European
Historical
Economics
Society

EHES Working Paper | No. 213 | April 2021

The Sleeping Giant Who Left for America: The Determinants and Impact of Danish Emigration During the Age of Mass Migration

Nina Boberg-Fazlić,
University of Southern Denmark

Markus Lampe,
WU Vienna University of Economics and Business,
CEPR

Paul Sharp,
University of Southern Denmark, CAGE, CEPR

EHES Working Paper | No. 213 | April 2021

The Sleeping Giant Who Left for America: The Determinants and Impact of Danish Emigration During the Age of Mass Migration*

Nina Boberg-Fazlić,
University of Southern Denmark

Markus Lampe,
WU Vienna University of Economics and Business, CEPR

Paul Sharp[1],
University of Southern Denmark, CAGE, CEPR

Abstract

What determines emigration, and what impact does it have on the sending country? We consider the case of Denmark between 1868 and 1908, when a large number of people left for America. A significant fraction of these were *tyender*, a servant-like occupational group that was heavily discriminated against at the time, and who saw little opportunities for advancement at home. We exploit the fact that the Danish agrarian reforms between 1784 and 1807 had differential impacts on this class of landless laborers around the country, and use detailed parish-level data – police protocols of emigrants; population censuses and land registers – to show that areas with a more unequal distribution of land witnessed larger emigration. We then use income tax data, finding evidence of a positive income effect on the areas which saw most emigration.

JEL Codes: J15, N33, O15

Keywords: Agrarian reform, Denmark, emigration, landless laborers

[1]Corresponding Author: Paul Richard Sharp (pauls@sam.sdu.dk)

*We thank participants at the Annual Conference 2021 of the Economic History Society, the 6th Annual Meeting of the Danish Society for Economic and Social History and seminar participants at Aix-Marseille University and Martin Luther University of Halle-Wittenberg for helpful comments and discussions. Nina Boberg-Fazlić gratefully acknowledges funding from Independent Research Fund Denmark (grant no. 6109-00287B).

1 Introduction

"We have the power, but we don't all know it yet, but the "Tyende"-class is a sleeping giant which will soon wake"

- Carl Westergaard (1907)

Who leaves during periods of mass migration, and what is the impact on the sending country? We exploit the example of mass emigration from Denmark between 1868 and 1908 in order to shed light on this question. In the Danish case, a large share of those who left belonged to the class of *"tyender"* (singular: *tyende*), often translated as servants, but actually denoting anyone, both male and female, who lived in the house of their employer, and could thus for example be agricultural laborers or milkmaids, as well as domestic servants. This group was heavily discriminated against. The *Tyende* Law of 1854, introduced just five years after the otherwise liberal constitution of 1849 which ended a period of absolute monarchy, made *tyender* subject to the head of their household, prescribed heavier penalties for breach of contract for a *tyende* than for his or her master (including immediate redundancy), and gave their employer the right to freely dispose over their time, beyond that stated in their contract. It was only in 1915 that *tyender*, alongside women, gained full suffrage, and in 1921 the *Tyende* Law was replaced with an Employee Law which went some way to improving their rights and position in society. Thus, as Hvidt (1971) explains, improved social status, at least for those *tyender* who lived in the countryside, required property, and where this was not possible, it provided a potent motivation for many to leave for America to acquire land. He noted that this occurred to a greater extent in areas of better quality soil, although the reasons for this remained unclear to him.

In previous work (Boberg-Fazlic et al., 2020), we have demonstrated that a distinctive pattern of land inequality was determined in Denmark during a period of far-reaching agrarian reforms and enclosure between 1784 and 1807, and that this remained throughout the nineteenth century. Areas with better quality soil, as determined during the last Ice Age, witnessed higher rates of population growth, leading to greater land inequality in those areas, although this declined in the last decades of the century, and we speculated that this might have been due to differential levels of outmigration. Taking this as our starting point, and drawing on a variety of detailed parish-level historical sources, we first consider whether these areas did indeed witness large levels of emigration and whether *tyender* constituted a disproportionate share of this emigration. We then ask to what extent this impacted on living standards for those who stayed as measured by two income taxes imposed in 1870 and 1905, before and after the period of mass migration. We expect, and find evidence for, positive income effects of emigration.

We thus contribute to two interrelated questions: who decides to migrate (selection of migrants), and what was the impact on the sending country. The answer to the second depends on the first. The loss of skilled migrants, a "brain drain", might have deleterious effects. So-called negative selection of migrants, on the other hand, as we expect to find for the Danish case, might have the opposite impact.[1] Building on the Roy model, Borjas (1987) explains the conditions under which emigrants will be positively or negatively selected in a theoretical model. They will be positively selected from the source population if the rate of return to skills is higher in the host country than in the home country, and negatively selected if the opposite is the case. This also implies that migrants are more likely to be positively selected when income inequality in the home country is relatively lower than in the host country and vice versa.

The existing evidence on skill-selectivity is however mixed. Gould and Moav (2016) find that Israeli migrants to the US were positively selected, as would be predicted by the more compressed wage distribution in Israel. Chiquiar and Hanson (2005), on the other hand, find that Mexican migrants to the US come from the middle of the skill distribution, rather than the lower end as would be predicted by the model. Feliciano (2005) and Grogger and Hanson (2011) find that migrants are selected positively on educational attainment from almost every sending country in the world, even those with very high levels of income inequality.[2] From the perspective of economic historians, Carter and Sutch (1999) noted a consensus among economic historians that European emigrants to the US during the age of mass migration were positively selected, although this has been disputed in some later studies. For example, Wegge (1999, 2002, 2010) finds emigrants from Germany during the 1850's (i.e. when transportation costs were still very high) came from the middle of the skill distribution, and for Scandinavia, Abramitzky et al. (2012) find Norwegian emigrants during the age of mass migration to be negatively selected from urban areas. Somewhat relating to our work, Connor (2019) shows that landless were more likely to emigrate from Ireland and that this factor was more important than the individual's occupation itself. Sons of farmers and illiterate men, i.e. individuals who were more likely to emigrate, were concentrated in unequal communities. Thus, he finds evidence for the Roy model. In general, selection appears to weaken with the growth of networks abroad, due to the impact on lowering migration costs (Abramitzky et al., 2012; Knudsen, 2019; McKenzie and Rapoport, 2010).

Hatton and Williamson (1998) find that real wages in Ireland increased over the period 1851-1908 largely due to emigration. The magnitude of this effect depends on the assumptions made about the international mobility of capital. Begley et al. (2016) shows that this is especially true for agricultural labourers and domestic servants. There is also evidence for this from Sweden (see Hatton and Williamson, 1998; O'Rourke and Williamson, 1995), although Karadja and Prawitz (2019) argue that the wage increase in this case is due to increased bargain power, not because of diminished labour supply. Mishra (2007) looks at

[1] It is not necessarily bad for the receiving country, however. Boberg-Fazlic and Sharp (2020) have demonstrated that Danish settlements in the United States proved efficient transmitters of knowledge and technology from Denmark, once the latter began developing rapidly from the 1880s.

[2] This may be reconciled with borrowing constraints, such that the poorest residents of sending countries can no longer afford to move when transportation costs increase (McKenzie and Rapoport, 2010) or with a logarithmic rather than a linear utility function (Grogger and Hanson, 2011).

this for the case of Mexico and finds that workers gain but capital and land owners lose from emigration.

For rural Denmark, which is the focus of the present work, there have been relatively few studies. Stolz and Baten (2012) examine selectivity of migrants using anthropometric inequality measures and literacy, finding for Denmark that migrants were slightly negatively selected up to 1850 (when they were however very few in number), but that subsequently there is no clear sign of selection in terms of literacy.[3] Baines (1994) shows that emigration rates from Danish provinces were negatively correlated with income, and while this does not say anything about the selection of migrants from these areas, it does have the implication that income is a relevant control for our analysis below, and we proxy for this in a number of ways. Knudsen (2019), using evidence from first names, finds that Danish (and other Scandinavian) emigrants exhibited more individualistic traits than those who chose to stay in Denmark.[4]

The remainder of this paper proceeds as follows. The following section presents the historical background, Section 3 introduces our data and empirical strategy, and Section 4 considers the selection of migrants and determination of migration. Sections 5 and 6 present our empirical analysis, and Section 7 concludes.

2 Historical Background

Hvidt (1971) provides a comprehensive survey of Danish emigration based on data from police protocols which we discuss in detail below. Very few Danes left before 1866, just 14,000, the majority of them religious and political refugees, mostly Mormons and those fleeing Northern Schleswig after it was lost to Prussia in 1864. Others left from 1865 when a new constitution made Denmark less open with respect to religion (Furer, 1972, p.45). Emigration mostly took off from the late 1860s, however, with around 158,000 Danes leaving for the US between 1868 and 1900. We do not know how many returned, but Hvidt (1971) estimates that only around 10 percent did so. The greatest extent of Danish emigration was reached in 1882, when 11,000 Danes left in a single year, many of whom were seeking land and jobs in the interior areas of America (Furer, 1972, p.56).

In Denmark, the number and size of farms was constant, and thus access to land constrained due to institutional arrangements aimed at protecting family farms. Since the 1680s, legislation prohibited the incorporation of farms into estate demesnes, and over the eighteenth century minimum sizes for farms were defined below which farms were not allowed to fall through land sales, etc. These rules were made to ensure that families could make a living off the farms, an aim that was explicitly confirmed during the 1784-1807 agricultural

[3] We thank Joerg Baten for detailed results beyond those reported in the published version.

[4] Somewhat relating to this, Karadja and Prawitz (2019) argue that Swedish migrants were ideologically selected, with those voting right wing being more likely to move.

reforms. In the immediate post-reform period, in 1819, a decree defined an extension of the equivalent of 21 ha of good quality land (c. 4 Tdr Hartkorn[5] - a quality-adjusted measure of land) as the yardstick, and half of that as the lower bound. Such legislation was reconfirmed in different forms over the nineteenth century, although in the second half of the nineteenth century, after a new tax assessment, a new absolute lower boundary of 1 Tdr Hartkorn was established, which also marks the lower bound of the statistical category 'farm' in the agricultural censuses (Jensen, 1945).

In addition to access to land, there is some evidence that Danish farmers earned more money in the US than they did back home (Mackintosh, 1992, 1993). Hvidt (1971) argues that these higher wages presented one motivation for leaving, but that two other factors were also important. First, improvements in transportation technology meant that moving became more affordable and second, information flows via various channels including personal connections, advertising by shipping and railroad companies from the US, as well as by the US and State governments, and various organizations which aimed to promote this, made it more attractive and accessible. Figure 1 provides an overview of the migration flow from rural areas over time and figure 2 shows the spatial distribution of emigration across Denmark. In the left panel we show total emigration rates (share of population) and in the right panel the share of *tyender* in total emigration.

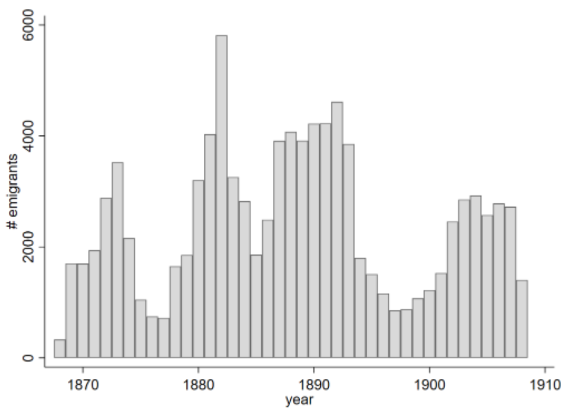
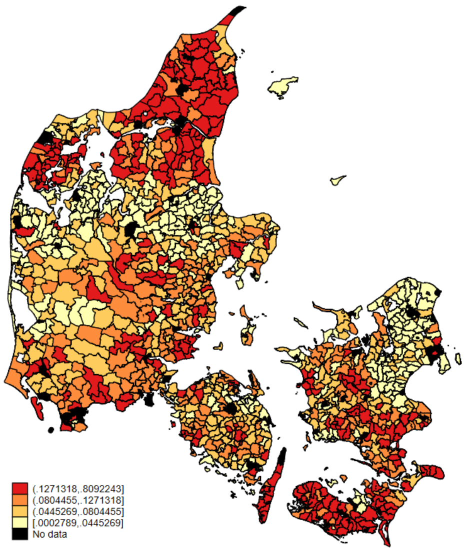


Figure 1: Number of emigrants from rural areas by year, 1868-1908. Source: Det Danske Udvandrerarkiv (2018)

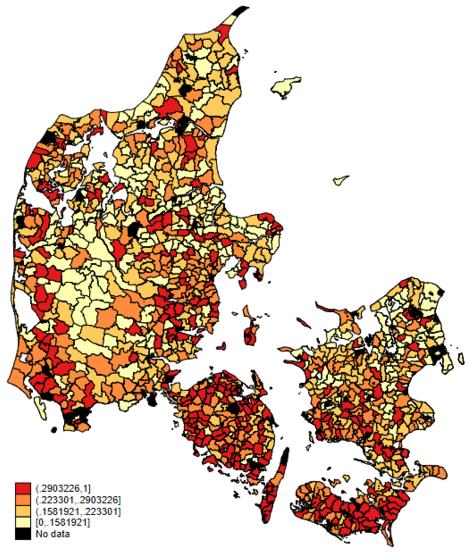
After a first smaller wave of emigration around 1870, numbers start to take off around 1880. This is especially true for emigration of *tyender* which comprises a significant share of rural emigration numbers between 1880 and 1895, as we present more evidence for below. Hvidt (1971)'s argument that emigration numbers increased after 1880 due to increasing incomes and declining transportation costs must have been particularly important for *tyender*, whose incomes would have been too low relative to the cost of emigration earlier. He discusses that two generations or so after the abolition of serfdom in 1800, contemporary observers suggested that the *tyende* class was becoming increasingly more politically aware in

[5] *Tdr Hartkorn* literally means barrels of hard grain.

the last decades of the nineteenth century. They desired upward social mobility, but in order to escape their position in society they needed to gain land. When this was not possible, they went to America. As noted above, *tyender* constituted an underclass of Danish society, and were discriminated against both socially and institutionally, and it took until 1921 for the status to be officially abolished. In fact, in the run up to the signing of the new constitution of 1915 which introduced full democracy to Denmark, there was much more political focus on giving votes to women than to the *tyender*, although pressure from campaigners such as Westergaard (quoted above) meant that they were permitted to vote for the municipality in 1908, and for parliament in 1915.



(a) Emigration rate



(b) Share of *tyender*

Figure 2: Total emigration by population 1860-1908 (left) and share of *tyender* in total emigration (right).

3 Data and Empirical Strategy

Our data is at the parish level and is taken from a variety of sources including Danish police protocols of emigrants, population censuses, land registers, and income tax records. We obtain the number of emigrants by counting their individual records as reported by the Danish police, digitized and maintained by the Danish Emigration Archives (*Det Danske Udvandrerarkiv*) and Aalborg City Archives (*Aalborg Stadsarkiv*) (Det Danske Udvandrerarkiv, 2018). As of 1868, every emigrant had to sign a contract with an emigration agent. These contracts had to be returned to the police to generate records for all emigrants. While the original contracts are lost, the police records are not. The data includes information on the name, birth date, date of emigration, occupation, birth place, last place of residence, travel destination, and the name of the ship the emigrant travelled on for 359,760 emigrants emigrating from (or via) Denmark (Copenhagen which was the only emigration port, although some also left via Hamburg in Germany) between 1868 and 1908. Of these 199,292 have a known last place of residence in Denmark. 141,411 emigrants travelled via Copenhagen but lived outside of Denmark (mostly Swedes). For the remaining 19,057 emigrants the place of residence could not uniquely be identified.

Our data on land inequality comes from our previous work (Boberg-Fazlic et al., 2020), where we digitized farm sizes and numbers on the parish levels from Danish land registers for the years 1682, 1834, 1850, 1860, 1873, 1885, and 1895. The data for the year 1682 is digitized from Pedersen (1928),[6] the year 1834 from *Commissionen for det statistiske Tabelværk i Danmark (1837)*, and the years 1850 to 1895 from *Statistiske Bureau (1852, 1864, 1877, 1888, 1896)*. The registers give the number of farms in different size categories (measured in *tdr hartkorn*) for every parish. Farms are divided into categories based on their size, but also based on ownership type, which could be owned, tenant or rented farms for example. Our measure of land inequality is thus not concerned with who actually owns the land, but who had access to its productive potential. This is usually referred to as 'operational unit inequality', and is more comparable to income inequality than wealth inequality. We refer to it simply as land inequality in the following. From these land registers, we calculated the Theil index on the parish level to measure land inequality. The Theil index exhibits several analytical advantages, which makes it preferable to the more widely used Gini coefficient for example. It satisfies the strong principle of transfers, such that a redistribution from one individual to a poorer one will lead to a decline in the Theil index proportional to the absolute distances between the individuals' incomes. This is especially desirable for our instrumental variables strategy, where we use the *changes* in the Theil index. Also, the Theil index ranks distributions unambiguously, such that two places with identical Theil indices will also have identical income distributions, which is not the case with the Gini coefficient. For more on this we refer to Boberg-Fazlic et al. (2020).

The data on land inequality only covers rural parishes and our proposed mechanism of land inequality to emigration can also only be present in rural areas. Our analysis is thus restricted to rural parishes. It is likely that those willing to emigrate would migrate to cities first and then take the journey to the U.S., for example, from there. Although the place of birth is recorded in the data it is much less complete than the last place of residence (it is only available for ca. 56,000 emigrants but out of these there could be more where we cannot uniquely identify the parish). Thus, in order to account for this type of migration we include the population weighted inverse distance of each parish to Danish market towns as a control variable. The market towns and their populations in the years 1801, 1840, 1860, 1870, 1880, 1890, 1901, and 1906 are taken from Statistiske Bureau (1880, 1906).[7]

Parish population numbers are taken from Danish censuses (Danish National Archives,

[6] Note that the title of the publication states 1688, but the data was collected in 1682. We therefore refer to this data as 1682.

[7] We have also calculated the inverse distance to all market towns weighted by the total emigration of the respective market town (instead of weighting by total population). Including this measure as a control variable does not alter the results.

1787, 1801, 1834, 1850; Statistiske Bureau, 1911), and we use several geographical characteristics of the parish as control variables. These include the parish area, distance to the nearest emigration port (either Copenhagen or Hamburg), distance to the coast as well as the longitude and latitude of the parish centroid. We calculate these using a shapefile of historical parish borders available from the Digital atlas over Denmark's historical administrative geography (*Digitalt atlas over Danmarks historisk-administrative geografi*, downloadable at: digdag.dk).

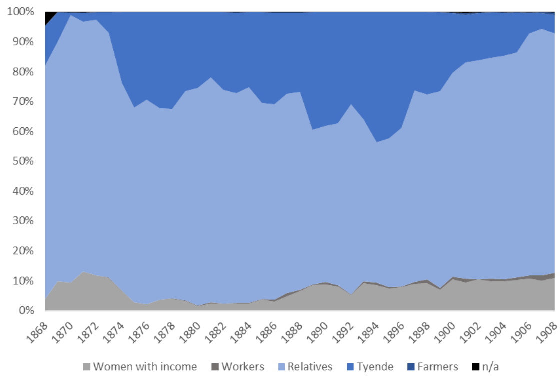
Finally, we combined all this with data from income tax records from the years 1870 and 1905 (Statistiske Bureau, 1873, 1905). Although, in principle, this data is also on the parish level, in many cases several (mainly two) parishes are reported together. We therefore aggregate all aforementioned data to the units in the tax records, which leaves us with 1,002 observational units (also termed parishes, although one "tax parish" usually included two church parishes). From the income tax records we digitize the number of taxpaying households, the amount of tax paid (in total) as well as total taxable income. The last two are measured in Danish crowns (*kroner*), where we convert numbers from the 1870 data given in the old currency *rigsdaler* into Danish crowns at the fixed ratio of 2 *kroner* to one *rigsdaler.* Due to Denmark's adherence to the gold standard, there was practically no inflation over this period, making the tax rates comparable over time, since there were no changes in tax liability rules over this period. The tax paid was calculated as 2.5% of taxable income less basic and child allowance. In rural areas (i.e. the areas included in the analysis) everyone with an income above 600 crowns had to pay income tax. On average 7.6% of the population paid income tax in 1870, increasing to 10.4% in 1905.

Naturally, between 1682 and 1895 there were some changes in parish borders. We convert all borders to those of 1682 by aggregating parishes that were split or consolidated later on. Market towns usually consisted of one parish for the town and one for the countryside. We are only interested in the countryside here, where the measure of land inequality can be applied to farms of different sizes. However, very often the parish of the market town and its rural part are reported together in some years and separately in other years. In these cases we exclude the rural parish of the market town in all years. We also do not include the island of Bornholm, as it is not covered by the land reports in all years. The same is true for the area of Southern Jutland, which was part of Denmark before the Second Schleswig war of 1864 and part of Prussia after this date. With these adjustments, we have data on 1,605 consistently reported parishes with stable borders across all years. Summary statistics for the variables used can be found in table A1 in the appendix.

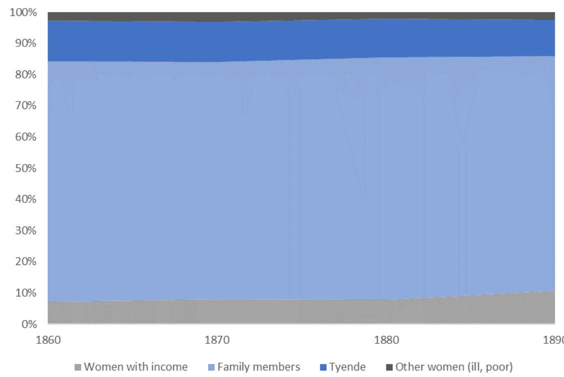
4 Explaining emigration patterns

4.1 Selection of emigrants

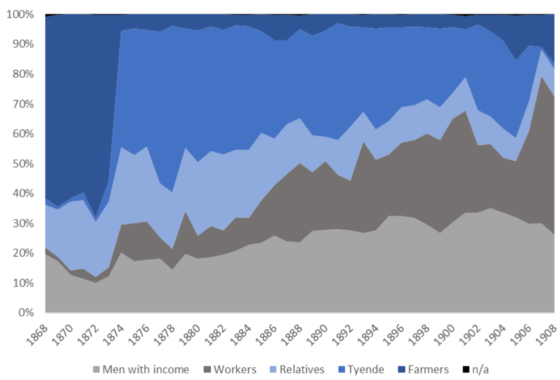
We compare the occupational structure of the emigrants to the total population as measured from census data. Using the summary tables from the census data does not allow for the division into farmers and workers. For the emigration data, we split emigrants into five different occupational groups: relatives, *tyender*, workers, farmers, and others. Relatives include occupational titles such as “wife”, “daughter”, “son”, or “mother in law”. *Tyender* include those specifically named “*tyende*” but also agricultural workers and crofters (*husmænd*) who were similarly lobbying for access to land (Solvang, 1985). It also includes the occupational title “*Karl*” (farmhand/servant). Workers are those titled worker or labourer (*arbejder*) without further specification. Occupational titles in the category farmer range from “farmer” (*bonde*) to “farm owner” (*gårdejer*). Others include anyone else: e.g. woodworkers, bakers, fishermen, gardeners, merchants, painters, dairy workers, masons, millers, shoemakers, tailors, smiths, carpenters.



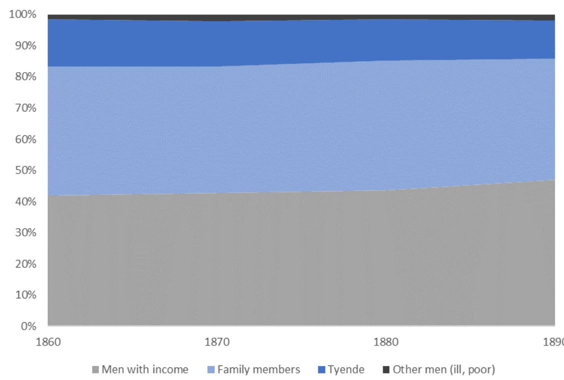
(a) Women, emigration data



(b) Women, census data



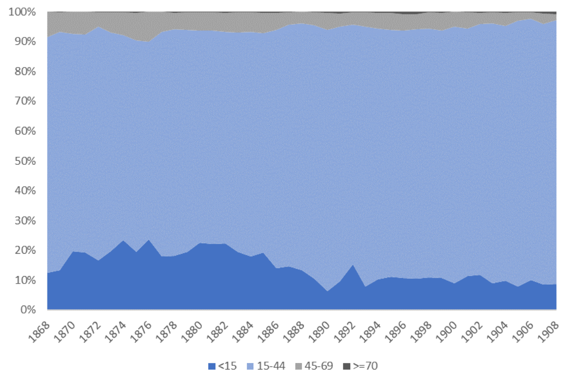
(c) Men, emigration data



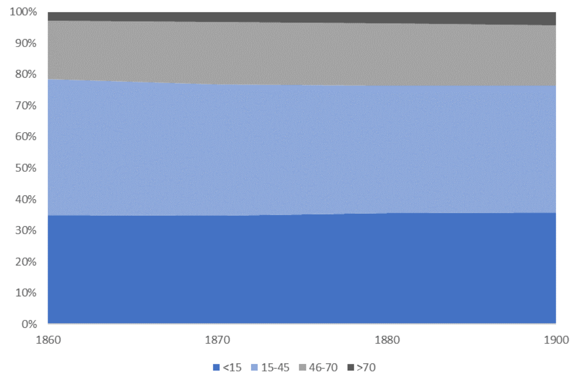
(d) Men, census data

Figure 3: *Tyender*, family members, and other income, men and women, comparing emigrants to total population.

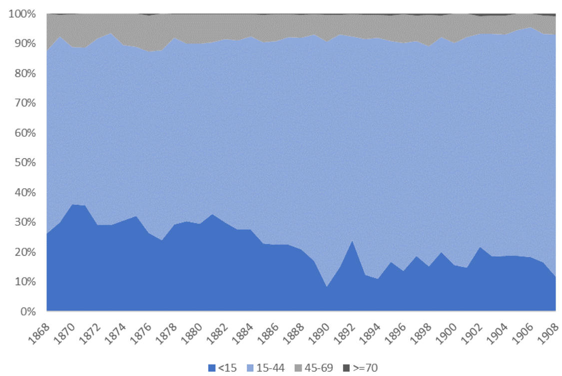
By far the largest share of women were relatives/family members, i.e. mainly wives and daughters. However, we also see a large overrepresentation of *tyender* in the emigration data amongst women. Amongst men, family members (i.e. under-age sons) are underrepresented. This is in line with the fact that young men (possibly couples) emigrated for economic reasons before starting a family. Again, *tyender* are overrepresented amongst emigrants and so are men with income. Here, it should be noted that the categories "workers" and "farmers" cannot clearly be allocated to either "*tyender*" or "men with income" in the census data. These occupations are self-reported, i.e. farmers include anyone specifying him- self as a farmer. This is likely to also include emigrants with the aspiration to own land and become a farmer in the destination country (Cohn, 1995; van Vugt, 1988). Also it is not not possible to know what emigrants categorized as "workers" actually did. As we only include rural areas in the analysis, however, these might well also be agricultural labourers. Taking into account that also these categories will partly be *tyender*, their overrepresentation becomes even larger.



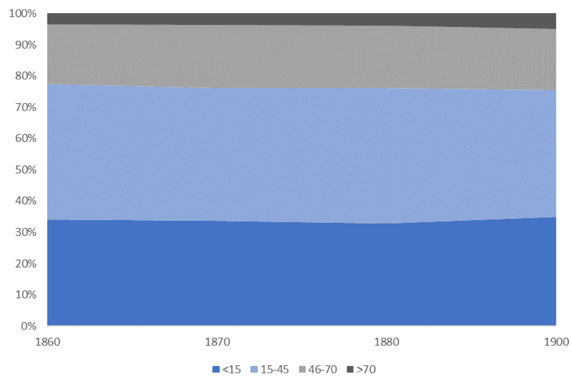
(a) Men, emigration data



(b) Men, census data



(c) Women, emigration data



(d) Women, census data

Figure 4: Age distributions, men and women, comparing emigrants to total population.

For both women and men the age group 15-44 years is highly overrepresented. Again, this confirms the idea that young individuals or couples emigrated for economic reasons in their prime working age.

4.2 OLS

To examine the impact of inequality on emigration, we start by estimating the following pooled OLS model for the years 1860 to 1895:

$$\ln(Emigration)_{p,t} = \alpha + \beta \times Theil_{p,t-1} + X_p'\gamma + \lambda_r + \mu_t + \epsilon_p \quad (1)$$

where $\ln(Emigration)_{p,t}$ is the natural logarithm of the number of emigrants leaving parish p during decade t: 1868-1878, 1879-1888, 1889-1898, and 1899-1908. $Theil_{p,t-1}$ is the lagged value of the Theil index for parish p, where we use the Theil index from 1860 for the decade 1868-1877, the Theil from 1873 for the decade 1879-1888, the Theil from 1885 for the decade 1889-1898 and the Theil from 1895 for the decade 1899-1908. λ_r represent regional fixed effects (defined as: Greater Copenhagen, Zealand, Funen, and Jutland) and μ_t time fixed effects (decades). $X_p'\gamma$ is a vector of geographical characteristics of parish p, including the natural logarithm of the parish area, the distance to the nearest emigration port (either Copenhagen or Hamburg), the distance to the coast, the share of parish area classified as boulder clay and the population weighted inverse distance to all Danish market towns (averaged over 1860-1906).

	(1) ln(Emigration)	(2) ln(Emigration)	(3) ln(Emigration)	(4) ln(Emigration)	(5) ln(Emigration)
L.Theil	0.667***	0.702***	0.708***	0.697***	0.346***
	(0.069)	(0.070)	(0.070)	(0.070)	(0.060)
BoulderClay	0.660***	0.685***	0.699***	0.698***	0.095
	(0.080)	(0.080)	(0.080)	(0.080)	(0.075)
ln(area)	0.716***	0.705***	0.699***	0.710***	-0.118**
	(0.037)	(0.037)	(0.037)	(0.038)	(0.053)
ln(CPH/HH)		0.210**	0.698***	0.701***	0.822***
		(0.092)	(0.133)	(0.133)	(0.112)
ln(MarketTowns)			0.853***	0.881***	0.614***
			(0.194)	(0.195)	(0.173)
ln(DistCoast)				-0.029	0.070***
				(0.019)	(0.017)
LnPop1860					1.101***
					(0.062)
Constant	-1.499***	-3.645***	-17.016***	-17.302***	-20.634***
	(0.157)	(0.973)	(3.078)	(3.072)	(2.718)
Region FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Observations	5,729	5,729	5,729	5,729	5,729
R-squared	0.27	0.27	0.27	0.27	0.39
Parishes	1,582	1,582	1,582	1,582	1,582

BoulderClay is the share of parish area classified as boulder clay. ln(CPH/HH) denotes the natural logarithm of the distance to Copenhagen or Hamburg, whichever is shorter. ln(MarketTowns) denotes the natural logarithm of the inverse population weighted distance to all market towns. ln(DistCoast) denotes the natural logarithm of the distance to the coast. Regions are defined as: Greater Copenhagen, Zealand, Funen, North and South Jutland. Standard errors clustered at the parish level in parentheses.
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$

Table 1: Pooled OLS estimations explaining emigration patterns with land inequality

There is a strong positive correlation between land inequality and emigration. This correlation remains when controlling for geographical characteristics of the parish and initial population. We cannot claim a causal relationship from this, however, motivating our use of an instrumental variable specification.

5 IV

As explained earlier, the migration decision in itself is not random. We do not model the individual migration decision, but rather use data on the parish level and include various control variables for observable characteristics which influence the migration decision, such that some of this concern is already mitigated (see also Dinkelman and Mariotti, 2016).

In Boberg-Fazlic et al. (2020) we show that the extensive land reforms, consolidating and enclosing the land of farms and abolishing common land, at the end of the 18th century led to higher land inequality subsequently. We also show that the increase in inequality is due to higher population growth (because of better land use), which implied more people without land as the distribution of land was more or less fixed for institutional reasons (the number of farms stayed rather constant as well as average the size of farms). With the extensive data on land inequality over the following century, we are able to show that the pattern of inequality is determined with the effect of the reforms. Whereas inequality steadily increases over the following century, the pattern of inequality across the country stays remarkably constant. We therefore propose an instrumental variables strategy, using the change in the Theil index over the reforms, i.e. the change in inequality from 1682 to 1834, as an instrument for the level of inequality in 1860, which then determines the level of emigration over the following decades.

First-stage estimation:

$$Theil_{p,1860} = \alpha + \beta \times \Delta Theil_{p,1682-1834} + X_p'\gamma + \lambda_r + \epsilon_p \tag{2}$$

Second-stage estimation:

$$\ln(TotalEmigration_{p,1868-1908}) = \alpha + \beta \times Theil_{p,1860} + X_p'\gamma + \lambda_r + \epsilon_p \tag{3}$$

where $\ln(TotalEmigration_{p,1868-1908})$ is the total number of emigrants from parish p emigrating between 1868 and 1908. $\Delta Theil_p$ is the change in the Theil index over the course of land reforms (1682 to 1834) in parish p and $Theil_{p,1860}$ is the level of the Theil index in parish p in 1860. $X_p'\gamma$ is a vector of control variables including geographical characteristics of parish p, namely the natural logarithm of the parish area, the distance to the nearest emigration port (either Copenhagen or Hamburg), the distance to the coast, the share of parish area classified as boulder clay and the population weighted inverse distance to all Danish market towns (averaged over 1860-1906). λ_r represent regional fixed effects (defined as: Greater Copenhagen, Zealand, Funen, North and South Jutland) and ϵ_p is the error term. We include the distance to either Copenhagen or Hamburg as it is likely that the propensity to emigrate is higher the closer the closer the parish is to a major port for emigration. As mentioned earlier, we include the weighted inverse distance to market towns to capture the propensity to migrate to towns first and from there overseas.

Using the change in land inequality between 1682 and 1834 poses no reverse causality problem on the level of inequality almost 100 years later. There might, however, be omitted variable bias of some factor which causes both the change in inequality throughout the reforms and the level of inequality in 1860, for example the size of the parish or its income level (although it is not clear how exactly these would be related to either the change or the level in inequality). To take account of this concern we include parish area and the level of population in 1860 to measure the size of the parish. To measure income, we include the distance to the coast and the share of parish area covered with boulder clay, although also the distance to Copenhagen or Hamburg and the inverse weighted distance to market towns will most likely be positively correlated with income. The distance to the coast is a measure of market access and access to coal. In Boberg-Fazlic et al. (2020) we showed that parishes with higher shares of soil classified as boulder clay, experienced higher increases in inequality. This was explained by the fact that boulder clay was high quality soil for agriculture, most suitable for the cultivation of barley, which was the most common crop at the time. The land reforms can be considered to have constituted a technology shock of better land use, of which benefits could be reaped to a larger extent in areas with better soil. Therefore, these areas also experienced higher rates of population growth, which – with farm sizes and numbers being more-or-less fixed – lead to more people without land. Thus, we propose that the effect of boulder clay goes through the increase in inequality caused by the reforms. It does, however, also measure soil quality which will be related to income levels in rural areas. Although it is not clear that parishes with higher income levels should have higher (or lower) inequality, we show specifications including it as a control variable. Also, it is not clear whether we would expect income to have a positive or negative effect on total emigration. Hvidt (1971) concludes that higher incomes meant that more people could afford to emigrate, along with lower transportation costs, but of course higher incomes would also ceteris paribus lower the returns to migration. Given the distance to the nearest port, transportation costs should be the same for the whole country, such that we might expect a positive effect of income on the level of emigration, captured by the included controls.

	(1)	(2)	(3)	(4)	(5)	(6)
	1st stage Theil	2nd stage ln(TotalEmigr.)	2nd stage ln(TotalEmigr.)	2nd stage ln(TotalEmigr.)	2nd stage ln(TotalEmigr.)	2nd stage ln(TotalEmigr.) weighted
$D.Theil_{1682-1834}$	0.439^{***}					
	(0.034)					
$Theil_{1860}$		0.679^{***}	0.761^{***}	0.846^{***}	0.846^{***}	0.786^{***}
		(0.197)	(0.211)	(0.204)	(0.204)	(0.212)
BoulderClay	0.109^{***}	0.857^{***}	0.870^{***}	0.864^{***}	0.864^{***}	0.900^{***}
	(0.029)	(0.104)	(0.103)	(0.102)	(0.102)	(0.102)
ln(CPH/HH)	-0.070^{*}		0.189^{*}	0.935^{***}	0.935^{***}	0.910^{***}
	(0.041)		(0.110)	(0.155)	(0.154)	(0.157)
ln(MarketTowns)	0.030			1.281^{***}	1.282^{***}	1.287^{***}
	(0.057)			(0.222)	(0.224)	(0.230)
ln(DistCoast)	-0.032^{***}				-0.001	-0.008
	(0.006)				(0.023)	(0.024)
ln(area)	0.053^{***}	0.843^{***}	0.830^{***}	0.810^{***}	0.810^{***}	0.833^{***}
	(0.013)	(0.041)	(0.042)	(0.041)	(0.043)	(0.043)
Constant	1.071	-0.292	-2.276^{*}	-22.508^{***}	-22.521^{***}	-22.295^{***}
	(0.917)	(0.249)	(1.231)	(3.556)	(3.554)	(3.584)
Region FE	Y	Y	Y	Y	Y	Y
Observations	1,582	1,582	1,582	1,582	1,582	1,582
KP F-statistic		176.29	160.42	159.14	165.53	147.44

BoulderClay is the share of parish area classified as boulder clay. ln(CPH/HH) denotes the natural logarithm of the distance to Copenhagen or Hamburg, whichever is shorter. ln(MarketTowns) denotes the natural logarithm of the inverse population weighted distance to all market towns. ln(DistCoast) denotes the natural logarithm of the distance to the coast. Regions are defined as: Greater Copenhagen, Zealand, Funen, North and South Jutland. Robust standard errors in parentheses.
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$

Table 2: IV estimations using the change in land inequality 1682-1834

Table 2 shows the first stage estimation in column (1), including all control variables. Columns (2) to (6) show second stage estimations including different sets of controls. Land inequality is significant in all specifications and the coefficient is remarkably constant. Column (7) shows the second stage specification including the full set of controls weighted by parish population in 1860. Again, neither the size of the coefficient nor its significance are affected. Taking the coefficient of column (6), if the Theil index increases by one standard deviation (0.35) emigration increases by 41 percent.

5.1 Robustness checks

As shown in figure 3, *tyender* were largely overrepresented and likely even more so than can be implied from the occupational titles directly. We therefore used all emigrants from the parish in the main results in the previous section, where we estimated the effect of land inequality on emigration. We can, however, also restrict our analysis to those emigrants with the explicit occupational title *tyende*. These estimations are presented in the appendix, see table A2. The first two columns show the OLS estimations as in table 1 and columns (3) and (4) show the IV estimations as in table 2. The results are hardly affected by restricting the emigrants, both in terms of significance but also in terms of the size of the coefficient. This is reassuring evidence for the mechanism we propose.

Table A3 in the appendix shows results divided by men and women. One could imagine that different factors drive emigration by sex. However, we do not find any significant differences. Figure 3a showed that a large fraction of female emigrants were family members (i.e. wives, mothers, mothers-in-law, daughters, etc.) and the factors explaining their emigration will thus be the same as the factors explaining their husband's/son's/father's emigration.

Finally, to account for spatial correlation, we also provide Conley standard errors for different cut-off levels. Results can be found in appendix table A4. The significance of our estimates is not affected.

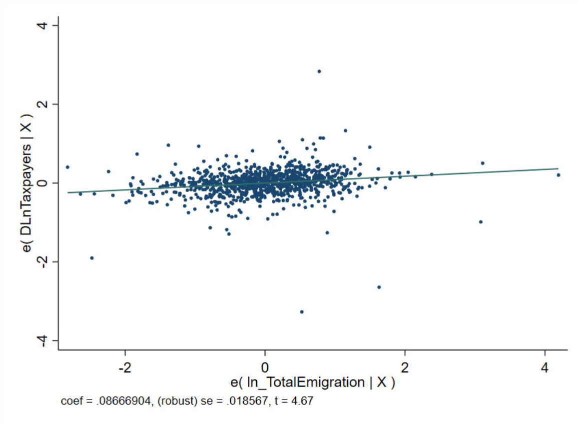
6 Effects on living standards in Denmark

We use income tax data from 1870 and from 1905 to look at the effect of emigration on local incomes (of the sending parish). Here, we specifically model possible income convergence across parishes. Usually, it is difficult to disentangle the effects of internal migration and emigration and the evidence is mixed (see, for example, Ozgen et al. (2010), who find little evidence for convergence). Enflo et al. (2014), for example, find evidence for migration contributing to wage convergence within Sweden, mainly by lowering wages in Stockholm and Gomellini and O'Grada (2011) demonstrate regional income convergence for Italy.

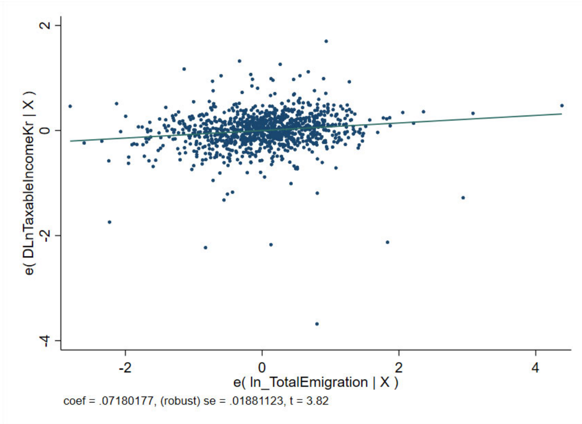
We specify the following pooled OLS model to investigate the effect of emigration on local incomes:

$$\Delta \ln(Y_{d,1870-1905}) = \alpha + \beta_1 \times \ln(Y_{d,1870}) + \beta_2 \times \ln(TotalEmigration_{d,1868-1908}) + X_d'\gamma + \lambda_r + \epsilon_d \quad (4)$$

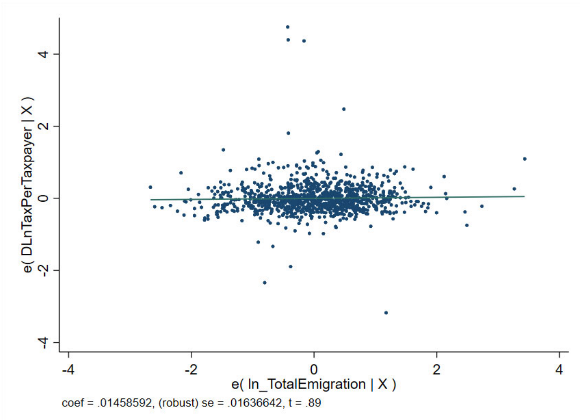
where $\Delta \ln(Y)_{d,1870-1905}$ is one of our outcome variables: *number of taxpayers*, *taxable income*, or *tax per taxpayer* in tax parish d. β_1 is the effect of the level of Y in 1870, i.e. it measures the extent of convergence across parishes. If β_1 is negative, there is convergence. Our main coefficient of interest is β_2 measuring the effect of emigration on the change in the outcome variable. The rest is as defined above. Figure 5 shows added variable plots for $\ln(TotalEmigration)$ from estimating equation 4 including the full set of controls. The respective tables can be found in the appendix (tables A5, A6, and A7).



(a) $\Delta \ln(numberoftaxpayers)$



(b) $\Delta \ln(taxableincome)$



(c) $\Delta \ln(taxpertaxpayer)$

Figure 5: Added variable plots for $\Delta \ln(numberoftaxpayers)$, $\Delta \ln(taxableincome)$, and $\Delta \ln(taxpertaxpayer)$, 1870-1905.

Clearly, there is a significant positive effect of emigration on the number of taxpaying households in the parish. In line with this, we also find a positive effect on the total taxable income in the parish. There is no robust effect, however, on the average tax paid per taxpayer (it is positive in some specifications, see table A7 in the appendix, but turns insignificant when adding control variables). Thus, more people move into the tax bracket in parishes with higher emigration. However, they seem to earn average incomes within the taxable incomes distribution (or the whole distribution is represented within the "new" taxpayers) as the average tax paid is not affected. Nevertheless, overall this constitutes a positive income effect of emigration for the parish as the number of people with relatively high incomes increases.

7 Conclusion

We have presented evidence that emigrants from Denmark during the Age of Mass Migration tended to be negatively selected. Many of them were from the class of *tyender* looking to improve their status through the acquisition of land, and we find that, consistent with this, emigration was greater from areas which had greater land inequality following the agrarian reforms a century previously. Finally, we find that areas of greater outmigration witnessed positive income effects, at least for a subset of those remaining.

It remains to be shown why incomes increased for the areas sending most migrants. We can think of two possible mechanisms: first, lower population density caused the marginal productivity of labour to increase, which would be consistent with evidence of generally increasing real wages over this period (Khaustova and Sharp, 2015). Second, labour shortages may lead to increasing mechanization, something also suggested by Hvidt (1971), who found evidence of labour shortages in the countryside. Various important new agricultural technologies were adopted at this time, most importantly the steam-powered automatic cream separator, and Denmark saw nothing less than a revolution in its agricultural sector from the early 1880s (Lampe and Sharp, 2018). Was there a switch to labour saving technology as (cheap) labour emigrated? This would be consistent with the work of Karadja and Prawitz (2019), who provide evidence from Sweden for substitution from labour to capital in agriculture as witnessed by the increased adoption of draft horses, which was a labour-saving technology. They also provide evidence for structural change in a companion paper, as measured by an increase in patents and fewer agricultural labourers but more industrial workers in high-emigration municipalities (Andersson et al., 2020).

Beyond the investigation of these mechanisms, we see at least two channels for future research. Following Karadja and Prawitz (2019) we could quantify the effect on Danish politics and the development of the welfare state. Second, it would be interesting to look at the impact of the arrival of the "sleeping giant" in the United States. Boberg-Fazlic and Sharp (2020) have already demonstrated that Danish migrants spurred the spread of the modern dairy industry, but the *tyende* class seem also to have been associated with a particular mentality, and a desire for social mobility, something which in the US would by the 1930s be referred to as the "American Dream".

References

Abramitzky, R., L. P. Boustan, and K. Eriksson (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.

Andersson, D., M. Karadja, and E. Prawitz (2020). Mass migration and technological change. Working Paper.

Baines, D. (1994). European emigration, 1815-1930: Looking at the emigration decision again. *Economic History Review 47*(3), 525–544.

Begley, J., F. Geary, and T. Stark (2016). Convergence in the pre-1914 Atlantic economy: what really happened to wages in Ireland between 1881 and 1911? *Cambridge Journal of Economics 40*, 43–67.

Boberg-Fazlic, N., M. Lampe, P. Martinelli Lasheras, and P. Sharp (2020). Winners and losers from enclosure: Evidence from Danish land inequality 1682-1895. Discussion Paper DP14616.

Boberg-Fazlic, N. and P. Sharp (2020). Immigrant communities and knowledge spillovers: Danish-Americans and the development of the dairy industry in the United States. CGO Working Paper 2020.010.

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *American Economic Review 77*(4), 531–553.

Carter, S. B. and R. Sutch (1999). Historical perspectives on the economic consequences of immigration into the United States. In C. Hirschman, P. Kasinitz, and J. DeWind (Eds.), *The Handbook of International Migration: The American Experience*. New York: Russell Sage Foundation.

Chiquiar, D. and G. H. Hanson (2005). International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States. *Journal of Political Economy 113*(2), 239–281.

Cohn, R. (1995). Occupational evidence on the causes of immigration to the United States. *Explorations in Economic History 32*, 383–408.

Commissionen for det statistiske Tabelværk i Danmark (1837). *En tabellarisk Fremstilling af Hartkornet, samtidlige Eiendommes Banktaxation of Bygningers Brandassurancesum, samt Jord-Eiendommes Fordeling m.v. i Kongeriget Danmark. (Statistisk Tabelværk, 2).* Bianco Luno & Schneider, Copenhagen.

Connor, D. S. (2019). The cream of the crop? Geography, networks, and Irish migrant selection in the age of mass migration. *Journal of Economic History 79*(1), 139–175.

Danish National Archives (1787). 1787 Census Denmark, (02/11/2018).

Danish National Archives (1801). 1801 Census Denmark, (20/11/2019).

Danish National Archives (1834). 1834 Census Denmark, (20/11/2019).

Danish National Archives (1850). 1850 Census Denmark, (02/11/2018).

Det Danske Udvandrerarkiv (2018). Københavns Politis Udvandrerprotokoller (hosted at Danish Demographic Database, delivery 12-02-2018).

Dinkelman, T. and M. Mariotti (2016). The long-run effects of labor migration on human capital formation in communities of origin. *American Economic Journal: Applied Economics 8*(4), 1–35.

Enflo, K., C. Lundh, and S. Prado (2014). The role of migration in regional wage convergence: Evidence from Sweden 1860-1940. *Explorations Economic History 52*, 93–110.

Feliciano, C. (2005). Educational selectivity in U.S. immigration: How do immigrants compare to those left behind? *Demography 42*(1), 131–52.

Furer, H. B. (1972). *The Scandinavians in America, 986-1970.* Dobbs Ferry, NY: Oceana Publications, Inc.

Gomellini, M. and C. O'Grada (2011). Outward and inward migrations in Italy: a historical perspective. Economic History Working Papers, Banca d'Italia, Number 8.

Gould, E. D. and O. Moav (2016). Does high inequality attract high skilled immigrants? *The Economic Journal 126*(593), 1055–1091.

Grogger, J. and G. H. Hanson (2011). Income maximization and the selection and sorting of international migrants. *Journal of Development Economics 95*(1), 42–57.

Hatton, T. J. and J. G. Williamson (1998). *The Age of Mass Migration: Causes and Economic Impact.* New York/Oxford: Oxford University Press.

Hvidt, K. (1971). *Flugten til Amerika, eller Drivkræfter I masseudvandringen fra Danmark 1868-1914.* Jysk Selskab for historie. Universitetsforlaget i Aarhus.

Jensen, H. (1945). *Dansk Jordpolitik 1757-1919. vol II: Fæstevæsenets Afvikling og Jordlovgivningen i Perioden 1810-1919.* Copenhagen: Gyldendal.

Karadja, M. and E. Prawitz (2019). Exit, voice, and political change: Evidence from Swedish mass migration to the United States. *Journal of Political Economy 127*(4), 1864–1925.

Khaustova, E. and P. Sharp (2015). A note on Danish living standards using historical wage series, 1731-1913. *The Journal of European Economic History 44*(3), 143–172.

Knudsen, A. S. B. (2019). Those who stayed: Selection and cultural change during the age of mass migration. Working paper.

Lampe, M. and P. Sharp (2018). *A Land of Milk and Butter: How Elites Created the Modern Danish Dairy Industry.* Chicago: University of Chicago Press.

Mackintosh, J. (1992). Elk Horn-Kimballton – the largest Danish settlement in America. In B. F. Larsen and H. Bender (Eds.), *Danish Emigration to the U.S.A.*, pp. 157–174. Aalborg, Denmark: Danes Worldwide Archives.

Mackintosh, J. (1993). The lure of prosperity: Economic development among Danish immigrants to Iowa. In B. F. Larsen, H. Bender, and K. Veien (Eds.), *On Distant Shores : Proceeding of the Marcus Lee Hansen Immigration Conference*, pp. 177–190. Aalborg, Denmark: Danes Worldwide Archives.

McKenzie, D. and H. Rapoport (2010). Self-selection patterns in Mexico-U.S. migration: The role of migration networks. *Review of Economics and Statistics 92*(4), 811–821.

Mishra, P. (2007). Emigration and wages in source countries: Evidence from Mexico. *Journal of Development Economics 82*, 180–199.

O'Rourke, K. H. and J. G. Williamson (1995). Open economy forces and late nineteenth century Swedish catch-up. A quantitative accounting. *Scandinavian Economic History Review 43*(2), 171–203.

Ozgen, C., P. Nijkamp, and J. Poot (2010). The effect of migration on income growth and convergence: meta-analytic evidence. *Papers in Regional Science 89*, 537–561.

Pedersen, H. (1928). *De danke Landbrug fremstillet paa Grundlag af Forarbejderne til Christian V.s Matrikel.* Gyldendal, Copenhagen.

Solvang, G. (1985). Husmandsbevægelsen i Danmark 1880-1920. En gennemgang af husmandsbevægelsens forudsætninger, politiske mål og midler samt bevægelsens omfang, resultater og betydning. *Bol og By: Landbohistorisk Tidsskrift 1*(1), 60–86.

Statistiske Bureau (1852). *En tabellarisk Oversigt over Hartkornets og Jordeiendommenes Fordeling i hvert af Kongeriget Danmarks Sogne og andre Bestanddele. (Statistisk Tabelværk Ny Række, 5).* Bianco Luno & Schneider, Copenhagen.

Statistiske Bureau (1864). *Tabeller over Hartkornets og Jordeiendommens Fordeling samt Belöbet af Gammelskat i hvert Sogn i Kongeriget Danmark den 1ste April 1860 (Statistisk Tabelværk, 3rd series, 4).* Bianco Luno & Schneider, Copenhagen.

Statistiske Bureau (1873). Statistiske Meddelelser, Ellevte Bind. Særskilte Oversigter over den i henhold til lov af 2den Juli 1870 i de enkelte skattekommuner paalignede overordentlige skat. Statens Statistiske Bureau, Copenhagen.

Statistiske Bureau (1877). *Tabeller over Hartkornets Fordeling i Kongeriget Danmark den 1ste April 1873 (Statistisk Tabelværk, 3rd series, 32).* Bianco Luno & Schneider, Copenhagen.

Statistiske Bureau (1880). Statistiske Meddelelser, Tredie Række, 4de Bind. Folkemængden i Kjøbstæderne og i Landsognene i Kongeriget Danmark efter Tællingen den 1. Februar 1880. Statens Statistiske Bureau, Copenhagen.

Statistiske Bureau (1888). *Danmarks Jordbrug, ordnede efter Størrelsen af deres Hartkorn den 1st Januar 1885. (Statistisk Tabelværk, 4th series, C, 5).* Bianco Luno & Schneider, Copenhagen.

Statistiske Bureau (1896). *Danmarks Jordbrug, ordnede efter Størrelsen af deres Hartkorn den 1st Januar 1895. (Statistisk Tabelværk, 4th series, C, 9).* Bianco Luno & Schneider, Copenhagen.

Statistiske Bureau (1905). Statistiske Meddelelser, Fjerde Række, Syttende Bind, Andet Hæfte. De nye statsskatter i skatteaaret 1904-05. Statens Statistiske Bureau, Copenhagen.

Statistiske Bureau (1906). Statistiske Meddelelser, Fjerde Række, En og Tyvende Bind, Første Hæfte. Folkemængden 1. Februar 1906 i Kongeriget Danmark. Statens Statistiske Bureau, Copenhagen.

Statistiske Bureau (1911). Statistiske Meddelelser, Fjerde Række, Syv og Tredivte Bind, Første Hæfte. Folkemængden 1. Februar 1911 i Kongeriget Danmark. Statens Statistiske Bureau, Copenhagen.

Stolz, Y. and J. Baten (2012). Brain drain in the age of mass migration: Does relative inequality explain migrant selectivity? *Explorations in Economic History 49*, 205–220.

van Vugt, W. (1988). Running from ruin?: The emigration of British farmers to the U.S.A. in the wake of the repeal of the corn laws. *Economic History Review 41*(3), 411–428.

Wegge, S. A. (1999). To part of not to part: Emigration and inheritance institutions in nineteenth-century Hesse-Cassel. *Explorations in Economic History 36*(1), 30–55.

Wegge, S. A. (2002). Occupational self-selection of European emigrants: Evidence from the nineteenth-century Hesse-Cassel. *European Review of Economic History 6*(2), 365–394.

Wegge, S. A. (2010). Who migrated and who stayed home? The village versus the individual: Self-selection of migrants out of the home population in 19th century Germany. unpublished.

8 Appendix

Table A1: Summary statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Emigration	6,328	15.53	25.75	0	527
Tyende	6,328	4.24	7.56	0	126
Theil	6,328	0.98	0.35	0.11	4.54
Δ Theil (1682-1834)	6,328	0.40	0.32	-1.03	2.27
Population 1860	6,328	767.5	688.5	55	1,0203
Distance to coast (km)	6,328	7.95	7.36	0.02	44.59
Distance CPH/HH (km)	6,328	165.24	75.20	2.80	295.12
Market town	6,328	9,111	5,318	2,820	127,821
Boulder clay	6,328	0.48	0.32	0	1
Area	6,328	22.70	18.43	1.73	184.12
Income taxpayers 1870	1,000	97.68	93.66	6	2,264
Income taxpayers 1905	1,000	151.21	195.1	11	4,188
Taxpayer rate 1870	1,000	0.08	0.03	0.004	0.88
Taxpayer rate 1905	1,000	0.10	0.04	0.002	0.97
Taxable income (Kr.) 1870	1,000	76,570	103,677	2,675	2.766,325
Taxable income (Kr.) 1905	1,000	264,892	489,663	12,643	14.300,000
Income/taxpayers 1870	1,000	24.8	15.35	3.26	146.05
Income/taxpayers 1905	1,000	16.01	62.88	0.39	1,605

Table A2: Robustness check: using *tyender* emigration only

	(1) ln(Tyende) OLS	(2) ln(Tyende) OLS, weighted	(3) ln(TotalTyende) IV	(4) ln(TotalTyende) IV, weighted
L.Theil	0.459***	0.499***		
	(0.054)	(0.090)		
Theil			0.779***	0.723***
			(0.212)	(0.218)
BoulderClay	0.612***	0.770***	0.923***	0.971***
	(0.065)	(0.109)	(0.100)	(0.100)
ln(CPH/HH)	0.422***	0.437*	0.792***	0.759***
	(0.116)	(0.229)	(0.151)	(0.152)
ln(MarketTowns)	0.418**	0.727**	0.873***	0.854***
	(0.170)	(0.310)	(0.216)	(0.221)
ln(DistCoast)	-0.054***	-0.082***	-0.034	-0.038
	(0.017)	(0.030)	(0.024)	(0.024)
ln(area)	0.517***	0.605***	0.714***	0.739***
	(0.033)	(0.062)	(0.043)	(0.043)
Constant	-10.298***	-13.832***	-18.296***	-17.778***
	(2.677)	(4.773)	(3.447)	(3.446)
Region FE	Y	Y	Y	Y
Year FE	Y	Y	N	N
Observations	4,503	4,503	1,519	1,519
KP F-statistic			153.68	136.63

BoulderClay is the share of parish area classified as boulder clay. ln(CPH/HH) denotes the natural logarithm of the distance to Copenhagen or Hamburg, whichever is shorter. ln(MarketTowns) denotes the natural logarithm of the inverse population weighted distance to all market towns. ln(DistCoast) denotes the natural logarithm of the distance to the coast. Regions are defined as: Greater Copenhagen, Zealand, Funen, North and South Jutland. Robust standard errors in parentheses.
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$

Table A3: Differential effects for men and women

	(1) ln(EmigrationMen) OLS, weighted	(2) ln(EmigrationMen) IV, weighted	(3) ln(EmigrationWomen) OLS, weighted	(4) ln(EmigrationWomen) IV, weighted
L.Theil	0.703***		0.822***	
	(0.115)		(0.125)	
Theil		0.712***		0.781***
		(0.202)		(0.236)
ln(CPH/HH)	0.811***	0.740***	1.056***	0.931***
	(0.276)	(0.145)	(0.315)	(0.167)
ln(MarketTowns)	1.524***	1.107***	1.853***	1.258***
	(0.389)	(0.214)	(0.454)	(0.243)
ln(DistCoast)	-0.039	-0.012	-0.042	-0.027
	(0.042)	(0.023)	(0.050)	(0.026)
ln(area)	0.818***	0.806***	0.863***	0.808***
	(0.064)	(0.041)	(0.075)	(0.046)
BoulderClay	0.875***	0.858***	0.899***	0.823***
	(0.141)	(0.097)	(0.170)	(0.108)
Constant	-23.762***	-19.033***	-30.389***	-23.162***
	(5.877)	(3.316)	(6.791)	(3.751)
Region FE	Y	Y	Y	Y
Year FE	Y	N	Y	N
Observations	11,039	1,577	10,661	1,523
KP F-statistic		147.46		134.26

BoulderClay is the share of parish area classified as boulder clay. ln(CPH/HH) denotes the natural logarithm of the distance to Copenhagen or Hamburg, whichever is shorter. ln(MarketTowns) denotes the natural logarithm of the inverse population weighted distance to all market towns. ln(DistCoast) denotes the natural logarithm of the distance to the coast. Regions are defined as: Greater Copenhagen, Zealand, Funen, North and South Jutland. Robust standard errors in parentheses.
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$

Table A4: Robustness check: accounting for spatial correlation using Conley standard errors

	(1) 2nd stage ln(EmigrationMen) 10 km	(2) 2nd stage ln(EmigrationMen) 25 km	(3) 2nd stage ln(EmigrationMen) 50 km	(4) 2nd stage ln(EmigrationMen) 100 km
Theil	0.846***	0.846***	0.846***	0.846***
	(0.204)	(0.204)	(0.205)	(0.205)
BoulderClay	0.864***	0.864***	0.864***	0.864***
	(0.102)	(0.102)	(0.103)	(0.102)
ln(CPH/HH)	0.935***	0.935***	0.935***	0.935***
	(0.154)	(0.154)	(0.155)	(0.155)
ln(MarketTowns)	1.282***	1.282***	1.282***	1.282***
	(0.224)	(0.224)	(0.225)	(0.224)
ln(DistCoast)	-0.001	-0.001	-0.001	-0.001
	(0.023)	(0.023)	(0.023)	(0.024)
ln(area)	0.810***	0.810***	0.810***	0.810***
	(0.043)	(0.043)	(0.043)	(0.043)
Constant	-22.521***	-22.521***	-22.521***	-22.521***
	(3.554)	(3.554)	(3.562)	(3.563)
Region FE	Y	Y	Y	Y
Observations	1,582	1,582	1,582	1,582
KP F-statistic				

BoulderClay is the share of parish area classified as boulder clay. ln(CPH/HH) denotes the natural logarithm of the distance to Copenhagen or Hamburg, whichever is shorter. ln(MarketTowns) denotes the natural logarithm of the inverse population weighted distance to all market towns. ln(DistCoast) denotes the natural logarithm of the distance to the coast. Regions are defined as: Greater Copenhagen, Zealand, Funen, North and South Jutland. Robust standard errors in parentheses.
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$

Table A5: Pooled OLS, dependent variable: $\Delta \ln(number of taxpayers)$

	(1) DLnTaxpayers	(2) DLnTaxpayers	(3) DLnTaxpayers	(4) DLnTaxpayers
LnTaxpayers1870	-0.179***	-0.261***	-0.310***	-0.488***
	(0.043)	(0.049)	(0.049)	(0.074)
ln_TotalEmigration		0.087***	0.100***	0.087***
		(0.019)	(0.019)	(0.019)
LnDPop			0.680***	0.701***
			(0.160)	(0.159)
BoulderClay				0.420***
				(0.061)
ln(DistPort)				-0.111
				(0.073)
ln(MarketTowns)				-0.031
				(0.105)
ln(DistCoast)				-0.052***
				(0.014)
ln(area)				0.291***
				(0.079)
Constant	1.179***	1.181***	1.473***	2.769*
	(0.186)	(0.181)	(0.164)	(1.413)
Region FE	N	N	Y	Y
Observations	1,000	1,000	1,000	1,000
R-squared	0.06	0.10	0.24	0.32

BoulderClay is the share of parish area classified as boulder clay. ln(CPH/HH) denotes the natural logarithm of the distance to Copenhagen or Hamburg, whichever is shorter. ln(MarketTowns) denotes the natural logarithm of the inverse population weighted distance to all market towns. ln(DistCoast) denotes the natural logarithm of the distance to the coast. Regions are defined as: Greater Copenhagen, Zealand, Funen, North and South Jutland. Robust standard errors in parentheses.
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$

Table A6: Pooled OLS, dependent variable: $\Delta \ln(taxableincome)$

	(1) DLnTaxableIncomeKr	(2) DLnTaxableIncomeKr	(3) DLnTaxableIncomeKr	(4) DLnTaxableIncomeKr
LnTaxableIncome1870	-0.278***	-0.352***	-0.343***	-0.477***
	(0.042)	(0.046)	(0.057)	(0.065)
ln_TotalEmigration		0.097***	0.101***	0.072***
		(0.019)	(0.019)	(0.019)
LnDPop			0.855***	0.818***
			(0.177)	(0.167)
BoulderClay				0.330***
				(0.079)
ln(DistPort)				-0.040
				(0.080)
ln(MarketTowns)				-0.026
				(0.119)
ln(DistCoast)				-0.040***
				(0.014)
ln(area)				0.295***
				(0.070)
Constant	4.282***	4.702***	4.653***	5.942***
	(0.456)	(0.471)	(0.555)	(1.678)
Region FE	N	N	Y	Y
Observations	1,000	1,000	1,000	1,000
R-squared	0.15	0.18	0.35	0.41

BoulderClay is the share of parish area classified as boulder clay. ln(CPH/HH) denotes the natural logarithm of the distance to Copenhagen or Hamburg, whichever is shorter. ln(MarketTowns) denotes the natural logarithm of the inverse population weighted distance to all market towns. ln(DistCoast) denotes the natural logarithm of the distance to the coast. Regions are defined as: Greater Copenhagen, Zealand, Funen, North and South Jutland. Robust standard errors in parentheses.
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$

Table A7: Pooled OLS, dependent variable: $\Delta \ln(taxpertaxpayer)$

	(1) DLnTaxPerTaxpayer	(2) DLnTaxPerTaxpayer	(3) DLnTaxPerTaxpayer	(4) DLnTaxPerTaxpayer
LnIncomeTaxPT1870	-0.453***	-0.458***	-0.434***	-0.415***
	(0.030)	(0.030)	(0.034)	(0.040)
ln_TotalEmigration		0.044***	0.035**	0.015
		(0.014)	(0.014)	(0.016)
LnDPop			0.159**	0.171**
			(0.077)	(0.076)
BoulderClay				0.095
				(0.080)
ln(DistPort)				-0.031
				(0.094)
ln(MarketTowns)				-0.176
				(0.135)
ln(DistCoast)				-0.016
				(0.014)
ln(area)				0.094***
				(0.032)
Constant	0.748***	0.585***	0.508***	2.242
	(0.090)	(0.093)	(0.117)	(2.196)
Region FE	N	N	Y	Y
Observations	1,000	1,000	1,000	1,000
R-squared	0.22	0.23	0.23	0.24

BoulderClay is the share of parish area classified as boulder clay. ln(CPH/HH) denotes the natural logarithm of the distance to Copenhagen or Hamburg, whichever is shorter. ln(MarketTowns) denotes the natural logarithm of the inverse population weighted distance to all market towns. ln(DistCoast) denotes the natural logarithm of the distance to the coast. Regions are defined as: Greater Copenhagen, Zealand, Funen, North and South Jutland. Robust standard errors in parentheses.
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$

EHES Working Paper Series

Recent EHES Working Papers

2021

EHES 212 The loss of human capital after the Spanish civil war
Blanca Sánchez-Alonso, Carlos Santiago-Caballero

EHES 211 Testing Marx. Income Inequality, Concentration, and Socialism in late 19th century Germany
Charlotte Bartels, Felix Kersting, Nikolaus Wolf

EHES 210 Inequality Beyond GDP: A Long View
Leandro Prados de la Escosura

EHES 209 Paving the way to modern growth. Evidence from Bourbon roads in Spain
Miquel-Àngel Garcia-López, Alfonso Herranz-Loncan, Filippo Tassinari, Elisabet Viladecans-Marsal

EHES 208 Why Covid19 will not be gone soon: Lessons from the institutional economics of smallpox vaccination in 19th Century Germany
Katharina Muhlhoff

EHES 207 Leaving Their Mark: Using Danish Student Grade Lists to Construct a More Detailed Measure of Historical Human Capital
Nicholas Ford, Kristin Ranestad, Paul Sharp

EHES 206 Violence in the Viking World: New Bioarchaeological Evidence
Joerg Baten, Giacomo Benati, Anna Kjellström

EHES 205 Productivity, Mortality, and Technology in European and US Coal Mining, 1800-1913
Javier Silvestre

2020

EHES 204 Globalization and Empire: Market integration and international trade between Canada, the United States and Britain, 1750-1870
Maja Uhre Pedersen, Vincent Geloso, Paul Sharp

All papers may be downloaded free of charge from: http://www.ehes.org/