SIEPR Discussion Paper No. 09-029

Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration

By
Ran Abramitzky, Leah Platt Boustan, and Katherine Eriksson
Stanford University

August 2010

Stanford Institute for Economic Policy Research
Stanford University
Stanford, CA 94305
(650) 725-1874

Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration*

Ran Abramitzky	Leah Platt Boustan	Katherine Eriksson
Stanford University	UCLA and NBER	UCLA

August, 2010

Abstract

The Age of Mass Migration (1850-1913) was among the largest migration episodes in history. Unlike today, the United States maintained an open border in this era. We compile a novel dataset of Norway-to-US migrants and estimate the return to migration while accounting for migrant selection. Our first method compares migrants to their brothers who remained in Norway; our second exploits the fact that, under primogeniture, older sons in land-owning families were less likely to migrate. We find that these migrants, unhindered by entry restrictions, were negatively selected from the sending population, and that the return to migration was relatively low.

JEL Code: J61, N30

Keywords: Migration, selection

* We have benefited from conversations with Timothy Bresnahan, Dora Costa, Pascaline Dupas, Joseph Ferrie, Claudia Goldin, Avner Greif, Timothy Guinnane, Rick Hornbeck, Seema Jayachandran, Lawrence Katz, Naomi Lamoreaux, Shirlee Lichtman, Robert Margo, Roy Mill, Joel Mokyr, Paul Rhode, Izi Sin, Gunnar Thorvaldsen, Gui Woolston, Gavin Wright and members of the KALER group at UCLA. We thank seminar participants at Harvard, Queen's, Simon Frasier, Toronto and Yale, as well as conference participants at the Economic History Association, the Social Science History Association and the Development of the American Economy and Labor Studies groups at the NBER. Matthew Baird helped to collect data from Ancestry.com. John Parman and Sula Sarkar generously shared data with us. We acknowledge financial support from the California Center for Population Research and the Center for Economic History at UCLA.

"Keep, ancient lands, your storied pomp!" cries she
With silent lips. "Give me your tired, your poor,
Your huddled masses yearning to breathe free,
The wretched refuse of your teeming shore.
Send these, the homeless, tempest-tost to me,
I lift my lamp beside the golden door!"

- Emma Lazarus, "The New Colossus" (1883)
  Displayed upon the Statue of Liberty in New York Harbor

1. Introduction

The Age of Mass Migration from Europe to the New World was one of the largest migration episodes in human history. Between 1850 and 1913, the United States absorbed nearly 30 million European immigrants. This paper asks two related questions about this migrant flow. First, were migrants positively or negatively selected from the European population? We test whether the US acquired higher skilled European migrants who were able to finance the voyage or whether it absorbed Europe's "tired, poor, huddled masses" who migrated to the US in search of opportunity. Secondly, what was the economic return to migrating from Europe to the United States in the late nineteenth century?

Understanding migrant selection is of particular importance in this historical period. Given the magnitude of the migration flow, the skill composition of departing migrants had potentially large implications for relative economic growth. If migrants were negatively selected, for example, this outflow may help to explain convergence between the US and poorer European countries. Furthermore, the US maintained a nearly open border in the late nineteenth century, allowing us to study the economic process of migrant self-selection without interference from the bureaucratic factors that govern migrant selection today. In contrast, in the current period, the immigrant flow is a product of both individual migration decisions and the application of

complicated entry rules and restrictions, obscuring the underlying economic forces. Thus, comparing our findings with contemporary studies can illuminate the effect of modern immigration policy on migrant selection.

Our empirical methods are also of general interest to labor economics and the economics of migration. Because migrants may not be randomly selected from the sending population, it is challenging to separately identify the return to migration and the selection into migration. Attempting to measure the return to migration with a naïve comparison of migrants and stayers would be confounded by migrant selection. For example, Europe-to-US migrants may earn more than men who remained in Europe because the brightest people, who would have enjoyed higher earnings even if they had stayed in Europe, are the most likely to move. Therefore, in the presence of positive selection, a naïve OLS estimate of the return to migration will be biased upward and similarly in the presence of negative selection it will be biased downward.

We use two alternative and complementary empirical strategies to estimate the return to migration while accounting for selection into migration. First, we use OLS regressions to compare the earnings of migrants to the earnings of their brother(s) who remain in Europe. This within-brother estimate eliminates selection across households, which can result from differing propensities to migrate for households whose members face poor economic opportunities in Europe or for households with the financial capacity to move.

Selection might also occur within households because brothers vary in their innate capacities or social roles. Our second approach addresses the potential for selection both across and within households. We use instrumental variable regressions to compare the earnings of migrants and non-migrants, using birth order to instrument for migration. According to inheritance customs, older brothers were more likely to inherit their parents' land;

correspondingly, we find that younger brothers were more likely to migrate to the US, especially from rural areas and from households that owned land. In some specifications, we also make use of the fact that primogeniture customs were stronger on the western coast of Norway than in the more developed Southeast. In particular, we allow the effect of birth order and its interaction with land ownership to vary by region, thereby controlling for other social or biological aspects of birth order that may be correlated with later labor market outcomes.

Beyond providing more accurate estimates of the return to migration, these methods allow us to infer the nature and extent of migrant selection both across and within households. Specifically, a comparison between the within-brother estimate and the naïve OLS estimate reveals the nature of selection across households, and a comparison between the IV estimate and the within-brother estimate reveals the nature of selection within households.

We focus on Norwegian migrants to the US. Norway had one of the highest out-migration rates among European sending countries, with over a quarter of its population eventually migrating to the US. More importantly, Norway completely digitized two censuses from the period (1865 and 1900), allowing us to follow large samples of migrants and non-migrants over time in Norway and to the US (Ferrie, 1996). Specifically, we create a novel data set of all Norwegian–born men in the US in 1900 using US Census records from the genealogy website, Ancestry.com. We then match men by name and age to their birth family in Norway in 1865. For comparison, we are also able to follow migrants' brothers and a sample of other men to the Norwegian labor market in 1900.

We observe each individual's occupation in either the US or the Norwegian labor market in 1900.[1] We then assign individuals the mean earnings for their occupation in either Norway or

[1] In principle, one could also study migrant selection by comparing the education levels or literacy rates of migrants to men who remained in Norway. However, the Norwegian Census did not collect information on literacy or years

the US (in real PPP-adjusted 1900 US dollars). For simplicity, we often refer to this occupation-based earnings measure as "earnings," but it can also usefully be thought of as an occupational ranking. While this measure captures two components of the return to migration, namely the potential for higher mean earnings in the US for each occupation and the potential for occupational upgrading, it cannot account for the potential for a higher return to skill *within* occupation in the US. Despite this drawback, the historical data has an important advantage over its modern counterparts. Due to privacy restrictions, the individual names that we use to match migrants to their birth families are only released 70 or more years after the initial Census was taken, rendering historical Census data the only large data set available for sibling comparisons or household-based instruments for migration.

We find evidence of negative selection both across- and within-households. That is, men with poorer economic prospects in Norway were more likely to move to the US in the late nineteenth century.[2] Both of our estimation strategies suggest that the return to migration is understated by 20-30 percent due to a process of negative selection. Once accounting for migrants' negative selection, we estimate a return to migration of 60 percent for men born in urban areas and 120 percent for men born in rural areas. Such returns are lower than contemporary estimates for the return to migration from Mexico to the United States (200-400 percent; see Hanson, 2006).

The remainder of the paper proceeds as follows. Section 2 discusses the historical context and related literature on the age of mass migration and migrant selection. Section 3 describes the data and the procedures we used to match migrants to their birth families in Norway. Section 4

---

of schooling in 1900. 97 percent of Norwegian-born men in the relevant age range who are observed in the US Census in 1900 report being literate.

[2] Again, we note that we can only measure selection across occupations, and migrants may have been the brightest and most motivated among the low-skilled.

presents the results of our two estimation strategies – namely conducting a within-brother analysis and using birth order as an instrument for migration. Section 5 concludes.

2. Contemporary and historical literature on migrant selection

A. Migration selection in the Roy model

Borjas (1987, 1991) modified the Roy model of occupational choice to generate predictions about the nature of migrant selection (Roy, 1951).[3] In this framework, migrant selection is determined by the relative variances of the income distribution in the sending and destination economies. If the destination country exhibits higher returns to skill, and therefore greater levels of income inequality, than the source country, migrants will be drawn from the top end of the source country's skill distribution. If, instead, the destination country offers lower returns to skill and is therefore more equal than the source, migrants will be disproportionately drawn from the low end of the source country's skill distribution.

Work on contemporary immigrant flows has found only mixed support for the Borjas model of migrant selection. Feliciano (2005) and Grogger and Hanson (2008) show that migrants are positively selected on educational attainment from almost every sending country in the world, even those countries with very high levels of income inequality.[4] As a result, studies of individual migrant flows will either confirm or contradict the Borjas model depending on whether the level of income inequality in the sending country in question is high or low.[5] For example, Chiquiar and Hanson (2006) observe that Mexican migrants to the United States are

[3] For an alternative view on migrant selectivity, see Chiswick (1999, 2000).
[4] Puerto Rico is one notable exception. Migrants leaving the highly unequal island for the more equal mainland are negatively selected whereas arrivals to the island are positively selected (Borjas, 2008).
[5] Because the data requirements for studying self-selection are so stringent, many studies of particular source and destination pairs focus on unique settings – including the Israeli *kibbutz* (Abramitzky, 2008) and the Pacific Islands (Akee, 2010; McKenzie, Gibson, and Stillman, forthcoming; McKenzie and Gibson, forthcoming).

drawn from the middle, rather than the low end, of the Mexican skill distribution, despite the fact that income inequality is higher in Mexico than in the United States. However, Gould and Moav (2010) find that Israeli migrants to the United States are positively selected, at least on observable skills, as would be predicted by the more compressed distribution of Israeli wages.

Scholars have attempted to reconcile the Borjas model with the facts about positive selection in a variety of ways. A new generation of models incorporates borrowing constraints and shows that, as the cost of migration increases, the poorest residents of sending countries can no longer afford to move (Borger, 2009; McKenzie and Rapoport, forthcoming).[6] Alternatively, Grogger and Hanson (2008) demonstrate that a classic Roy model with a linear, rather than a logarithmic, utility function generates predictions of positive selection whenever the skill-related differences in *wage levels*, rather than the relative returns to skill *in percentage terms*, are high. In this framework, positive selection is a likely outcome in the contemporary world given the dramatic difference in wage levels between developed and developing countries.

B. Migration selection in the Age of Mass Migration

Between 1850 and 1913, more than 40 million Europeans moved to the New World, nearly two-thirds of whom settled in the United States.[7] Initially, migrants from the British Isles and Germany constituted the majority of the migrant flow to the US. These early migrants were joined by Scandinavians and other Northern Europeans in the 1870s and by Southern and Eastern Europeans in the 1880s. Norway experienced one of the highest out-migration rates in the 1880s, during which time 95 of every thousand Norwegians left the country.

[6] The cost of migrating from Mexico to the US is around $2000 in 2000 US dollars, or 35 percent of the annual earnings of a low-skilled Mexican worker (Hanson, 2006).
[7] This paragraph is based on Hatton and Williamson (1994, 1998).

With the shift from sail to steam technology on the Atlantic, the cost of migration fell dramatically over the nineteenth century (Keeling, 1999). Hatton and Williamson (2006) emphasize that the declining cost of migration, coupled with rising real incomes in the newly industrializing countries of Southern and Eastern Europe, relaxed the financial constraints on households that were previously too poor to pay for passage to the New World. As a result, they argue, migrant selection became increasingly negative over the century as migration became affordable to a greater share of the European population.

Certainly, Hatton and Williamson's prediction is true at the "macro" level; migrants were increasingly drawn from poorer Southern and Eastern European countries over time as transport costs fell. Yet, to our knowledge, Wegge (2002, 2010) are the only papers to provide individual-level evidence on migrant selection in the nineteenth century.[8] Wegge documents intermediate selection for the emigration flow leaving Germany in the 1850s: members of the highest- and lowest-skill occupations were less likely to migrate than were workers in the mid-skill range, such as machinists, metal workers and brewers. She concludes that the poorest migrants may have lacked the resources necessary to finance their trip. This result is specific to the 1850s when the cost of passage to the New World exceeded the total annual earnings of the average German laborer. Following Hatton and Williamson's logic, we would expect the later Norwegian migration to be more negatively selected than the earlier German flow.

An application of the Roy model to our historical context also generates a prediction of negative selection. According to the Roy model, migrants will be negatively selected when the income distribution in the destination country is more compressed than that of the sending country. In the modern context, the Scandinavian countries are more equal than the United

[8] For work on migrant selection in other historical periods, see Ferrie (1999) on rural-to-urban migration in the US, Margo (1990) on black migrants leaving the US South, and Abramitzky and Braggion (2006) on indentured servants to New World colonies.

States. However, in the late nineteenth and early twentieth century, the opposite was true. Figure 1 compares the occupation-based earnings distribution in the United States and Norway in 1900. We array individuals from lowest- to highest-paid with earnings represented in US dollars and the Norwegian distribution rescaled to share the US mean (the data are described in more detail in Section 3). Low-skilled workers in the US out-earned similar Norwegians, while Norwegians above the 90th percentile of the income distribution commanded higher earnings than their US counterparts. These occupation-based earnings distributions suggest that Norway offered a higher return to skill than did the United States circa 1900, which is consistent with the historical evidence on 90-50 ratios in the two countries (Soltow, 1965; Goldin and Katz, 1999).[9]

Not only could men at the low end of the Norwegian skill distribution expect a higher return to migration at a point in time but low-skilled men in the US also benefitted from substantial occupational mobility over the life-cycle. Ferrie and Long (2004) document that only 18 percent of men in the US who held an unskilled, blue collar job in 1850 remained unskilled workers by 1880. By matching across two Norwegian Censuses, we find that 47 percent of men in unskilled, blue collar occupations in Norway in 1875 remained unskilled workers 25 years later. Low skilled workers in the US were twice as likely to move up the occupational ladder over their lifetime, especially by becoming owner-occupier farmers.

In historical terms, the costs of migration were relatively low in the late nineteenth century. We estimate that the total cost of migration, including foregone earnings during the

[9] Soltow (1965) compares the average earnings for men in the top decile of the income distribution to mean earnings in urban places in Norway in 1890. He finds a [90-100]/mean ratio of 4.25. Goldin and Katz (1999) instead calculate a more conventional 90/50 ratio of 1.71 for 12 urban industries in the US in 1890. To compare these two figures, we use two adjustment factors: (1) the ratio of median to mean income of 0.93 from the 1911 Canadian Census (Green and Green, 2008) and (2) the ratio of average earnings in the top decile to earnings at the 90th percentile of 2.56 from the 1917 US tax returns (Piketty and Saez, 2003). By this method, the pseudo [90-100]/mean ratio for the US in 1890 is 4.06, which is a bit lower than Norway. We note that our adjustment factors may inflate the US ratio, particularly because top-end inequality in the US was likely higher in 1917 than in 1890. Even by this conservative measure, we find that Norway was less equal than the US.

voyage, would represent around 18 percent of the annual earnings of a Norwegian farm laborer.[10] Migrant networks also helped to defray the cost of passage for new arrivals; 40 percent of Norwegian migrants during this period travelled on pre-paid steamship tickets financed by friends or relatives (Hvidt, 1975, p. 129).

3. Data and Matching

A. Occupation and earnings data in Norway and the United States

Our goal is to identify Norwegian men in migrant and non-migrant households for whom we can observe labor market outcomes later in life. We rely on three Census sources, the complete digitized Norwegian Censuses of 1865 and 1900 and a data set containing the full population of Norwegian-born men in the US in 1900. We create the full sample of Norwegian immigrants living in the US in 1900 using the genealogy website Ancestry.com. The Norwegian Census data are archived at the North Atlantic Population Project (NAPP).

We observe labor market outcomes in 1900, when the men in our sample are in their 30s and 40s. Neither the US nor the Norwegian Census of 1900 contains information on wages or income. Instead, we assign men the mean income earned by members of their occupation.[11] Men living in the United States are matched to income data from the 1901 Cost of Living Survey while men living in Norway are matched to mean income by occupation tabulations for the year 1900 published by Statistics Norway and other sources (Haines and Preston, 1991; *Statistik*

[10] Norwegian farm laborers earned around $175 in 1900 US dollars. For this calculation, we assume that migrants lost 20 days of work for the passage and the resettlement. However, it is interesting to note that Armstrong and Lewis (2009) report that the typical Dutch migrant to Canada in the 1920s saved around $150 (in 1900 US dollars) for the cost of the voyage and resettlement, nearly a full year's salary for a Norwegian farm laborer.

[11] For men living in the US, we code occupation by hand using the digital images of Census manuscripts available on Ancestry.com.

Aarbog, 1900; Grytten, 2007).[12] The 1901 Cost of Living Survey reports income information for more than 300 occupations in the US. At least one member of our sample is employed in 189 of these categories. We convert Norwegian wages to real, PPP-adjusted US dollars using the 1900 exchange rate and price levels reported in Grytten (2004). The Data Appendix provides more detail on the data sources and assumptions underlying these estimates.

The 1901 Cost of Living survey may overstate the return to migration both because the survey was conducted in urban areas and because the majority of survey respondents were native born. We also assign to our US observations the mean occupation-earnings measures from the 1915 Iowa Census, which better represents the urban and rural status of our population. Furthermore, native-born workers, who make up a large share of respondents in the Cost of Living Survey, may have earned more than the typical Norwegian migrant. According to worker surveys conducted by the Immigration Commission in the 1900s, Scandinavian migrants earned 15 log points below native-born workers of native parentage in the same industry due either to discrimination or a lack of US-specific skills (Hatton and Williamson, 1998, p. 146-47). As a result, our estimates of the return to migration may be overstated but our conclusions about migrant selection will be unaffected.

Table 1 reports the ten most common occupations for our sample of matched brothers in Norway and the United States. 42 percent of Norway-to-US migrants in our sample worked in farm occupations, compared to only 26 percent of our sample of Norwegian stayers. Migrants to the US were also far more likely to report being general laborers (8 percent versus 1.4 percent).

[12] *Statistics Norway* reports daily wage rates. We convert these wage rates into annual earnings figures by assuming that Norwegians worked six-day work-weeks and were unemployed for 0.66 months during the year (= 297 days of work per year, on average). Our estimate for months spent unemployment is based on reported unemployment for Norwegian migrants in the 1900 US Census.

Other common occupations in both countries include carpenters, fisherman and sawmill operatives.

Our unavoidable reliance on mean earnings by occupation prevents us from measuring the full return to migration. Conceptually, the return to migration can be decomposed into: (1) the presence of higher wages in the US in the typical occupation; (2) the possibility that migrants are able to switch from low-paying to high-paying occupations upon arriving in the US; and (3) the existence of a higher within-occupation return to ability in the US. Our estimate of the return to migration captures only the first two aspects of the total return.

We face a related limitation in our ability to describe the extent of migrant selection. Positive selection, for instance, could be generated either by high migration rates among men from occupations with high mean earnings or by high migration rates among men at the 80^{th} or 90^{th} percentile of the wage distribution *within* their occupation. The reverse is true, of course, for negative selection. With our data, we can document the fact that more (fewer) common laborers move to the US but we will not be able to observe whether the best (worst) among the laborers made the journey.

B. Matching Norwegian-born migrants to their birth families

We construct our data set of matched migrant and non-migrant Norwegian men in two steps. First, we match Norwegian-born men from the 1900 US Census to their birth families in the 1865 Norwegian data. Then, we match Norwegian men living in Norway in 1900 to the 1865 Census. Through this process, we identify some households that contain both a matched migrant and a matched stayer. Our baseline method ("**Match 1**") uses an iterative matching strategy pioneered by Ferrie (1996). We describe this procedure in detail:

(1) We identify 55,079 Norwegian-born men between the ages of 38-50 in the 1900 US Census. Men in this age range were likely to be living with their birth families in 1865.[13]

(2) We standardize all first and last names in the US and the complete 1865 Norwegian Census to address orthographic differences between phonetically equivalent names using the NYSIIS algorithm (see Atack and Bateman, 1992).

(3) 30,629 of the Norwegian-born men in the US are unique by first name, last name and birth year. We match these unique observations back to 1865 using the following iterative procedure. We start by looking for a match by name and exact birth year. If we find a *unique* match here, we stop and consider the observation "matched." If we find multiple matches for the same birth year, the observation is thrown out. If we do not find a match at this first step, we try matching first within a one-year band (older or younger) and then with a two-year band around the reported birth year. If neither of these attempts produces a match, the observation is considered to be "unmatched."

This procedure creates a sample of 4,563 men from migrant households – 3,920 men who lived in the United States in 1900 and 643 who lived in Norway in 1900 – and more than 21,000 men from non-migrant households.[14] Our matched sample produces an emigration rate of 12 to 15 percent, which corresponds to the emigration rate in the aggregate data.[15] We achieve a match

[13] We restrict our attention to men who are at least three years of age in 1865 to ensure that all observations can match to a two-year age band around the reported age (see step 3). We omit men who are reported to be less than one year of age in the 1865 Census because of concerns about data quality in this subsample.

[14] Matches conducted wholly within Norway can make use of an individual's province or municipality of birth in addition to his name and age. Adding an extra matching variable would allow us to distinguish between some men who have the same name and age, thereby increasing our sample size. However, we maintain the more restrictive match by name and age alone for both US-to-Norway and Norwegian matches in order to ensure that migrants and their brothers face the same probability of entering the sample.

[15] Semmingsen (1978, p. 99) reports that over 700,000 Norwegians left the country from 1866 to 1915 out of a base population of 2.2 million, an out-migration rate of 32 percent. 75 percent of Norwegian emigrants settled in the United States. According to Ferenczi and Willcox (1929), 64 percent of the Norwegian migrants who entered the

rate of 13 percent among Norwegian-born men with a unique name-birth year combination who lived in the US in 1900 and XX percent among men in Norway in 1900. These rates are comparable to Ferrie's (1996) backwards match rate of 19 percent within the United States over a single Census decade. Ferrie's match rates are slightly higher than ours because his match is conducted over a single decade and he is able to match not only on name and age but also on state of birth.

We are concerned that the iterative nature of this method will produce false matches. False matches may occur because we may keep men in the sample who have both an exact match and a "close" match (within a one- or two-year band around the reported birth year). We design a second matching procedure ("**Match 2**") to address this concern. Match 2 conducts a single match for a restricted sample of men who are unique by name within a *five-year age band* in both Censuses (two years around the reported age in each direction). In so doing, we limit the potential for false matches in 1900 but we also reduce the ultimate size of the sample. The result is a brothers sample of 1,420 observations – 1,273 men who lived in the United States in 1900 and 147 who lived in Norway – along with nearly 13,000 men from non-migrant households that can be matched between the Norwegian Censuses.

C. Comparing matched samples to the full population

In order to enter our matched samples, an individual must be present in both the 1865 Norwegian Census and either the US or Norwegian Census of 1900. That is, men who die or move to another country between 1865 and 1900 will not be part of the matched samples, nor will men with relatively common names for whom we cannot find a unique match. As a result,

---

United States between 1865 and 1915 arrived by 1900. Taken together, these numbers suggest that 15 percent of Norwegians left for the United States between 1865 to 1900 (= 0.32 emigration rate · 0.75 to US · 0.64 by 1900).

our matched samples may not be representative of either the 1865 population in Norway or of the 1900 population of Norwegian-born men. In particular, we expect that our matched samples may be drawn from households with a higher socio-economic status, as such households used a wider array of given names (Gjerde, 1985, p. 48).

Table 2a compares the attributes of men in the two complete matched samples to the Norwegian population in the same age range in the 1865 Census, while Table 2b compares matched migrants to Norwegian-born men living in the United States and matched stayers to men living in Norway in 1900. In 1865, differences between the matched sample and the population can be driven either by differences in naming practices or by mortality bias. In contrast, all individuals in the 1900 data survived until 1900 (by definition), with any remaining differences due to the matching procedure itself.[16]

The most striking difference between our matched samples and the Norwegian population in 1865 is the probability of living in an urban area. In that year, only 14 percent of the Norwegian population lived in an urban area, compared to 24 percent of Match 1 and 32 percent of Match 2.[17] As a result, the remainder of the table compares our matched samples to the population by urban status. As we expected, the commonness of one's name is a proxy for urban location and socio-economic status. Men in rural areas are much more likely to have a common name; the median rural man shared his first and last name with 119 others, while the median urban man shared his name with only 10 others. Name frequency in the matched samples is substantially lower than that of the population because of our unique name requirement.

[16] Mortality bias could work in either direction; on the one hand, there is a well-known, positive relationship between income and health but, on the other hand, there was a significant health penalty associated with living in urban areas in the late nineteenth century (Frank and Mustard, 1994; Haines, 2001).

[17] Norwegian households were defined by the Census as urban if their municipality of residence was considered to be a town. However, many towns contained agricultural land on their periphery. Therefore, the urban designation likely includes some households with "rural" characteristics.

Within urban areas, matched men are up to twice as likely to come from a household whose head is a merchant or who reports having some assets. In rural areas, men in our matched samples are just as likely as the population to live in households whose head was a farmer and who reported owning land. Our matched samples are demographically similar to the population in terms of age, number of siblings, and birth order.

Table 2b compares matched stayers and matched migrants to the full population of Norwegian-born men living in Norway and the United States, respectively, in 1900. Both matched migrants and matched stayers are 3 to 6 percentage points more likely to live in an urban area in 1900, perhaps because of their urban upbringing, but the gap in urban status is not as large as it was in 1865. Matched migrants are also somewhat less likely to live in a Norwegian enclave as measured by the share of the population in the county of residence that are Norwegian born. Overall, men in our matched samples earn 3 percent more than the comparable population in 1900. Although our sample is not fully representative of the population from which it is drawn, it is important to note that the direction and extent of this bias is similar in both sending and destination areas and therefore will not affect our estimates of the economic return to migration or our conclusions about migrant selection.

We address a few additional limitations of our matching procedure here. Our method will not capture migrants who anglicize their name upon arrival in the US, which is of concern if changing one's name is correlated with economic success. Following Fryer and Levitt (2004), we use the complete 1880 US Census to construct indices of a name's distinctively Norwegian content. Our name index ranges from zero to two, with a value of zero reflecting the fact that no men in the US with a given first and last name were born in Norway and a value of two assigned

to men whose first and last names are both distinctively Norwegian.[18] Men in our matched samples have index values of 1.51 to 1.60, compared to 1.37 for unmatched Norwegian men in the US in 1900, suggesting that our matched samples are, indeed, more likely to have distinctively Norwegian names. However, we find no evidence that the "Norwegian-ness" of a man's name is related to our occupation-based earnings measure.[19]

Moreover, our sample of matched migrants will not include temporary movers who returned to Norway before 1900. According to the aggregate statistics, 25 percent of the Norwegian migration flow eventually returned to Norway (Semmingsen, 1978, p. 20).[20] Return migrants may have been disproportionately drawn from the upper or lower end of the income distribution, either because unsuccessful migrants return home to lean on their familial support network or because the most successful migrants are able to build up a certain level of savings most quickly in order to return home.

The availability of an intermediate US Census in 1880 allows us to test the selectivity of the return migration flow. We identify over 25,000 Norwegian-born men in the relevant age range in the 1880 Census. We are able to locate 14 percent of these men in either the US or the Norwegian Censuses of 1900; one-third of these had returned to Norway. We compare the economic outcomes of migrants who eventually returned to Norway and those who remained in the US in 1880, when both sets of migrants were still living in the US. Figure 2 reveals few

[18] Our first name index is equal to pr(first name | Norwegian-born) / {pr(first name | Norwegian-born) + pr(first name | born elsewhere)} and likewise for our last name index. The full measure adds these two indices together.
[19] We regress ln(earnings) on the full name index and a quadratic for age for Norwegian-born men in the 1900 IPUMS in the relevant age range. The coefficient on the name index is 0.018 (s.e. = 0.017). By this estimate, the average difference in the index value of 0.2 between matched and unmatched men would translate into a 0.4 percent difference in earnings which is both small and statistically insignificant.
[20] The United States only began tracking return migration in 1907-1908. Gould (1980) reports a much lower return migration rate (6.7 percent) for Norwegians for the 1907-1913 period.

discernable differences in the occupational distributions of these two groups.[21] Men who eventually returned to Norway are slightly over-represented at the bottom end of the occupational distribution but the mean occupation score of returners and persisters are statistically indistinguishable.

4. Estimating the return to migration and the nature of self-selection

A. Comparing the mean earnings of migrants and stayers

Our matched data sets provide a unique opportunity to account for selection bias in the estimated return to migration. However, as a benchmark, we begin by comparing the earnings of all Norwegian-born men living in the United States to all men in Norway in 1900 in the relevant age range. This population-wide comparison provides a basic OLS estimate of the return to migration.

Combining all Norwegian-born men between the ages of 38 and 50 from the 100 percent 1900 Norwegian Census and the 1 percent sample of the 1900 US Census, we estimate:

$$\ln(\text{Earnings}_i) = \alpha + \beta_1(\text{Migrant}_i) + \beta_2(\text{Age}_i) + \beta_3(\text{Age}_i^2) + \varepsilon_i \qquad (1)$$

where Earnings_i denotes the mean earnings of members of individual i's occupation in 1900 in his country of residence, Migrant_i is a dummy variable equal to one if individual i lives in the United States in 1900, and Age_i and Age_i^2 are individual i's age and age-squared in 1900.[22] The US Census data are taken from the Integrated Public-Use Microdata Series (IPUMS).[23]

[21] The occupation score variable is contained in the IPUMS data set and is calculated by matching occupations to their median earnings in 1950.
[22] Over 95 percent of both US and Norwegian observations have a recorded occupation.
[23] We also try using the "year of immigration" Census variable to restrict our sample to men who were at least 18 years old at the time of immigration to exclude men who arrived in the US as children. We find qualitatively similar results for the regressions reported in Table 3 and all subsequent tables.

The coefficient of interest in equation 1 is β_1, which measures the difference in the earnings of migrants and non-migrants. The first column of Table 3 shows that Norwegian migrants to the United States earned 60 log points (82 percent) more than men living in Norway in 1900.[24] Columns 2 through 4 reproduce the OLS estimates from equation 1 for our matched samples. The implied return to migration of 59-62 log points (80-85 percent) is similar to the 82 percent return to migration estimate for the population as a whole, despite the distinctive features of the matched samples reviewed in the previous section. The third column of Table 3 assigns US migrants in the first matched sample the average earnings for their occupation from the 1915 Iowa Census (appropriately deflated), which is more representative of the urban/rural composition of Norwegian migrants. The implied return to migration of 57 log points (77 percent) suggests that the baseline estimates may be overstated by up to five log points due to the data's urban bias.

We conclude that our matched samples produce OLS estimates of the return to migration that are similar to those found in the full population. In the remainder of the paper, we present results based on the matched samples.

B. Comparing the occupational distribution of migrants and stayers

We begin by comparing the occupational distributions of Norwegian migrants to men who remained in Norway in 1900. Differences between migrants and stayers could be due either to initial migrant selection or to occupational upgrading (or downgrading) in the US. On the one hand, higher rates of occupational mobility in the US may have allowed migrants to climb the

[24] The return to migration increases with both age and years spent in the United States. The average 36 year old migrant earned 53 log points (71 percent) more than his counterpart in Norway, whereas the average 50 year old migrant experienced a return of 65 log points (91 percent).

occupational ladder, while, on the other hand, migrants may have lacked the US-specific skills necessary to hold a highly-paid occupation.

Figure 3 presents the occupational distributions of migrants and stayers, with occupations arrayed from lowest- to highest-paid according to the average US earnings in that occupation.[25] We omit farmers, the largest occupational category, for reasons of scale but results are qualitatively similar when farmers are included. For men born in urban areas, migrants are more likely to hold low-paying jobs such as day laborer or servant, while the occupational distribution for men remaining in Norway is skewed toward higher-paying jobs (for example, merchants). Men born in rural areas are employed in similar jobs in both countries.

The occupational distributions in Figure 3 assign all occupations a common skill price (in this case, US earnings). After doing so, the large return to migration presented in Table 3 disappears for all men and become substantially negative for urban-born men. The negative "return" to migration of 17 log points for the urban sub-sample suggests the presence of either occupational downgrading or negative selection or both. We argue that the pattern is unlikely to be due to occupational downgrading alone. Any disadvantage resulting from a lack of US-specific skills that would lead migrants to hold jobs below their true skill level would likely erode over time. In contrast, we find a similarly negative "return" to migration for migrants who had been in the US for at least 20 years using this common skill price approach (18 log points).

C. *Direct measures of family background in migrant and non-migrant households*

[25] Chiquiar and Hanson (2006) conduct a similar exercise for Mexican migrants to the United States using the 2000 Census. They assign migrants the earnings that they would have received, given their education and experience level, if they had remained in Mexico. We use US earnings, rather than Norwegian earnings, because the US earnings data are richer, reflecting nearly 200 occupational categories.

Ideally, we could disentangle migrant selection from occupational downgrading by comparing the occupations held by migrants and stayers in Norway before migration took place. However, the only pre-migration outcomes we have for migrants are from the 1865 Census when these men were, on average, 8.5 years old. Instead, we compare the economic outcomes of fathers in the 1865 Norwegian Census and the 1886 Land Registers. We subdivide households with at least one matched member into migrant and non-migrant households. Note that households with one (or more) matched migrants can be neatly classified as "migrant" households; however, we cannot rule out the possibility that unmatched sons in "non-migrant" households moved to the US. Given this caveat, any observable differences in fathers' outcomes will understate the true difference between these two groups.

Table 4 compares the occupations, asset holdings, and property tax values of the heads of migrant and non-migrant households. We find that, in urban areas, migrants are drawn from the lower end of the economic spectrum. The heads of migrant households are 4.3 percentage points less likely to hold an occupation with above-median earnings and also 4.3 percentage points less likely to own assets, defined as owning a business or serving as a master craftsman in an artisanal workshop. In rural areas, the patterns of selection are less clear. Migrants are no less likely to hail from a household whose head holds a job with above-median pay. However, the heads of migrant households are 2.5 percentage points less likely to own assets. In rural areas, asset ownership is nearly synonymous with land ownership. As a result, heads of migrant household are 1.9 percentage points less likely to match to the Land Register of 1886. Conditional on matching to these tax records, though, we find no difference between migrant and non-migrant households in the property tax bill, suggesting that, among those households that did own land, the size of landholdings were nearly equivalent. These direct measures of family

background confirm the previous pattern of negative selection among migrants from urban areas. For rural migrants, the family background measures show no evidence of positive selection and, in some cases, are consistent with negative selection.

D. Comparing migrant and non-migrant brothers within households

In this section, we explore household-level selection into migration by contrasting two estimates of the return to migration: OLS estimates conducted with all men in the matched samples and estimates obtained solely by comparing migrant and non-migrant brothers within households. Within-brother estimates will eliminate bias due to aspects of family background that are correlated both with the probability of migration and with labor market outcomes later in life. By comparing OLS and within-household estimates, we can infer the direction and magnitude of this form of selection into migration.

More formally, consider decomposing the individual error term (ε_i) in equation 1 into two components: $\alpha_j + \nu_{ij}$, where α_j is shared between siblings in the same household *j* and ν_{ij} is idiosyncratic to individuals. Adding household fixed effects to equation 1 will absorb the household portion of the error term (α_j). If households that send migrants to the United States are negatively selected on wealth or social connections, we would expect the within-household estimate to be greater than the OLS coefficient ($\beta^H > \beta_1$) and vice versa if households are positively selected.[26] It is important to note that we cannot interpret β^H as the "true" return to migration if the individual component of the error term (ν_{ij}) is also correlated with earnings.

Table 5 uses our matched samples to compare between- and within-household estimates of the return to migration. In order to contribute to the within-brothers estimation, a household

[26] See Griliches (1979); Altonji and Dunn (1996); Aaronson (1998); and Sacerdote (2004) for examples of within-sibling estimates in other contexts. Ashenfelter and Krueger (1994), Behrman, Rosenzweig and Taubman (1996) and Behrman and Rosenzweig (2002) use pairs of identical twins to estimate the returns to schooling.

must contain at least two members whose names are unique enough to allow them to match between 1865 and 1900. We begin in the first row of each panel by conducting OLS on this restricted sample. Men born in rural areas experienced a higher return to migration than did men born in urban areas. In Match 1, the estimated return to migration is 66 log points (93 percent) for men born in rural areas as compared to only 35 log points (42 percent) for men born in urban areas.[27]

The second row in each panel adds household fixed effects. We find strong evidence of negative selection across households for migrants leaving urban areas. In our urban sample, the return to migration estimates *increase* by 20 to 30 percent when restricting our comparison to brothers in migrant households. This pattern suggests that the migration flow from Norwegian cities and towns was drawn from households with either lower average ability, fewer connections, or less wealth. In contrast, the estimated return to migration falls slightly when comparing migrants who originated in rural areas to their brothers who remained in Norway. We conclude that the direction of migrant selection varied by place of origin. While migrant households originating in rural areas are nearly representative of the population, migrant households in urban areas appear to be negatively selected.

E. Birth order as an instrument for migration

Comparing the within-brothers and OLS estimates of the return to migration reveals selection in the type of households that sent migrants to the US. However, even within households, brothers differ in unmeasured personal attributes (ν_{ij}). In this section, we turn to a complementary instrumental variables estimation that can address both across- and within-

[27] The return to migration in this sub-sample is somewhat lower than in the matched samples as a whole, perhaps because households with two matched members are more likely to have a high socio-economic status.

household forms of selection. In particular, we aim to find an individual characteristic that is correlated with the propensity to migrate but is not otherwise associated with labor market potential.

Historically, Norway relied on a primogeniture system of inheritance wherein the eldest brother stood to inherit the family assets and the corresponding obligation to care for his aging parents. In this social context, younger brothers, who had to "make their own way" in the world, may have been more likely to migrate to the US. In his detailed social history of migration from western Norway, Gjerde (1985) argues that migration was one solution for younger siblings who were constrained by the "system of primogeniture…[under which] they could be nourished and remain on the farm, but they could not marry until they acquired livelihoods that would sustain new families" (p. 86). Consistent with this qualitative view, we find that oldest brothers were more likely to remain in their municipality of birth and work as owner-occupier farmers, two characteristics that are consistent with having inherited the family farm.[28]

We examine the relationship between birth order and migration status in a difference in differences framework. Specifically, we subdivide men's childhood households into those with and without assets in 1865 and compare the migration rates of oldest brothers to brothers further down the birth order. Our first stage equation relates the probability of migration to the interaction between being an eldest brother and living in a household with assets:

$$Migrant_{ij} = \alpha + \gamma_1(oldest_{ij} \text{ x } assets_{ij}) + \gamma_2(oldest_{ij}) + \gamma_3(assets_{ij}) + \Gamma_1 A_{ij} + \Gamma_2 F_j + \Gamma_3 P_j + \varepsilon_{ij}. \qquad (2)$$

[28] In 1900, 27 percent of oldest brothers from the 1865 Census lived in their municipality of birth compared to 23 percent of other men. Similarly, 32 percent of oldest brothers versus 28 percent of other men were owner-occupier farmers. Both differences are highly statistically significant. See Guinnane (1992) and Wegge (1999) for empirical work on the relationship between inheritance systems and immigration in other European contexts.

We control for dummy variables for single year of age (A_{ij}), total number of siblings in the household (F_j) and province of residence in 1865 (P_j). To minimize measurement error in the designation of oldest brothers, the sample is limited to men whose mothers were young enough for a (near)-complete household structure to be observed in the Census.[29] For this analysis, we focus on men born in rural areas, as primogeniture was not a strongly held custom in cities.

Our reduced form equation relates birth order and childhood household assets to earnings later in life:

$$Earnings_{ij} = \alpha + \zeta_1(oldest_{ij} \times assets_{ij}) + \zeta_2(oldest_{ij}) + \zeta_3(assets_{ij}) + Z_1A_{ij} + Z_2F_j + Z_3P_j + \varepsilon_{ij}. \quad (3)$$

The contemporary literature on birth order leads us to expect that $\zeta_2 > 0$, or that older siblings out-earned their younger counterparts (Black, Deveraux and Salvanes, 2005; Price, 2008). However, by receiving an inheritance, older brothers were tied to the land and may have missed the opportunity to reap the return to moving to the US; therefore, we expect that the coefficient on the interaction between birth order and household assets will be negative ($\zeta_1 < 0$).

An instrumental variables estimator of the return to migration follows naturally from this framework. The Wald estimator is simply ζ_1/γ_1, or the reduction in earnings associated with being an oldest brother in a household with assets scaled by the associated decline in the probability of migrating to the US. The identifying assumption underlying this instrument is that birth order does not differentially influence labor market outcomes in households with and without assets beyond its effect on the probability of migration. However, older brothers in

[29] We restrict the sample to men whose mothers were 42 years old or less in 1865. This cut-off was selected according to the following logic: (1) Age at first birth was high in Norway during this period; in the 1865 Census, only 13 percent of women had a (surviving) child by the age of 23. (2) Furthermore, children lived with their parents until their late adolescence; 91 percent lived in their childhood until at least the age of 19. Together, these two facts imply that household structure would be incomplete for only 1.2 percent of 42-year old mothers in 1865 (= 0.13 with child by age 23 x 0.09 who left home by age 19). Our results are robust to increasing the age cut-off to 45.

households with assets may experience the direct benefits of receiving an inheritance as well as any complementary investments in human capital that their parents may provide. In this case, our IV estimates of the return to migration will be biased downward. We test the identifying assumption below by considering the effect of birth order on labor market outcomes among men who remain in the Norway.

The first panel of Table 6 examines the effect of birth order and household assets on the probability of migrating to the US. In households without assets, eldest brothers are slightly less likely than their younger counterparts to move to the US but this difference is not statistically significant. In contrast, older brothers in wealthier households are 4.9 percentage points less likely to migrate than are brothers further down the birth order (with migration rates of 9.0 percent and 13.9 percent, respectively). The full difference in differences suggests that, relative to households without assets, oldest brothers in households with assets are 3.2 percentage points less likely to migrate than their younger brothers.

The second panel of Table 6 compares the earnings of older and younger brothers by household type. As we predicted, absent the mobility restrictions associated with inheriting land, older brothers earn slightly more than their younger counterparts later in life (1.2 percent). However, in households in which the eldest brother stands to inherit land or a family business, older brothers actually earn 2.7 percent *less* than their younger siblings. The earnings penalty is not consistent with older brothers from these households receiving special advantages or additional investments associated with inheriting land. Rather, we conjecture that this earnings gap reflects the return to migration that their more rootless younger brothers were able to enjoy. Overall, relative to households without assets, older brothers earn 4.0 percent less than their younger siblings. Taken together, these estimates imply a return to migration of 125 percent (= -

0.040/ -0.032). The OLS estimate of the return to migration for this selected sample of rural men is 92 percent. The larger IV coefficients suggest that the simple earnings comparison is biased downward by around 30 percent by negative selection.

The final row of Table 6 compares the effect of birth order by household type among men who remained in Norway in 1900. In this case, any differences in earnings cannot be due to migration to the US. In all households, older brothers earned slightly more than their younger counterparts but the positive effect of birth order on labor market outcomes is stronger among households without assets. Again, we suspect that whatever advantage came with inheriting land is outweighed by the associated decline in mobility, in this case internal mobility from rural to urban areas. When we restrict the sample to men who still lived in their province of birth in 1900, the difference in birth order effects between households with and without assets disappears.

Overall, we find that younger siblings are more likely to migrate to the US, particularly from households with assets. Larger estimates of the return to migration in an instrumental variables context provide further evidence of migrants' negative selection from the sending population.

5. Conclusion

We know surprisingly little about how migrants during the age of mass migration were selected from the European population and the economic return from their journey. In this paper, we construct a unique data set of Norwegian-to-US migration and use two alternative approaches to estimate the return to migration in the presence of selection into migration. The first method compares the earnings of men who moved to the US and their brothers who stayed behind in

Norway. This approach eliminates the component of migrant selection that takes place across households. Our second approach uses birth order to instrument for migration. Younger brothers in rural areas were more likely to migrate to the US, perhaps because their older siblings inherited the family farm under Norway's system of primogeniture. The IV estimates eliminate selection both across and within households.

We estimate returns to migration from Norway to the US that range from 60 percent for men leaving urban areas to 120 percent for men leaving rural areas. The return to migration in the late nineteenth century appear to be substantially smaller than the 200-400 percent return that Hanson (2006) estimates for the migration from Mexico to the United States today. The contemporary return to migration may be higher than in the past because of the sizeable bureaucratic costs of legal immigration and the comparable costs for the undocumented of evading detection, both of which reduce the supply of immigrants to the country. In the late nineteenth century, the border was open to all prospective migrants and, therefore, the return to migration were relatively low.

We find that migrant selection generates substantial bias in the naïve estimates of the return to migration. Both across- and within-households, men who faced poor economic prospects in Norway were more likely to migrate to the United States. Due to this negative selection, the naïve return to migration is underestimated by 20-30 percent. The fact that migrants to the US appear to have been drawn from the lower end of the occupational distribution is consistent with high rates of occupational mobility in the United States relative to Norway in the late nineteenth century (Ferrie and Long, 2004; Ferrie, 2005). A standard Roy model of migration, as in Borjas (1987), predicts that men at the lower end of the occupational distribution would have more to gain by moving to the New World – and that is indeed what we

find. The fact that European migrants, when unhindered by entry restrictions, were negatively selected from the sending population, may explain why some countries explicitly select for more skilled applicants in their immigration policies today.

Data Appendix

Our earnings measure is based on assigning men the mean income earned by members of their occupation. Men living in the United States are matched to income data from the 1901 Cost of Living Survey while men living in Norway are matched to mean income by occupation tabulations for the year 1900 published by Statistics Norway and other sources (Haines and Preston, 1991; *Statistik Aarbog*, 1900; Grytten, 2007). We convert Norwegian wages to real, PPP-adjusted US dollars using the 1900 exchange rate and price levels reported in Grytten (2004).

These sources do not report information on earnings for a few large occupations, including farmers and fisherman and white collar workers in Norway. This appendix explains how we estimate the earnings of men in these occupations.

A. Estimating farmers' income

Standard sources do not report information on earnings for owner-occupier farmers in either the United States or Norway. We follow Mitchell, et al. (1922) in estimating the net earnings of owner-operator farmers from farm revenue and expenditures data. For the United States, we use data on farmers in Minnesota, the most common state of residence in our sample, from the 1900 Census of Agriculture. For Norway, we use data for the total value of farm products for the 1900 harvest found in the 1907 Census of Agriculture (*Jordbruksteljinga*).

Estimated earnings for farmers in the United States

	Statistics per farm
INCOME	
Value of farm products not fed to livestock	$753
Value of house rent and food/fuel produced on farm and consumed by family	$200 (*)
Gross earnings	***$953***
EXPENDITURES	
Labor, fertilizers	$98
Feed, seed, threshing	$75 (^)
Taxes	$27 (#)
Maintenance charges (building, machinery)	$62 (+)
Total	***$262***
NET EARNINGS	$691

(*) = Ratio of rent and food/fuel consumed to value of products sold from Goldenweiser (1916).
(^) = Ratio of feed, seed, and threshing charges relative to labor and fertilizers from Goldenweiser (1916).
(#) = Assume tax rate of 0.6% on total value of farm.
(+) = Assume maintenance charge (depreciation) of 0.05 on buildings and 0.15 on machinery. Values of buildings and machinery reported in 1900 Census of Agriculture.

The 1907 Census of Agriculture reports the total value of farm product, rather than average value per farm. According to the 1900 Census, total farm output in Norway is produced by 133,400 owner-operators, 73,200 farm laborers, 24,500 tenant farmers and 35,800 individuals who report being "farmers and fisherman." To estimate the earnings of owner-occupiers, we need to subtract the value added by tenant farmers and the composite "farmer and fisherman" category; farm labor is already accounted for on the expenditures side of the ledger. The average farm laborer earned $185 a year (US $1900). We assume that, with free mobility, tenant farmers would have earned the same amount as farm laborers (in expectation). Therefore, we subtract $4.5 million (=24,500 · $185) from the total value of farm product. Furthermore, we assume that men who report being "farmers and fisherman" earn a subsistence living and eat what they produce. Thus, we divide total farm product less $4.5 million by the number of owner-operators.

Estimated earnings for farmers in Norway

	Statistics per farmer
INCOME	
Value of farm products	$397 (+)
Value of house rent and food/fuel produced on farm and consumed by family (not reported)	$106 (*)
Gross earnings	$503
EXPENDITURES	$109 (*)
NET EARNINGS	$393

(+) = Unlike the US Census of Agriculture, the value of farm products is derived from transaction data, rather than farmer estimates. Therefore, we assume that the grain used on the farm to feed livestock is already excluded from the total.
(*) = We assume the same ratios as used for the US calculation.

B. Estimating fishermen's income

The 1906 Statistics Annual (*Statistik Aarbog*) reports the total value of cod, herring, mackerel, salmon, merlan, lobster and oysters sold in 1900. The 1910 volume *Gages Annuels des Domestiques et Salaires des Ouvriers* indicates that, in deep-sea fishing expeditions, fishermen typically received 35-55 of the catch. We divide this total by the 41,680 fisherman in the 1900 Census.

Estimated earnings for fisherman in Norway

	Statistics per fisherman
Value of products sold	$416
Share provided to fisherman	$145-$228. [We use $200.]
Value of direct consumption of fish	$48 (*)
TOTAL INCOME	$248

(*) Between 1830-1871, the average family spent 8 percent of their expenditures on fish (Grytten, 2004). The average Norwegian family's income was $300 (in US $1900), implying an expenditure of $24 on fish. The families of fisherman likely ate more fish than the average family. We double this value to $48.

C. White collar workers

With the exception of primary school teachers, we have yet to locate income data for white collar workers in Norway in 1900 (24 percent of the labor force). At the moment, we assign these workers the relevant income level from the United States deflated by the average Norway-US income gap. If the return to skill was higher in the United States than in Norway, this procedure will understate the total return to migration.

References

Aaronson, Daniel, "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes," *Journal of Human Resources*, 33 (1998), 915-946.

Abramitzky, Ran and Fabio Braggion, "Migration and Human Capital: Self Selection of Indentured Servants to the Americas," *Journal of Economic History,* 66 (2006), 882-905.

Abramitzky, Ran, "The Effect of Redistribution on Migration: Evidence from the Israeli Kibbutz," *Journal of Public Economics*, 93 (2009), 498-511.

Akee, Randall K. Q., "Who Leaves and Who Returns? Deciphering Immigrant Self-Selection from a Developing Country," *Economic Development and Cultural Change*, 58 (2010) 323-344.

Altonji, Joseph P. and Thomas A. Dunn, "Using Siblings to Estimate the Effect of School Quality on Wages," *Review of Economics and Statistics,* 78 (1996), 665-671.

Armstrong, Alexander and Frank D. Lewis, "Capital Constraints and European Migration to Canada: Evidence from the 1920s Passenger Lists." Manuscript, 2009.

Ashenfelter, Orley and Alan Krueger, "Estimates of the Economic Return to Schooling from a New Sample of Twins," *American Economic Review,* 84 (1994), 1157-1173.

Atack, Jeremy and Fred Bateman, "'Matchmaker, Matchmaker, Make Me a Match': A General Personal Computer-Based Matching Program for Historical Research," *Historical Methods,* 25 (1992), 53-65.

Black, Sandra, Paul Devereaux and Kjell Salvanes, "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education," *Quarterly Journal of Economics*, 120:2, (2005), 669-700.

"British Pound-Norway Krone Exchange Rate." *Global Financial Data*. <http://www. globalfinancialdata.com/>

Borger, Scott, "Self-Selection and Liquidity Constraints in Different Migration Cost Regimes," Manuscript, 2009.

Borjas, George, "Self-Selection and the Earnings of Immigrants," *American Economic Review,* 77 (1987), 531-553.

----------, "Immigration and Self-Selection," in *Immigration, Trade and the Labor Market*, John M. Abowd, and Richard B. Freeman, eds. (Chicago: University of Chicago Press, 1991).

----------, "The Economics of Immigration," *Journal of Economic Literature,* 32 (1994), 1667-1717.

----------, "Labor Outflows and Labor Inflows in Puerto Rico," *Journal of Human Capital*, 2:1 (2008), 32-68.

Chiquiar, Daniel and Gordon Hanson, "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States," *Journal of Political Economy,* 113 (2005), 239-281.

Chiswick Barry, "Are Immigrants Favorably Self-Selected?" *American Economic Review* 89 (1999), 181-185.

----------, "Are Immigrants Favorably Self-Selected? An Economic Analysis," in *Migration Theory: Talking Across Disciplines*, Caroline Brettell, and James Frank Hollifield, eds. (New York: Routledge, 2000).

Chiswick, Barry and Timothy Hatton, "International Migration and the Integration of the Labor Markets," in *Globalization in Historical Perspective*, Michael D. Bordo, Alan M. Taylor, and Jeffrey G. Williamson, eds. (Boston: National Bureau of Economic Research , 2003).

Feliciano, Cynthia, "Educational Selectivity in US Immigration: How Do Immigrants Compare to Those Left Behind?" *Demography*, 42 (2005), 131-152.

Ferenczi, Imre, and Walter Willcox, *International Migrations*. (New York, National Bureau of Economic Research. 1929).

Ferrie, Joseph, "A New Sample of Males Linked from the Public Use Micro Sample of the 1850 U.S. Federal Census of Population to the 1860 U.S. Federal Census Manuscript Schedules," *Historical Methods,* 29 (1996), 141-156.

----------, "How Ya Gonna Keep 'Em Down on the Farm [When They've Seen Schenectady]?: Rural to Urban Migration in Nineteenth Century America, 1850-70," working paper, Northwestern University, 1999.

----------, "The End of American Exceptionalism? Mobility in the U.S. Since 1850." *Journal of Economic Perspectives*, 19 (2005), 199-215.

Ferrie, Joseph and Jason Long. "Geographic and Occupational Mobility in Britain and the US, 1850-1881." Manuscript, 2004.

Frank, John W. and J. Fraser Mustard, "The Determinants of Health from a Historical Perspective," *Daedalus*, 123 (1994).

Fryer, Roland G. and Steven D. Levitt, "The Causes and Consequences of Distinctively Black Names," *Quarterly Journal of Economics*, 119:3 (2004), 767-805.

"Genealogy, Family Trees and Family History Records online." *Ancestry.com.* The Generations Network, Inc. < http://www.ancestry.com/>

Gjerde, Jon, *From Peasants to Farmers: The Migration from Balestrand, Norway to the Upper Middle West*, (Cambridge: Cambridge University Press, 1985).

Goldenweiser, E. A, "The Farmer's Income," *American Economic Review*, 6 (1916), 42-48.

Goldin, Claudia and Lawrence F. Katz, "The Returns to Skill across the Twentieth Century United States," NBER Working Paper No. 7126, 1999.

Gould, Eric D. and Omer Moav, "When is 'Too Much' Inequality Not Enough? The Selection of Israeli Emigrants." Manuscript, 2010.

Green, Alan G. and David A. Green, "Immigration and Canada's Wage Structure in the First Half of the Twentieth Century," Manuscript, 2008.

Griliches, Zvi, "Sibling Models and Data in Economics: Beginnings of a Survey." *Journal of Political Economy*, 87 (1979), S37-S64.

Grogger, Jeffrey and Gordon H. Hanson, "Income Maximization and the Selection and Sorting of International Migrants," NBER Working Paper No. 13821, 2008.

Grytten, Ola H., "A Consumer Price Index for Norway, 1516-2003." In *Historical Monetary Statistics for Norway, 1819-2003.* Øyvind Eitrheim, Jan T. Klovland and Jan F. Qvigstad, eds. Norges Banks Occasional Papers no. 35, 2004.

Grytten, Ola H., "Nominal Wages in Norway 1726-1940 by Occupation," In *Historical Monetary Statistics for Norway – Part II*. Øyvind Eitrheim, Jan T. Klovland and Jan F. Qvigstad, eds. Norges Banks Occasional Papers no. 38, 2007.

Guinnane, Timothy W., "Intergenerational Transfers, Emigration, and the Rural Irish Household System," *Explorations in Economic History*, 29 (1992), 456-476.

Haines, Michael R. and Samuel H. Preston, *Fatal Years: Child Mortality in the Late Nineteenth-Century America*. (Princeton, NJ: Princeton University Press, 1991).

Haines, Michael R., "The Urban Mortality Transition in the United States, 1800-1940," *Annales de Demographie Historique*, 1 (2001), 33-64.

Hanson, Gordon, "Illegal Migration from Mexico to the United States," *Journal of Economic Literature*, 44 (2006), 869-924.

Hatton, Timothy and Jeffrey Williamson, "What Drove the Mass Migration from Europe in the Late Nineteenth Century?" *Population and Development Review*, 20 (1994), 533-559.

----------, *The Age of Mass Migration: Causes and Economic Impact*. (Oxford: Oxford University Press, 1998).

----------, "International Migration in the Long-Run: Positive Selection, Negative Selection and Policy Williamson and Hatton," In *Labor Mobility and the World Economy*. Rolf J. Langhammer and Federic Foders, eds. (Berlin: Springer, 2006).

Hvidt, Kristian, *Flight to America*. (New York: Academic Press, 1975).

Keeling, Drew, "The Transportation Revolution and Transatlantic Migration, 1850-1914," *Research in Economic History*, 19 (1999), 39-74.

Margo Robert, *Race and Schooling in the South, 1880-1950: An Economic History*. (Chicago: The University of Chicago Press, 1990).

McKenzie, David and John Gibson, "The Microeconomic Determinants of Emigration and Return Migration of the Best and Brightest: Evidence from the Pacific," *Journal of Development Economics*, forthcoming.

McKenzie, David, John Gibson and Steven Stillman, "How Important is Selection? Experimental vs. Non-experimental Measures of the Income Gains from Migration," *Journal of the European Economics Association*, forthcoming.

McKenzie, David and Hillel Rapoport, "Self-selection Patterns in Mexico-US Migration: The Role of Migration Networks," *Review of Economics and Statistics*, forthcoming.

Mitchell, Wesley, Willford I. King, Frederick R. Macaulay and Oswald W. Knauth, *Income in the United States, Its Amount and Distribution, 1909-1919*. (New York: National Bureau of Economic Research, 1922).

North Atlantic Population Project and Minnesota Population Center. *NAPP: Complete Count Microdata*. NAPP Version 2.0 [computer files]. Minneapolis, MN: Minnesota Population Center, 2008. [http://www.nappdata.org]

Norwegian Historical Data Centre. *Land Registers 1886*. [computer files]. Troms, Norway: University of Troms, 2010. [http://www.rhd.uit.no/matrikkel/excel_e.html]

Piketty, Thomas and Emmanuel Saez, "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, 118:1 (2003), 1-39.

Price, Joseph, "Parent-Child Quality Time: Does Birth Order Matter?" *Journal of Human Resources*, 43:1 (2005), 240–265.

Roy, A.D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3 (1951), 135-146.

Ruggles, Stephen, et al., *Integrated Public Use Microdata Series: Version 4.0*. Minneapolis, MN: Minnesota Population Center, 2008.

Sacerdote, Bruce, "How Large Are The Effects From Changes In Family Environment? A Study of Korean American Adoptees," *Quarterly Journal of Economics*, 122 (2007), 119-157.

Semmingsen, Ingrid, *Norway to America: A History of the Migration*. (Minneapolis: University of Minnesota Press, 1978).

Soltow, Lee, *Toward Income Equality in Norway* (Madison: University of Wisconsin Press, 1965).

Statistiske Centralbureau, *Arbeidslonninger* (*Gages Annuels des Domestiques et Salaires des Ourvriers*). (Kristiania: I Kommission Hos H. Aschehough & Co. 1900, 1910).

Statistiske Centralbureau, *Jordbrukstaellingen I Kongeriket Norge*. (Kristiania: I Kommission Hos H. Aschehough & Co. 1909).

Statistiske Centralbureau, *Statistik Aarbog for Kongeriget Norge (Annuaire Statistique de la Norvege)*.(Kristiania: I Kommission Hos H. Aschehough & Co. 1905.)

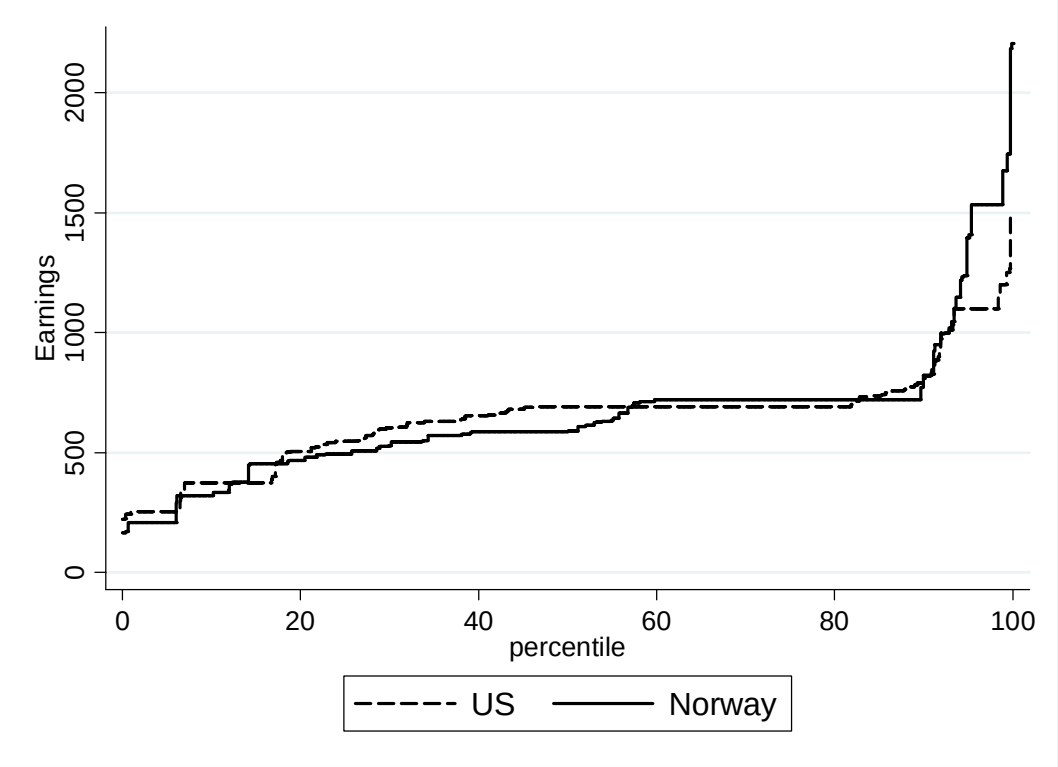
US Bureau of the Census, *Census of Agriculture*. (Washington, DC: Government Printing Office, 1900).

Wegge, Simone A., "To Part or Not to Part: Emigration and Inheritance Institutions in Nineteenth-Century Hesse-Cassel," *Explorations in Economic History*, 36(1999), 30-55.

----------, "Occupational Self-selection of Nineteenth-Century German Emigrants: Evidence from the Principality of Hesse-Cassel," *European Review of Economic History*, 6 (2002), 365-394.

----------, "Who Migrated and Who Stayed Home? The Village versus the Individual. Self-Selection of Migrants Out of the Home Population in 19th Century Germany." Manuscript, 2010.

Figure 1: Income distribution in US and Norway in 1900



Notes: US and Norwegian distributions contain all men aged 38 to 50 in the respective Censuses of 1900. Individuals are assigned the mean earnings for their occupation and are arrayed from lowest- to highest-paid occupation. The Norwegian distribution is rescaled to have the same mean as the US distribution (the actual Norwegian mean was $350 and the US mean was $643 in 1900 US dollars).

Figure 2: Comparing the occupational distributions of Norway-to-US migrants who either stay in the US or who return to Norway between 1880 and 1900

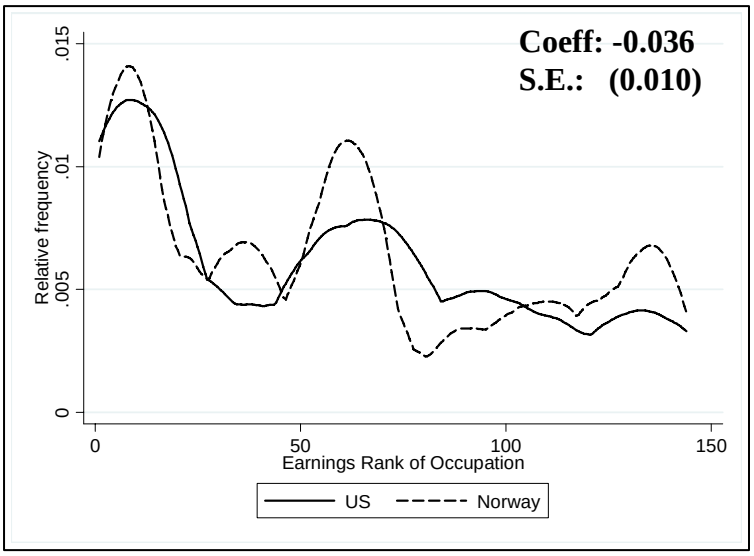
	N	Mean occupation score (1880)	% with occupation score < 12
Unmatched	21,949	14.68	35.20
Matched	3,597	14.65	36.50
- Matched (US)	2,392	14.77	35.50
- Matched (Norway)	1,205	14.39	38.40*

Notes: Return migrants are defined as Norwegian-born men observed in the 1880 US Census who are matched to the 1900 Norwegian census (N = 1,205). Persistent migrants are Norwegian-born men in the US Census of 1880 who are matched to the 1900 US Census (N = 2,392). For comparison, unmatched men are Norwegian-born men in the 1880 US Census who cannot be matched to either Norway or the US in 1900.

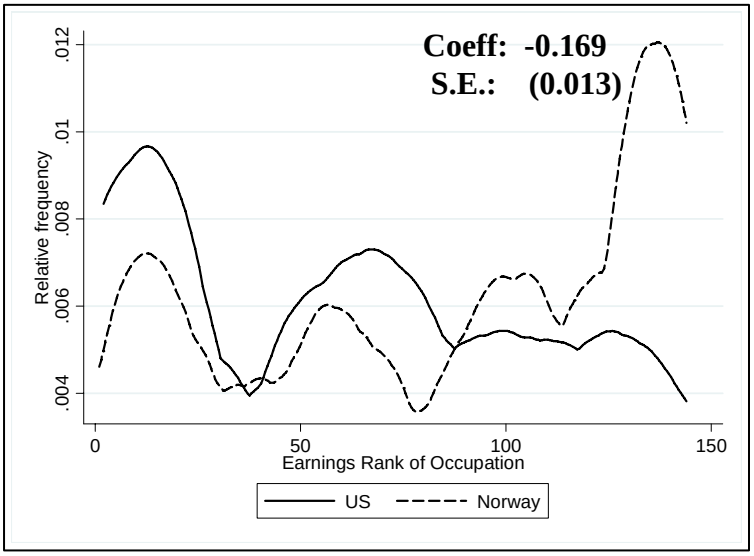
The occupation score measure, which is taken from the 1880 IPUMS sample, is constructed by ordering occupation according to their median earnings in 1950. The mean occupation score and share of the sample with an occupation score in the bottom quartile (score < 12) are reported in the accompanying table. On both measures, the differences between matched and unmatched men are not statistically significant. We mark differences between return and persistent migrants that are statistically different at the 10 percent level with an *.

Figure 3: Comparing the occupational distributions of Norwegian-born men in the US and Norway in 1900

A. Born in rural areas



B. Born in urban areas



Notes: Each figure presents the relative frequency of 144 earning categories (representing 189 distinct occupations) for Norwegian-born men in the US and in Norway. All men are assigned the mean US earnings in their occupation. Men are divided by rural or urban place of birth. Farmers, the largest occupational category, is excluded from the figure for reasons of scale. We report coefficients and standard errors from OLS regressions of ln(earnings) on a dummy for living in the US controlling for a quadratic in age.

Table 1:
Common occupations held by Norwegian-born men in the US and Norway

A. Top 10 occupations in matched sample, Norwegian-born men living in the US in 1900

Rank	Occupation	Frequency	Percentage	Earnings
1	Farmers and Planters	1690	36.99	691
2	Laborers (General)	382	8.32	373
3	Carpenters and Joiners	274	5.97	630
4	Farm Laborers	211	4.6	255
5	Painters, Glaziers, and Varnishers	112	2.44	624
6	Sailors	103	2.24	467
7	Saw and Planing Mill Workers	73	1.59	572
8	Machinists	56	1.22	736
9	Fishermen and Oystermen	55	1.2	550
10	Steam Railroad Laborers	52	1.13	462
Total		3008	65.7	

N= 4591. Occupation data collected by hand from Census manuscripts on Ancestry.com. Annual earnings by occupation data from the 1901 Cost of Living Survey reported in Haines and Preston (1991) in year 1900 dollars. Average income of owner-occupier farmers is estimated using data from the US Census of Agriculture. The Data Appendix provides details on this procedure.

B. Top 10 occupations in matched sample, Norwegian-born men living in Norway in 1900

Rank	Occupation	Frequency	Percentage	Earnings
1	General Farmers	148	17.11	393
2	Dealer, Merchant, etc.	44	5.09	837
3	Husbandmen or Cottars	28	3.24	113
4	Farmer and Fisherman	27	3.12	321
5	Carpenters	27	3.12	312
6	Farm Workers	27	3.12	175
7	Ship Masters and Captains	16	1.85	297
8	Fishermen	15	1.73	248
9	Sawyers and Sawmill Operatives	14	1.62	269
10	Boot and Shoe Makers and Repairers	14	1.62	276
Total		360	41.62	

N= 865. Historical International Standard Classification of Occupations (HISCO) occupation categories. Annual earnings by occupation data from Statistik Aarbog (1900) and Grytten (2007). Values reported in year 1900 dollars. Average income of owner-occupier farmers and fisherman is estimated using data from the Norwegian Census of Agriculture. The Data Appendix provides details on this procedure.

Table 2a:
Comparing the two matched samples to the Norwegian population in 1865

	Means			Differences	
	Population	Match 1	Match 2	Pop/Match 1	Match 1/2
Urban	0.140	0.237	0.316	0.096	0.078
	(0.341)	(0.425)	(0.460)	(0.002)	(0.005)
			A. Urban = 1		
Name frequency	368.19	17.67	3.82	-350.51	-13.85
	(1318.47)	(255.62)	(5.68)	(17.23)	(4.39)
Median	*10*	*4*	*2*		
Age	8.428	8.563	8.663	0.135	0.099
	(3.691)	(3.789)	(3.773)	(0.053)	(0.082)
Number of siblings	4.036	4.157	4.277	0.118	0.119
	(1.818)	(1.915)	(1.963)	(0.025)	(0.040)
Sibling rank	2.541	2.665	2.728	0.113	0.063
	(1.492)	(1.573)	(1.629)	(0.021)	(0.034)
Father is merchant	0.056	0.084	0.107	0.028	0.023
	(0.230)	(0.107)	(0.309)	(0.006)	(0.006)
Assets	0.137	0.195	0.233	0.058	0.037
	(0.344)	(0.396)	(0.423)	(0.005)	(0.008)
			B. Urban = 0		
Name frequency	842.46	29.76	7.35	-812.69	-22.41
	(2127.66)	(51.81)	(8.59)	(15.45)	(0.595)
Median	*119*	*14*	*4*		
Age	8.547	8.890	8.932	0.348	-0.041
	(3.657)	(3.721)	(3.708)	(0.028)	(0.051)
Number of siblings	4.145	4.206	4.254	0.057	0.048
	(1.798)	(1.812)	(1.843)	(0.014)	(0.024)
Sibling rank	2.690	2.722	2.754	0.025	0.031
	(1.546)	(1.574)	(1.611)	(0.011)	(0.021)
Father is farmer	0.435	0.421	0.449	-0.014	0.029
	(0.495)	(0.493)	(0.449)	(0.004)	(0.006)
Assets	0.626	0.629	0.633	0.002	0.004
	(0.483)	(0.483)	(0.481)	(0.003)	(0.006)

Notes: Column 1 contains means and standard deviations (in parentheses) of individual characteristics for the full population between the ages of 3 and 15 in Norway in 1865. Columns 2 and 3 reports similar statistics for the two matched samples. The number of observations underlying each column is 245,765,

24,853, and 10,758 respectively. Column 4 reports coefficients and standard errors for differences between the first matched sample and the total Norwegian population and column 5 looks at differences between the two matched samples. The number of siblings is inclusive of the individual. Oldest siblings have a sibling rank of one.

Table 2b:
Comparing matched samples to the Norwegian population in 1900 and to Norwegian migrants in the United States in 1900

	Means			Differences	
	Population	Match 1	Match 2	Pop/Match 1	Match 1/2
			A. In Norway in 1900		
Age	43.84	43.86	43.91	0.025	-0.052
	(3.72)	(3.72)	(3.73)	(0.028)	(0.045)
Married	0.875	0.856	0.847	-0.019	-0.009
	(0.331)	(0.351)	(0.359)	(0.002)	(0.004)
Children	2.941	2.823	2.740	-0.119	-0.093
	(2.481)	(2.464)	(2.462)	(0.019)	(0.030)
ln(earnings)	5.769	5.797	5.857	0.027	0.060
	(0.423)	(0.455)	(0.487)	(0.004)	(0.006)
Urban residence	0.233	0.291	0.338	0.057	0.047
	(0.423)	(0.454)	(0.473)	(0.004)	(0.007)
N	105,057	16,916	7,781	121,973	24,697
			B. In US in 1900		
Age	43.39	43.51	43.46	0.121	0.043
	(3.734)	(3.788)	(3.734)	(0.165)	(0.120)
ln(earnings)	6.384	6.421	6.433	0.037	0.022
	(0.322)	(0.309)	(0.300)	(0.013)	(0.010)
Share urban in county	0.347	0.385	0.402	0.030	0.024
	(0.370)	(0.385)	(0.390)	(0.017)	(0.013)
Share Norwegian in county	0.071	0.065	0.067	-0.006	0.002
	(0.063)	(0.064)	(0.066)	(0.003)	(0.002)
Norwegian name index	1.374	1.602	1.510	0.228	-0.129
	(0.485)	(0.371)	(0.424)	(0.008)	(0.015)
N	647	3,461	1,243	4,008	4,704

Notes: Column 1 contains means and standard deviations (in parentheses) of individual characteristics for the full population of Norwegian migrants between the ages of 38 and 50 living in United States in 1900. Columns 2 and 3 report similar statistics for the two matched samples. Column 4 reports coefficients and standard errors for differences between the first matched sample and the total Norwegian population and column 5 looks at differences between the two matched samples. The Norwegian name index is equal to the sum of the probabilities that a man is born in Norway conditional on having a given first or last name; the index ranges from zero to two. The share urban is equal to the share of the migrant's county who lives in a town with 2,500 or more residents. The share Norwegian is equal to the share of the migrant's county who were born in Norway.

Table 3:
OLS regressions of the return to migration from Norway to the US

Dependent variable = ln(earnings)

	Full population	Match 1 – COL data	Match 1 – Iowa data	Match 2 – COL data
=1 if migrant	0.603	0.620	0.575	0.586
	(0.015)	(0.008)	(0.008)	(0.014)
N	144,266	21,975	21,247	8,769

Notes: Standard errors are reported in parentheses. All regressions control a quadratic in age. The first column contains a representative sample of the population of Norwegian-born men between the ages of 38-50 in 1900 from the 100 percent 1900 Norwegian Census and 1 percent 1900 US Census sample (IPUMS). Columns 2 and 3 report estimates from the first matched sample based on an iterative matching strategy that searches first for an exact match and then for matches in a one- or two-year age band. In column 3, the US migrants are assigned earnings from the 1915 Iowa Census (appropriately adjusted for inflation). Column 4 reports estimates from the second matched sample, which instead requires that matched observations be unique within a five-year age band.

Table 4:
Economic outcomes of heads of migrant and non-migrant households

	Migrant HHs	Non-migrant HHs	Difference
Urban			
Above median occupation	0.476	0.520	-0.043
			(0.017)
Assets	0.198	0.241	-0.043
			(0.015)
N	1075	3993	5068
Rural			
Above median occupation	0.653	0.648	0.006
			(0.010)
Assets	0.639	0.665	-0.025
			(0.010)
Match to tax records	0.095	0.114	-0.019
			(0.007)
Value 1	1.755	1.788	-0.033
			(0.129)
Value 2	2.618	2.777	-0.156
			(0.245)
N	2380/226	15755/1804	18135/2030

Notes: Results for Match 1. This table compares households with at least one matched member in 1865. Households are placed in the "migrant" category if they have at least one matched migrant. Means are reported in columns 1 and 2 and differences between migrant and non-migrant households, along with standard errors, are reported in column 3. Above median occupations are those earning more than $393 per year in 1900 US dollars. We assign income levels to fathers using mean Norwegian earnings by occupation in 1900. Assets is an indicator variable equal to one for men who own a business, own land, or are master craftsman in an artisanal workshop. Tax records refers to the 1886 Land Registers. For fathers who match to the tax records, we report the value in dalers (value 1) and marks (value 2) of property taxes owed, a proxy for total land value. Sample sizes refer to the total number of matched households and to the number of matched households that can be found in the tax records, respectively.

Table 5:
OLS and within-household estimates of the return to migration
Households with two or more members in the matched sample

Dependent variable = ln(earnings); Coefficient on =1 if migrant

	Rural, 1865	Urban, 1865
Panel A: Match 1		
OLS	0.661	0.351
	(0.021)	(0.029)
Within household	0.643	0.422
	(0.027)	(0.038)
N	3,165	1,406
Panel B: Match 2		
OLS	0.622	0.286
	(0.045)	(0.041)
Within household	0.578	0.373
	(0.067)	(0.067)
N	880	654

Notes**:** Standard errors are reported in parentheses. Each cell contains coefficient estimates of log earnings on a dummy variable equal to one for individuals living in the United States in 1900. Regressions also include controls for age and age squared. Panels 1 and 2 reports results from the first and second matched samples respectively. In each panel, the first row conducts an OLS regression for the restricted sample of households that have at least two matched members in the data set. The second row adds household fixed effects. The second and third columns conduct similar analyses for men who lived in rural or urban areas respectively in 1865.

Table 6: IV estimates of the return to migration
Using the interaction of birth order and household assets as an instrument

	Assets	No assets	Difference: Assets - no assets
	A. First stage: Probability of living in US in 1900		
Oldest	0.090	0.138	-0.047
			(0.012)
Younger	0.139	0.155	-0.016
			(0.015)
Diff: Oldest - younger	-0.049	-0.017	-0.032
	(0.012)	(0.015)	(0.018)
	B. Reduced form: Ln (earnings)		
Oldest	5.826	5.862	-0.036
			(0.016)
Younger	5.854	5.851	0.003
			(0.018)
Diff: Oldest - younger	-0.027	0.012	-0.040
	(0.016)	(0.020)	(0.024)
		Implied return to migration:	125.5%
			(74.8)
	C. Testing exclusion restriction: Ln(earnings), Men living in Norway in 1900		
Diff: Oldest - younger	0.008	0.039	-0.031
	(0.015)	(0.020)	(0.024)

Notes: Standard errors are reported in parentheses. The sample includes men in the first matched sample who live in a rural area in 1865, whose mother is 42 years old or less in 1865, and who have six or fewer siblings. The first panel presents implied coefficients from the first stage regression (equation 2). The dependent variable is an indicator for living in the US in 1900. The key right hand side variables are dummies for being an oldest brother, having assets in one's childhood home, and an interaction of the two. The regressions also include dummy variables for age, number of siblings and province of residence. The second panel presents implied coefficients from a reduced form equation whose dependent variable is the logarithm of occupation-based earnings in 1900 (equation 3). The implied return to migration is a Wald estimate for which the instrument for migration is the interaction between being an oldest brother and hailing from a household with assets. The third panel reports implied coefficients from the reduced form equation conducted on the subsample of men living in Norway in 1900. There are 6,742 observations underlying the first and second panels and 5,710 observations underlying the third panel.