteborgsemigranten: Databasen emigranten populär'.
- Esaiasson, P. (1990), *Svenska valkampanjer 1866-1988*.
- Hatton, T. J. (1995), 'A model of scandinavian emigration, 1870–1913', *European Economic Review* **39**(3), 557–564.
- Karadja, M. & Prawitz, E. (2016), 'Exit, voice, and political change: Evidence from swedish mass migration to the united states'.
- Minnesota Population Center (2017), *North Atlantic Population Project: Complete Count Microdata: Version 2.3 [dataset]*, University of Minnesota, Minneapolis, MN.
URL: <https://doi.org/10.18128/D040.V2.3>
- Öhngren, B. (1977), *GEOKOD: en kodlista för den administrativa indelningen i Sverige 1862-1951*, Acta Universitatis Upsaliensis.
- Pehrson, L. (2014), *Den nya staden: Utvandringen till Amerika II*, Albert Bonniers Förlag.
URL: <https://books.google.se/books?id=6cGKAwAAQBAJ>

- Statistics Sweden (n.d.), ‘Befolkningsstatistik’, <https://www.scb.se/hitta-statistik/sverige-i-siffror/manniskorna-i-sverige/in-och-utvandring/>. Accessed: 2018-01-09.
- Sundbärg, G. (1913), ‘Emigrationsutredningen: Betänkande’, *Norstedt & Söner* .
- Swedish National Archives (2016), ‘Historical gis maps 1569-1998’, <https://riksarkivet.se/psidata>. Accessed: 2018-01-08.
- Wegge, S. A. (1998), ‘Chain migration and information networks: Evidence from nineteenth-century hesse-cassel’, *The Journal of Economic History* **58**(4), 957–986.
- Wegge, S. A. (2002), ‘Occupational self-selection of european emigrants: Evidence from nineteenth-century hesse-cassel’, *European Review of Economic History* **6**(3), 365–394.
- Wicksell, K. (1882), *Om utvandringen, dess betydelse och orsaker. Föredrag, etc*, Albert Bonniers Förlag.

Appendix

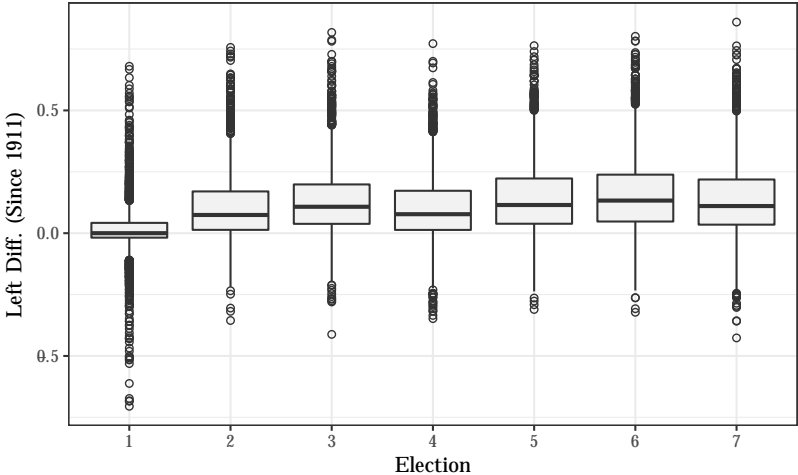


Figure 2: Left Diff. distribution by election

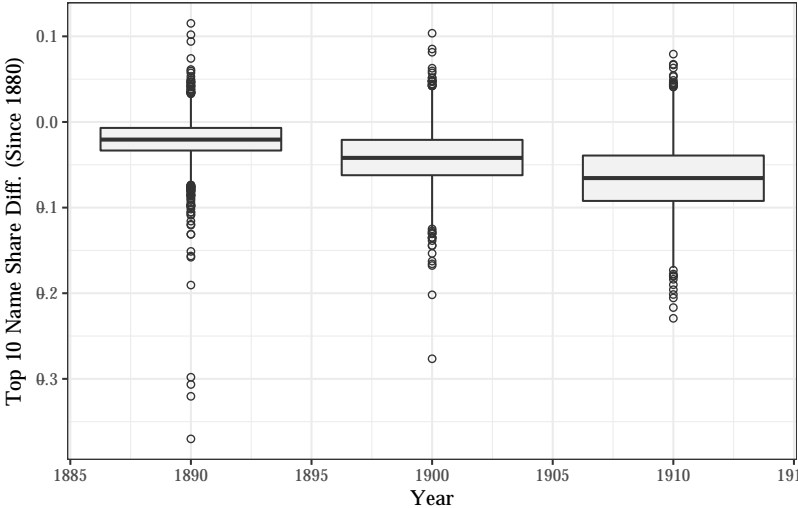


Figure 3: Top 10 Name Share Diff. distribution by year

Table 5: Average left-wing share 1911-1921

	(1)	(2)	(3)
	Av. Left Share	Av. Left Share	Av. Left Share
1910 Log Population	0.0353*** (0.00671)	0.0354*** (0.00675)	0.0355*** (0.00677)
1870-1911 Log Emig.	-0.0186 (0.0184)	-0.0185 (0.0185)	-0.0189 (0.0186)
1910 Top 3 Name Share	-0.104 (0.139)		
1910 Share Urban	-0.0213 (0.0147)	-0.0221 (0.0149)	-0.0237 (0.0149)
1910 Mean Age	-0.0300*** (0.00314)	-0.0300*** (0.00313)	-0.0298*** (0.00309)
Log Port Dist.	0.00974 (0.0167)	0.0102 (0.0168)	0.00986 (0.0168)
1910 Share Male	0.320 (0.177)	0.321 (0.176)	0.322 (0.174)
1910 Voter Turnout	0.135* (0.0532)	0.136* (0.0531)	0.136* (0.0533)
1910 Union Share	1.365** (0.418)	1.364** (0.418)	1.370** (0.417)
1910 Top 5 Name Share		-0.0934 (0.138)	
1910 Top 10 Name Share			-0.131 (0.140)
Constant	0.623** (0.203)	0.622** (0.202)	0.650** (0.209)
Fixed Effects	Yes	Yes	Yes
Observations	2325	2325	2325
Adj. R-sq	0.606	0.606	0.607

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

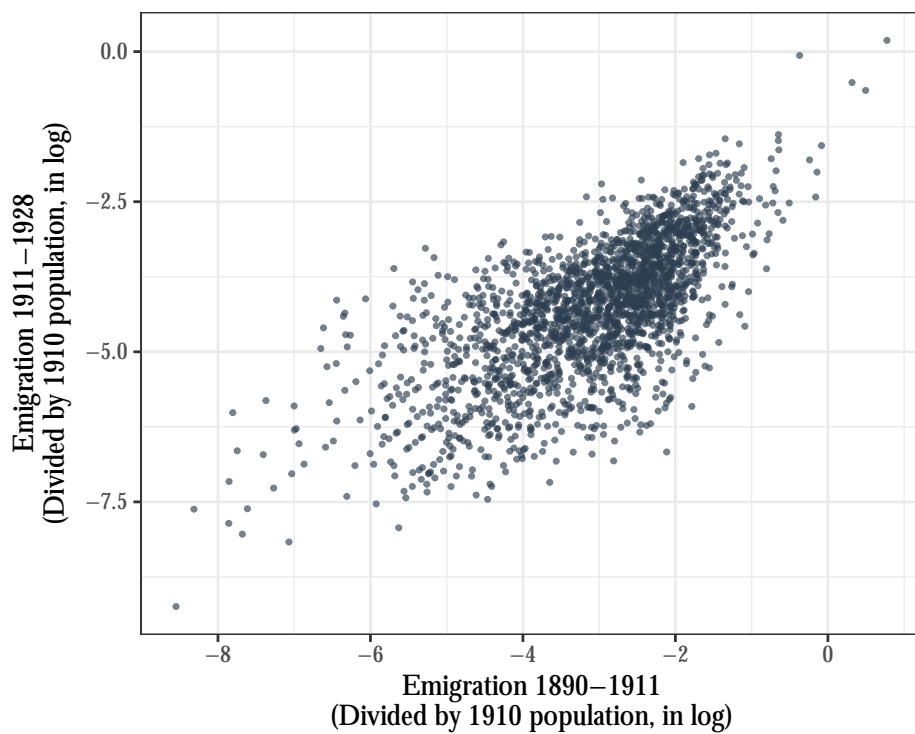


Figure 4: The correlation between emigration before and after the 1911 election. *Source: Clemensson 2006; my calculations)*

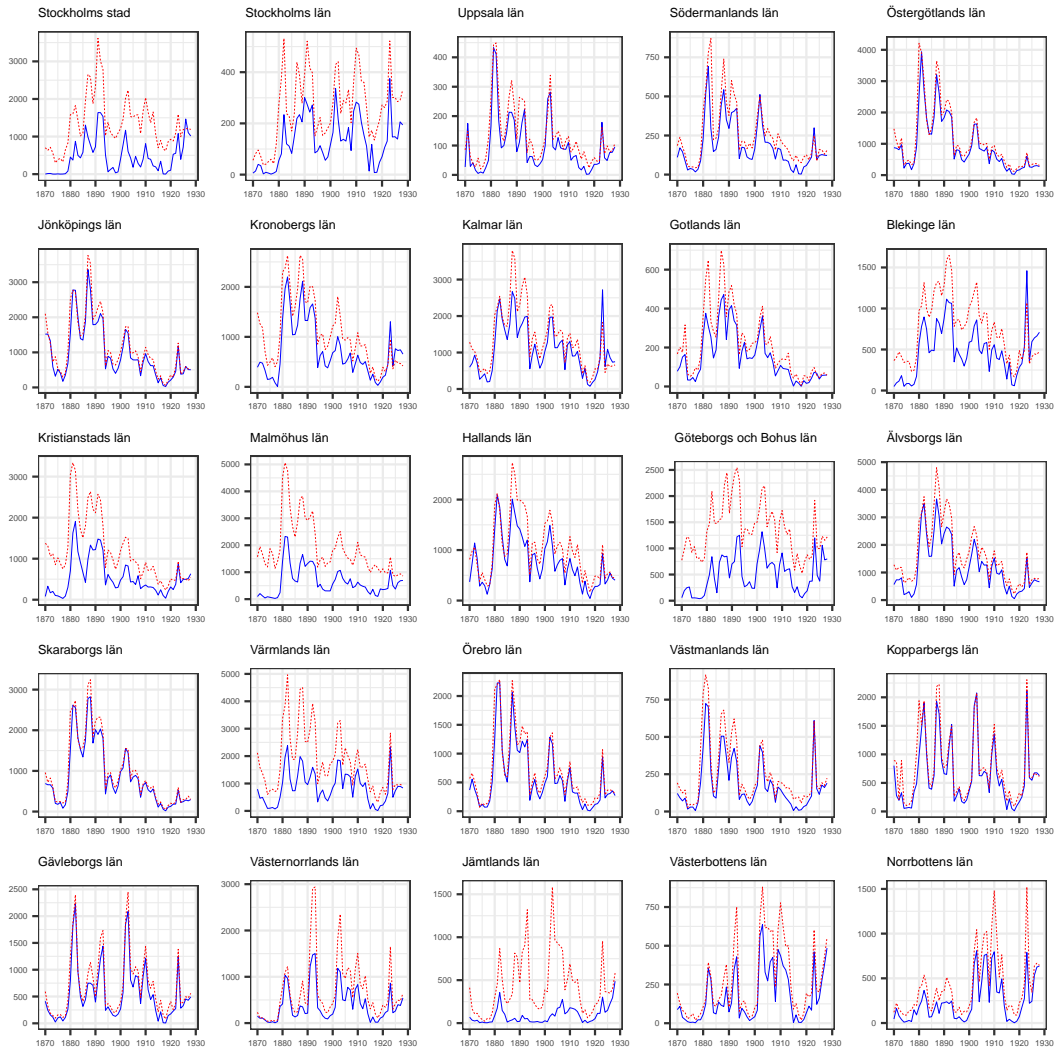


Figure 5: The number of matched emigrants per county-year (solid blue) vs. the total number according to aggregated SCB statistics.