# Discussion Papers
## Department of Economics
## University of Copenhagen

No. 19-01

Those Who Stayed: Individualism, Self-Selection and Cultural Change during the Age of Mass Migration

Anne Sofie Beck Knudsen

# Those Who Stayed: Individualism, Self-Selection and Cultural Change during the Age of Mass Migration*

Anne Sofie Beck Knudsen[†]

January 2019

## Abstract

This paper examines the joint evolution of emigration and individualism in Scandinavia during the Age of Mass Migration (1850-1920). A long-standing hypothesis holds that people of a stronger individualistic mindset are more likely to migrate as they suffer lower costs of abandoning existing social networks. Building on this hypothesis, I propose a theory of cultural change where migrant self-selection generates a relative push away from individualism, and towards collectivism, in migrant-sending locations through a combination of initial distributional effects and channels of intergenerational cultural transmission. Due to the interdependent relationship between emigration and individualism, emigration is furthermore associated with cultural convergence across subnational locations. I combine various sources of empirical data, including historical population census records and passenger lists of emigrants, and test the relevant elements of the proposed theory at the individual and subnational district level, and in the short and long run. Together, the empirical results suggest that individualists were more likely to migrate than collectivists, and that the Scandinavian countries would have been considerably more individualistic and culturally diverse, had emigration not taken place.

**Keywords:** Culture, individualism, migration, selection, economic history
**JEL classification:** Z10, F22, O15, R23, N33

# 1 Introduction

People of Western societies are unique in their strong view of themselves as independent from other human beings (Henrich et al. 2010). This culture of individualism has roots in the distant past and is believed to have played an important role in the economic and political development of the region (Gorodnichenko and Roland, 2011a; Talhelm et al. 2014, Olsson & Paik 2016, Buggle 2017). Underlying the latter argument is evidence that differences in individualism and its counterpart, collectivism, impact processes of innovation, entrepreneurship, cooperation, and public goods provision.[1] Yet, little is known about what has influenced the evolution of individualism over time and across space within the Western world. This reflects a general gap in the literature of economics, which is becoming increasingly interested in understanding how culture shapes and is shaped by the economy.[2]

In this paper, I explore the relationship between individualism and a common example of human behavior: migration. I propose a theory, where migration flows generate cultural change towards collectivism and convergence across migrant-sending locations. To examine the relationship empirically, I take a historical perspective and study the Scandinavian experience in one of the largest migration events in modern history, the Age of Mass Migration. Here, I find that Scandinavians who grew up in individualistic households were more likely to emigrate during this period. This is in accordance with the *voluntary settlement hypothesis* in social psychology that people of individualistic mind-sets suffer lower costs of leaving existing social networks behind (Kitayama et al. 2006). From this, I expect migration to imply a reduction in individualism in migrant-sending districts, which persists over time if culture is transmitted across generations. On top of that, I expect migration to generate cultural convergence, because the related cultural change on the one hand is faster in initially more individualistic locations, and on the other hand slows down as emigration accumulates and migrant selectivity weakens. Empirical analyses of historical panels on Scandinavian individualism and emigration reveal these patterns of cultural evolution. Moreover, the cultural change that took place during the Age of Mass Migration was sufficiently profound to leave a long-run impact on contemporary Scandinavian culture.

Several circumstances make the Age of Mass Migration an ideal case for the empirical objective of this paper. During the period, millions of people left Europe to settle in New World countries such as the United States. Sweden, Norway, and Denmark experienced some of the highest emigration rates in Europe during this period, involving the departure

---

[1]See Morris et al. (1993), Greif (1994), Wagner III (1995), Chen et al., (1998), van Everdingen and Waarts (2003), Bozeman (2007), and Taylor and Wilson (2012).

[2]See review by Nunn (2012), Alesina and Giuliano (2015), and the few studies on cultural change in other traits that include Voigtländer and Voth, (2012), Guiliano and Nunn (2017), Fouka, (2018), and Bentzen (forthcoming).

of approximately 25% of their populations. Besides representing the largest migration event in Scandinavian history, global regulatory policies on migration were particularly loose at this point in time, which enables the identification of self-selective processes under limited governmental influence. In addition, the historical context allows me to study long-run cultural implications of migration in sending locations.

An empirical challenge in cultural economics is the lack of historical data, which makes detailed studies of persistence and change over time difficult. Drawing on research in the field of social psychology, I construct aggregate series and individual level indicators of individualism based on the uncommonness of first names, which has been documented to reflect a desire to stand out rather than fit in (Zweigenhaft, 1981; Twenge et al., 2010; Emery, 2013). Properties of a first name choice reflect the preferences of the name-givers, and the uncommonness of a first name thus measures the individualism inherited from home. The predetermined nature of first names provides a useful source of identification. I validate the measure using contemporary and historical indicators of individualism.

With a measure of individualism in hand, I proceed to empirically examine the elements of the proposed relationship between migration and individualism. First, I test the voluntary settlement hypothesis. For this purpose I construct a novel database of migrants and non-migrants who lived in Scandinavia just before or during the Age of Mass Migration. The Scandinavian countries were particular meticulous in registering not just their home populations but everyone that left to settle elsewhere. I am thereby able to combine historical population census records with detailed passenger lists from ships that carried migrants abroad. The result is a database of 1,253,317 Scandinavian first-time emigrants that cover 62% of the total emigration flows of the period. For a subset of these emigrants, I identify their childhood households in the population census records. Analysis of the data confirms that individuals who grew up in individualistic households were more likely to emigrate. This remains true with controls for other circumstances that may have shaped the cultural, social, and economic prospects of young individuals such as religiosity and the ownership of assets. Individualistic emigrants were furthermore less likely to seek the company of fellow Scandinavians or preserve customs from home, once arriving at their destination.

For use in analyses of aggregate district effects, I quantify the cultural shock of selective emigration as the percentage point drop in the prevalence of individualists in the migrant-sending population that occurs due to emigration. This is feasible because I have information on rates of emigration and the gap between emigrant and population individualism. Accumulating these shocks over the entire period of the Age of Mass Migration reveals an overall reduction in individualism from emigration of approximately 3.9%-points in Denmark, 10.1%-points in Sweden, and 13.1%-points in Norway.

In the second part of the empirical analysis, I analyze if the cultural shocks of emigration were in fact transmitted to new generations so that cultural change would persist

3

over time. I take advantage of the cohort structure of the historical population census records and construct a set of decennial panels that cover the period 1730-1910. Individualism is here measured across districts and decades as the uncommonness of first names in the corresponding birth cohorts. I include cohorts born before the start of the Age of Mass Migration to check that cultural differences also persisted in periods without mass emigration. Fixed-effects estimations document that the cultural shocks of emigration pushed migrant-sending district culture in a collectivistic direction and this is robust to the inclusion of control variables that capture alternative district, cohort, and emigrant characteristics.

Third, I assess the long-run cultural impact of the Age of Mass Migration on contemporary cultural differences in Scandinavia. For each historical population census, I measure individualism by the degree of uncommonness of first names among children, and I accumulate the cultural shocks of emigration that occurred since the historical census year. The empirical results show that, whereas subnational cultural differences persisted over a period of up to 170 years, the cumulative sum of shocks was associated with a substantial decline in contemporary individualism.

Finally, I present evidence that documents the presence of cultural convergence before and especially during the Age of Mass Migration. Cross-district variation in individualism was falling over time, and emigration appears to have contributed significantly to this trend. I also find support of the underlying convergence mechanism: Past individualism positively predicts the size of emigration-induced cultural shocks, but its impact diminishes as emigration accumulates over time. This is consistent with the idea that emigrant culture converged towards population culture as more collectivists joined the flows of migration.

Taken together, the empirical results of this paper show that individualism exhibited significant but imperfect persistence across localities from before the Age of Mass Migration and until today. Cultural change and convergence took place as the level of individualism impacted and was impacted by one of the largest migration events of modern history. I thereby draw attention to a mechanism of cultural change that has not been studied before but is of broad relevance. The findings may inform an otherwise charged debate on the drivers and consequences of migration, which are gaining increasing attention as present international migration flows are growing.

Below, the contribution of this paper is put in relation to the existing literature. The next section describes the historical context, and Section 3 discusses the conceptual considerations. Section 4 motivates and describes the main data sources and empirical measures. Section 5 outlines the empirical frameworks used to validate the theoretical predictions of the paper. Sections 6, 7 and 8 empirically examine the predictions regarding voluntary settlement, cultural change, and cultural convergence. Section 9 concludes.

**Related Literature**   This paper provides the most comprehensive test of the voluntary settlement hypothesis on selective migration so far produced. The hypothesis was formed as a version of the frontier hypothesis, first put forth by Turner (1920), on American individualism being the result of a long history of living on an expanding frontier that was characterized by isolation and harshness. Instead, emphasis was put on the voluntary movement of particularly autonomous and independent individuals into frontier regions, and how this self-selection shaped subsequent cultural developments (Kitayama et al., 2006; Varnum and Kitayama, 2010). Empirical evidence of these theories has so far been limited to documenting correlations between levels of individualism and a history of receiving migrants.[3] More compelling is the work, parallel to this study, by Bazzi, Fiszbein, and Gebresilasse (2018), who track the American frontier over time and analyze its individualistic environment in detail. The present paper distinguishes from the mentioned studies, as I narrow the focus on the individualistic selection of migrants and examine the phenomenon in the original home locations of the migrants. I also investigate its consequences for aggregate cultural change in migrant-sending localities, which represents a natural yet unexplored extension of the hypothesis.

The results of this paper contribute more generally to the understanding of determinants and consequences of migration. Studies of migrant selection have traditionally focused on the impact of economic circumstances and networks (see Abramitzky et al., 2017, for review), and of particular relevance is the work by Abramitzky et al. (2013) that documents how historical Norwegian emigrants were negatively selected in terms of economic prospects at childhood. An exception is Jaeger et al. (2010) who find evidence of a link between risk attitudes and migration, based on ex post characteristics of migrants and stayers in the contemporary German population. The literature on implications of emigration on migrant-sending communities is also related to this study. This includes the literature on modern globalization and brain drain (Docquier and Rapoport, 2012) and recent studies that link historical Swedish mass emigration to the rise of labor unions and innovative activity across Swedish localities (Anderson et al., 2017; Karadja and Prawitz, forthcoming).

Another relevant strand of literature is the field on long run development, which is concerned with factors that have ultimately shaped the world as it is today. Among other factors, this literature focuses on the role played by culture, and it is identifying the historical origins of contemporary cultural differences and their impact on economic and institutional developments. This includes a growing number of studies on individualism (Greif, 1994; Gorodnichenko and Roland, 2011a, 2011b, 2017; Olsson and Paik, 2016;

---

[3]Studies in social psychology have found that recently settled countries and regions in the United States and Japan are more individualistic than others (Kitayama et al. 2006, Varnum & Kitayama 2010), that within-country residential mobility is higher in individualistic countries and locations (see Oishi 2010 for a review), and that college students with a history of migration identify with more individualistic values (Oishi, Lun & Sherman, 2007).

Buggle, 2017; Bazzi et al., 2018). There are less studies on the processes of cultural change (with examples being Voigtländer and Voth, 2012; Guiliano and Nunn, 2017; Fouka, 2018; and Bentzen, forthcoming).[4] This paper contributes to this literature by providing insights on a specific channel of cultural change, migration, and its reciprocal relationship with the cultural trait itself in the short as well as the long run.

One reason for the gap in the literature regarding cultural change is the lack of long series of cultural indicators. Like any other cultural trait, quantitative measures of individualism and collectivism are usually constructed based on surveys. Values linked to individualism are independence and personal achievement, while collectivism is linked to values of conformity, tradition, and the preservation of group harmony. A significant limitation of survey based measures is that they only cover the past 30 years and cannot be replicated retrospectively. Here one has to rely on alternative sources of cultural information. Various examples of human behavior reveal this kind of information, for instance how we communicate with each other, consume and produce arts, and raise our children. Efforts in quantifying this information is ongoing in various disciplines of the social sciences.[5] Despite the amount of available historical material, larger scale historical datasets constructed in this manner are few. The study by Barry et al. (1959) of differences in child rearing across ethnic societies is a notable exception. In economics, important examples include Becker and Pascali (forthcoming) and Voigtländer and Voth (2012) who capture anti-Semitic beliefs in book titles and by the prevalence of pogroms, and Murrell et al. (2011) who identify the diffusion of Whig culture in 17th century England by the use of words related to freedom in titles of written publications. Similarly, research by Galor, Özak, and Sarid (2017, 2018) study the origins and economic consequences of global language structures.

## 2    Historical Context

Between 1850 and 1920 more than 55 million Europeans emigrated to the New World, with the majority settling in the United States (Taylor and Williamson, 1997). Migration was facilitated by a near absence of regulatory migration policies and the shift from wind to steam driven ship transportation. Scandinavia experienced some of the highest emigration rates in Europe during the time period. According to the numbers in Figure 1, emigration peaked in the 1880's where around 8% of the Swedish and Norwegian population left in a

---

[4]Guiliano & Nunn (2017) show that locations subject to climatic instability experience more rapidly changing cultures. Fouka (2018) documents that the implementation of forced assimilation policies in the US impacted the cultural integration of German immigrants negatively, and Bentzen (forthcoming) finds that religiosity increases in locations immediately after being hit by earthquakes. When studying the long term persistence of anti-Semitic beliefs, Voigtländer & Voth (2012) show that growing city size and trade weakens persistence.

[5]For a review of the psychological studies on contemporary cultural products see Morling & Lamoreaux (2008).

single decade. While Danish emigration flows were smaller, nearly 15% of the population left altogether. Total emigration amounted to around 38% and 26% in Norway and Sweden respectively - numbers only matched by Ireland and Italy (Hatton and Williamson, 1992). The intensity of emigration varied considerably across Scandinavian districts as can be seen from the map in Figure 1, where major emigration ports are also shown.

Figure 1: Scandinavian Emigration 1850-1920
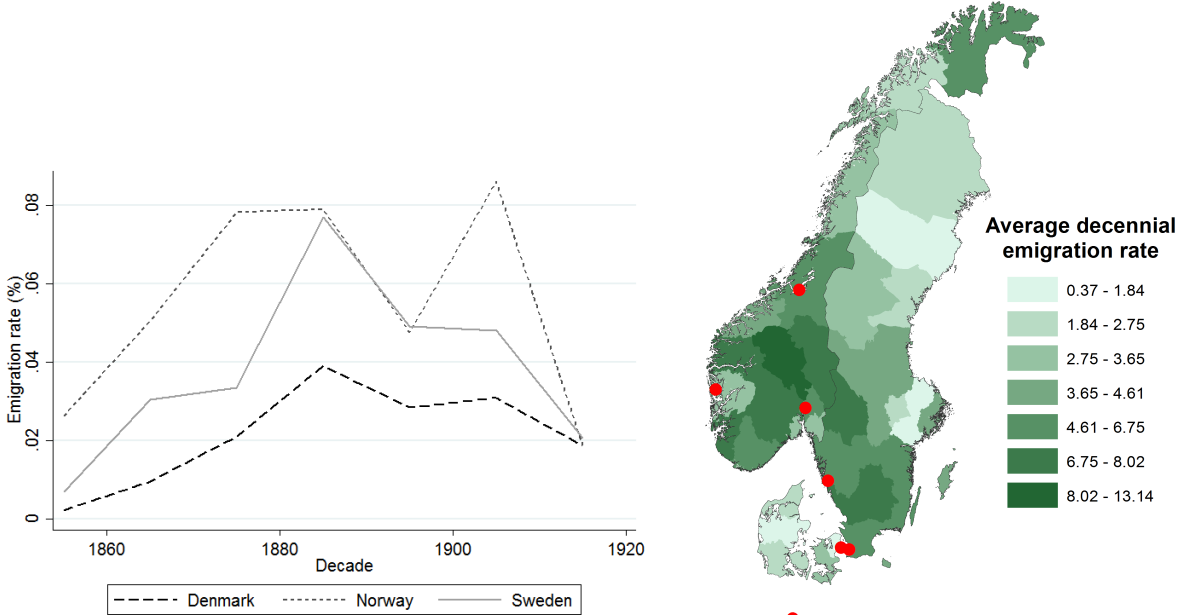


Notes: Decennial rates of emigration over time (left) and across districts (right). See section 3.3 for a description of data sources. Districts represent the current subnational divisions and refer to 21 Swedish counties (län), 19 Norwegian counties (pre-2018 fylke), and 10 Danish provinces (landsdele). Major emigration ports for the Scandinavian emigrants (the red dots) include Gothenburg, Malmö, Oslo, Trondheim, Bergen, Copenhagen, and Hamburg.

In related empirical research, the causes of migration during this time period have especially been linked to economic circumstances. Years of failed harvests, structural transformation, and economic inequality generated a push of particularly low skilled and agricultural workers who benefitted from better job opportunities and access to cheap land in the New World (Hatton and Williamson, 1992; Abramitzky et al., 2012, 2013; Karadja and Prawitz, *forthcoming*). Yet, historians agree that migrants were motivated by more than hopes of escaping poverty (Semmingsen and Haugen, 1978; Indseth, 2006; Sønnichsen, 2013). Stories on the *'American Dream'* and the view of the United States as the *'Land of Opportunities'* were core to the migration discourse. Private letters, diaries, and newspaper articles of the time reveal that ideas of personal freedom and social equality embodied in the American society were of great value to the migrants. In the United States, people were free to pursue own goals. The countries of the New World thereby represented a contrast to Scandinavia that, besides offering poor economic prospects, were characterized by social rigidity, religious intolerance, and limited suffrage.[6] The historical

---

[6]People of minority religious denominations were for instance among the first to emigrate, including

evidence suggests that cultural motivations were part of the migration decision in addition to economic incentives. Since individualism was generally on the rise in Western countries in this time period (Macfarlane, 1991), cultural motivations could even have played a role in triggering the Age of Mass Migration.

# 3  Conceptual Considerations

The act of migration involves leaving familiar and established social networks behind. Depending on the value placed on these networks, and circumstances at destination, the migration experience is potentially associated with great psychological distress (Eisenbruch, 1991; Bhugra and Becker, 2005). This observation has motivated research to study non-economic drivers and consequences of migration. In social psychology, one of the more prominent theories, the *voluntary settlement hypothesis*, proposes that people of a stronger individualistic cultural beliefs are more likely to migrate and settle somewhere new than their collectivistic counterparts (Kitayama et al., 2006).

An individualistic culture is commonly defined as centering on the belief that human beings constitute autonomous units in loosely-knitted social networks. Here, individual identity is derived from inner attributes, abilities, and personality traits. In contrast, a collectivist culture emphasizes the individual as embedded in larger and interconnected networks, where identity is derived from social relations. This difference in self-concepts translates into different sets of cultural values, which includes the emphasis of distinction over conformity or the pursuit of individual versus collective goals in individualistic and collectivistic cultures, respectively (Heine and Ruby, 2010).

Based on these characteristics, it seems likely that this cultural trait plays an important role in the migration decision. First, people of an individualistic mindset suffer a lower cost of abandoning existing social networks - both in terms of loss of identity and support system. Second, they may experience a personal satisfaction from doing something out of the ordinary that requires and shows independent effort. While migration is certainly driven by other factors than individualism and collectivism, one would expect individualists to respond more readily in a setting of various push and pull effects.

If people migrate based, in part, on individualistic cultural values, migration will have implications on the overall evolution of cultures. Emigration must be associated with an immediate reduction in the prevalence of individualists in the migrant-sending population. This is a simple composition effect that I refer to as the *cultural shock of emigration*. The magnitude of the shock depends on the rate and selectivity of emigration. For a shock to have persistent cultural effects, channels of intergenerational cultural transmission need to be sufficiently efficient. Theoretical and empirical research has established that individual

---

Danish and Norwegian Mormons and members of the Swedish free church movement. Some were fleeing actual religious persecution.

preferences and values are passed along from parents to children and influenced by other members, so-called role models, of society (Cavalli-Sforza and Feldman, 1981; Bisin and Verdier, 1998, 2000, 2001; Dohmen et al., 2012; Chowdhury et al. 2018). In a related study I document the presence of these channels in the cultural transmission of individualistic versus collectivistic cultural values in the same historical setting as the present study (Knudsen, 2018).

With channels of intergenerational cultural transmission in place, emigration is expected to push the cultures of migrant-sending locations in a relatively more collectivistic direction. The intensity with which this happens need not be proportional to the initial shock. The shock may alter both the supply of cultural role models and the benefits from being one or the other cultural type. An example is if parents of distinct cultural values become a minority.[7] Such mechanisms may strengthen or weaken the initial impact of the shock.

Another expectation is that emigration generates subnational cultural convergence. Locations that are initially inhabited by more individualistically minded people will due to self-selection experience larger out-flows of emigrants and thereby a larger drop in individualism than their collectivistic neighbors. Over time, as these locations become relatively more collectivistic, fewer people emigrate because of their individualistic traits. Instead, when migrant networks abroad are created and a common acceptance and experience with emigration spreads at home, the social cost of migration diminishes and more collectivists are expected to join the migration flows.

The above considerations may be summarized in three testable predictions:

**Prediction 1** (voluntary settlement). *People that emigrated in the Age of Mass Migration were more individualistic than the people that stayed behind.*

**Prediction 2** (cultural change). *The reduction in home population individualism directly associated with selective emigration persisted over time through channels of intergenerational cultural transmission.*

**Prediction 3** (cultural convergence). *The speed of cultural change caused by selective migration was faster in individualistic locations and diminished across all locations as emigration accumulated, generating cultural convergence over time.*

---

[7]Parents of distinct cultural values may alter their socialization efforts when the cultural environment changes and they become a cultural minority. On the one hand minority individualists may be more determined in ensuring the survival of their (individualistic) traits. On the other hand, one could imagine a social cost of being individualist in an increasingly collectivistic environment, whereby more individualists may become collectivists. Likewise, finding a spouse that shares one's individualistic traits can be more difficult.

# 4 Data Sources and Measurement

## 4.1 Population Census Records

One of two main data sources for this paper is a set of historical census records on the Scandinavian population over the period 1845-1910. This paper benefits from the on-going work by national statistical offices and academic institutions with digitizing these archives.[8] Of the digitized population census records, I use those on entire populations that contain information on residence and birthplace district. The resulting set of censuses cover the years 1845 and 1880 for Denmark, 1880, 1890 and 1900 for Sweden, and 1865, 1900, and 1910 for Norway. Besides documenting demographic and economic circumstances of all individuals and their respective households, the censuses include information on full names from which cultural indicators can be constructed (see section 4.3).

I define subnational districts according to present-day, EU standard NUTS 3 boundaries. These include the 21 Swedish counties (*län*), 19 Norwegian counties (*pre-2018 fylke*), and 10 Danish provinces (*landsdele*) that are illustrated in Figure 1. For Sweden and Norway these districts are associated with local government, and they go back to the historical period of this paper. The Danish provinces are used for statistical purposes and carry no governmental authority. They reflect a collection of more or less stable county borders.

## 4.2 Emigration Data

I construct two datasets on the subnational emigration patterns in Scandinavia during the Age of Mass Migration: One with information at the individual migrant level and one at the subnational district level.

### 4.2.1 A Database of Scandinavian Migrants

I use detailed passenger lists to create a database of a majority of the emigrants that left Scandinavia in this time period. In 1868 and 1869 national laws were passed in each Scandinavian country, requiring travel agencies to record personal details of all passengers leaving on ships bound for international destinations. These laws were inspired by a similar system in Hamburg and implemented to ensure the proper conduct of passengers' money and onboard safety. The quality of this information can be trusted to be high as it was reported to the national police, whose control and approval was required for the agencies to continue their business (Bender, 2007).

---

[8]The Swedish and Norwegian data is made available by the North Atlantic Population Project at the Minnesota Population Center (2017), while the Danish data is obtained through the Danish National Archives.

From the most comprehensive digitized versions of these records, information on emigration year and last place of residence is collected for 1,253,317 Scandinavian first-time emigrants over the period 1869-1920. Foreigners and visitors are identified and removed from the database. For 88% of these migrants, information on full name and gender is also available, and so is information on birthplace for 25% of these.[9] The passenger lists do not cover all emigrants of the period, but compared to official counts of emigrants described below, 62% are accounted for in the migrant database of this paper. Around 300,000 emigrated before the first data point and are not included. The database is nonetheless the most comprehensive of its kind.

To test the overall reliability of the passenger list data, I compare it to the corresponding population census records. In particular, I compare the change from one population census to the next in the number of people that share the same first name, birth decade, and residence district to the number of emigrants with the same characteristics leaving between these census years. The results in Appendix Table A.9 show that emigration significantly and strongly predicts actual change along these dimensions in the population. This finding speaks to the high quality of the migrant database, especially in light of unobserved effects of mortality and internal migration that hamper a clean empirical test.

I add internal out-migrants to the database for comparison. These are identified as individuals that lived outside their birth district in any of the population censuses, which was 10-15% of the population.

### 4.2.2 Aggregate Migration Flows

Part of the empirical analysis focuses on the relationship between emigration and cultural change at the subnational district level. Rather than relying entirely on passenger list data for that purpose, I use additional official emigration statistics from Norway and Sweden. These data were constructed at the time based on church records, and they are considered a more complete source of information on *aggregate* migration flows than the passenger list data, where a significant number of emigrants are unaccounted for. The underlying individual level church record data is not available in any complete format and is therefore not used in this paper. Reassuringly though, a comparison between the two data sources in Appendix Table A.10 reveals a high degree of within-country correlation over time. This also holds when dropping passenger list entries that lack information on

---

[9]The main passenger list data sources are: The Copenhagen Police Emigration Protocols 1869-1908 (Denmark), "Emihamn" Emigration Records 1869-1950 (Sweden), and Norwegian regional emigration lists 1869-1930. The "Emihamn" records contain no information on birthplace. Instead two additional sources are used: "Emibas Göteborg" and "Emibas Värmland" that cover emigrants that resided in the Göteborg or Värmland districts at emigration. Information on residence and birthplace is originally stored in more or less precise strings in the source data, which I have transformed to codes identical to those used in the population census records.

full name and birth decade.

In Denmark, where no official subnational emigration statistics exist, aggregate passenger list numbers are used instead. This solution is not flawless since the Danish passenger lists lack a few important sources of information, including Mormon emigration and documentation from travel agencies in Horsens, Vejle, and Ribe (Bender, 2007). I therefore calculate two alternative Danish district emigration variables. One uses passenger list counts for the period 1850-1920 from ships leaving Hamburg in Germany that carried a total of 2/3 of all Danish emigrants (Bender, 2007). Another combines the Danish passenger list counts with the Hamburg data from before 1869, Mormon data from 1872-1887, and the total number of ship tickets sold by the Ribe, Vejle, and Horsens agencies.[10] Appendix Table A.11 shows that the number of Hamburg list emigrants is not particularly correlated with the passenger numbers preserved by the Danish police, which underlines the importance of checking future empirical results with both variables.

No data on internal flows of migration from this period exists, and the focus thus remains on emigration in the aggregate analyses.

## 4.3   Measuring Individualism Using First Names

The usual quantitative indicators of cultural values that are based on questionnaires and experiments do not exist historically. Instead, I construct indicators of individualism from the distribution of first names in the historical district populations. This approach is based on research in sociology and psychology that argues that the commonness of first names can be seen as reflecting the core difference between collectivistic and individualistic cultural values. While a common first name suggests a desire of the name-givers for their child to conform and fit in, an uncommon first name signals independence, originality, and the wish to stand out and differentiate oneself from the surrounding social environment (*see references below*).

First names are generally acknowledged to closely reflect the cultural and social preferences of the name-givers (Mateos, 2013). Across societies, giving a name is regarded as one of the most important acts surrounding the birth of a child. It is often the center of ceremonial celebration and connected to future hopes and aspirations on behalf of the child (Tan, 2004). First names are used to distinguish people from each other, but they also signal attachment to a number of groups such as gender, age, and nationality. The balance between independent and group identity has been identified as the most important driver of first name choice in the seminal work by Lieberson (2000).

---

[10]None of the additional data sources are readily available at the individual level. Aggregate data from the Hamburg lists are produced by searching the records at ancestry.com. The subnational distribution of emigrants from the Mormon lists were constructed by counting the number of results of searching for the names of the Danish regions. In case of the local Vejle, Ribe, and Horsens, aggregate emigration numbers were based on the assumption that most of their customers resided in the same region as the location of the ticket office. Sources are found in Bender (2007).

That a preference for uncommonness in first name choice reflects individualistic values is not just a theoretical deduction. Studies on families in the United States show that parents who choose less common names for their children motivate this by a desire for the children to be unique and different (Zweigenhaft, 1981; Emery, 2013). Bearers of uncommon names have furthermore been documented to score low on values associated with conformity (Schonberg and Murphy, 1974). Conversely, Dutch families that choose more common first names have been found to spend more time with other people of their local communities (Bloothooft and Groot, 2008).

Previous research has used aggregate trends in first name uncommonness to study cultural differences across the US and Japan (Varnum and Kitayama, 2010) and to document rising individualism over time for a number of Western countries (Lieberson and Lynn, 2003; Twenge et al., 2010). I follow the common practice of these studies and calculate district level individualism as *the share of a birth cohort that does not carry one of the ten most popular female or male names of this same cohort*. At the individual level this translates into a dichotomous variable of having an uncommon (non-top-ten) or common first name.

Using first names to construct indicators of culture is associated with a lot of advantages. Cultural indicators can be constructed at any point in time and at any aggregation. With name statistics on entire populations, perfectly representative measures of the target population can be calculated. Important to this paper is that they can be analyzed back in time, where cultural indicators of the modern type do not exist. Another useful feature is the fact that the child has no influence over its name. The properties of a first name reflect the cultural environment in which the child grows up and from which it learns. The uncommonness of a first name thereby measures inherited individualism of the person carrying the name. When growing up, a child may evolve a different set of cultural preferences as it learns and interacts with other members of society. The predetermined nature of first names provides a useful source of identification.

### 4.3.1 Practical Considerations

As mentioned above, I categorize a first name as being uncommon if it is not among the ten most popular female or male names in a specific birth cohort in the population censuses. Birth cohorts are defined by a birth decade and birth district to ensure both enough observations to calculate precise measures, that the measure reflects the culture of local social networks, and that changing naming fashions do not impact the measurement.[11] Despite the support of a broad literature on the use of this measure as an indicator of

---

[11] Had I instead looked at national naming patterns, the uncommonness measures would also capture national connectivity and attitudes towards the state. Similarly, if I had included the entire population across birth decades, changing naming fashions would impact the measurement. Using birth decades instead of single birth years ensures enough observations to calculate precise measures.

inherited individualism, a few practical challenges are worth further discussion.

First, the focus on the ten most popular male and female first names may seem arbitrary, although this is common practice. To get a sense of how well this upper tail represents the entire names distribution, Figure 2 illustrates the 100 most popular first names given to boys in each historical population census. Here, the 10 most popular names sufficiently reflect the concentration of the entire distribution without being affected by the skewedness of the few most popular names.[12] Nevertheless, I construct two additional indicators of uncommonness for robustness checks: A Gini coefficient that incorporates all first names and their popularity shares, and the average first name popularity share across individuals.

Figure 2: Distribution of the top 100 boy names in each population census
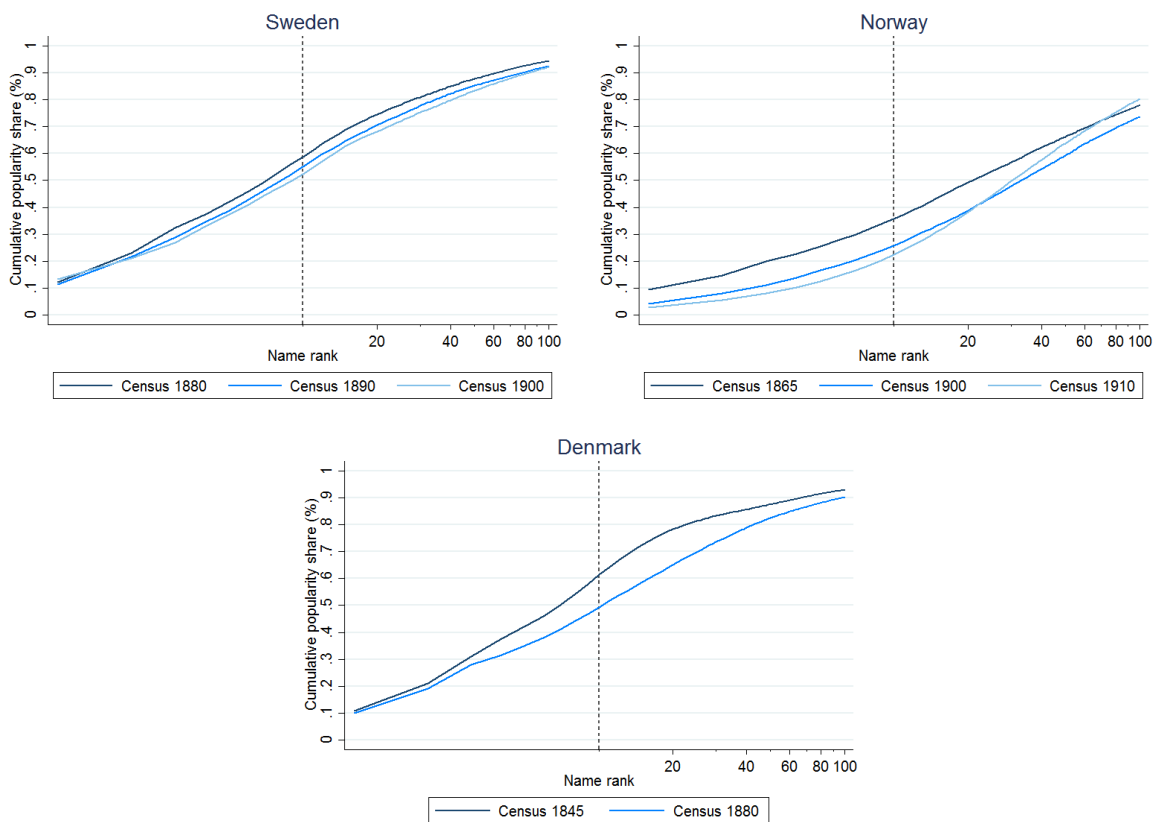


Notes: Popularity rank and cumulative popularity share of first names given to male children below the age of ten in each historical population census. The 100 first names illustrated cover an average of 85 pct. of all boys. The x-axis is presented in a logarithmic scale, and the vertical line divides the ten most popular first names from the rest.

Second, errors are present in all archival material. Names may be wrongly spelled due to illiteracy or different enumerator practices. In my baseline measure of first name uncommonness I use the raw data, because differences in spelling may also signal a preference for uniqueness. I do however check the robustness of all empirical results by using

---

[12] That the most popular first name poorly captures the overall distribution is also clear from Appendix Table A.7 that lists the most popular child first names of selected census records. Here, the most popular female name is often more widespread than the male counterpart while the opposite is true for first names below the second popularity rank.

the phonetically spelled version of first names, whereby small discrepancies in spelling (i.e. *Christian* versus *Kristian*) are eliminated. Third, some people are noted to have multiple first names (with an average of 1.5 names across all individuals in the censuses). I focus on the initial first name in the baseline measure and use the full list of first names in robustness checks.

Finally, I construct a number of variables to control for aspects in the first name uncommonness measure that do not reflect individualism. This includes gender and sibling structures and a measure of religiosity based on Biblical first names.[13] First names of Scandinavian origin are also identified to capture the intensity of nationalism and national descent.[14] I also construct indicators to account for the tradition in some families to choose a name of an older family member. Moreover, the first name of an individual may be uncommon not because of the emphasis on individuality but because of a social or physical distance to the dominant culture of a location. Although Scandinavia as a region is considered to be extraordinarily homogenous in terms of ancestry, shared history, and interconnectedness, different aspects of diversity may still be an issue in the measurement. I capture this with a measure of last name uncommonness, which since Darwin (1875) has been used to proxy genetic, ancestral, socio-economic, or geographic diversity (Cavalli-Sforza et al., 1994).[15] Where available, other aspects of social and economic circumstances are controlled for.

### 4.3.2 Validation

**Using Present-Day Indicators** To further validate that a preference for uncommon first names is linked to individualism, I compare contemporary first name patterns to better known indicators of individualism across and within countries. For this purpose, I collect all publically available first names popularity statistics for babies born in 2015 across the world.[16] Although Western countries make up the majority of the database, all continents are represented (see list in Appendix Table A.1). I calculate the baseline measure of first names uncommonness as the share of newborns with names not among

---

[13]Research has documented how firstborn children and boys in general receive more common names (Lieberson 2000). The underlying intuition is that since these individuals traditionally have been assigned the role of providers in the family, their ability to blend into general society has been more necessary.

[14]Lists of Scandinavian or Biblical origin first names are obtained from www.behindthename.com.

[15]Around 35% of the populations included in the data were given last names according to patronymic practices, whereby a suffix is attached to the father's name. These were more widespread in Norway, and less so in Denmark. In these cases, the commonness of last names does not carry direct historical signals. They do however still control for a family history of individualism, which may be impacted by the mentioned confounders.

[16]This is done by searching the web pages of national statistical offices, newspapers, and broader collections of first names databases like behindthename.com and nancy.cc. All searches are done in English and the national language (using Google Translate). For some countries only the most popular male or female name is identified. The result is a database with national level statistics for 51 countries and subnational district statistics for 15 of them. Of these, only 43 and 14 countries have statistics on the ten most popular country or district level names.

the ten most popular female or male names of the country (in 43 countries) or subnational district (in 13 countries). Some countries (including Sweden) report statistics on just the most popular first name. Although uncommonness calculated based on a single name is considered a less stable measure (see discussion above), I run additional regressions on a sample where these are included.

Results in Appendix Table A.2 show that country level first name uncommonness correlates positively with the Hofstede (2001, 2010) index of individualism, which has been used extensively in the empirical literature (Ozak and Galor, 2016; Varnum and Kitayama, 2010; Buggle, 2017; Gorodnichenko and Roland, 2011a, 2011b). Based on surveys among IBM employees across the world, Hofstede constructs six measures of core cultural traits, where Individualism is defined as a preference for a loosely-knit social framework in which individuals are expected to take care of themselves and their immediate families. The correlation is robust to controlling for continent fixed effects, number of births in 2015, three measures of diversity (ethnic, genetic, and linguistic), and other Hofstede cultural indicators. National first name patterns also correlate significantly with additional more or less known indicators of individualism and collectivism as seen in Appendix Table A.3.[17]

To ensure that the uncommonness of first names also correlates with culture within countries, I use individual level data from the European and World Values Surveys (EVS/WVS). I calculate the Welzel (2013) measure of *'emancipative values'* for each individual based on attitudes towards lifestyle liberty, gender equality, personal autonomy, and the voice of the people.[18] This measure is recommended by the World Values Survey Association as the best indicator of individualism in the survey. I add country fixed effects to study within-country variation, control for individual respondent characteristics, and cluster standard errors at the district level. Results in Appendix Table A.4 and A.5 show that individuals that live in districts with uncommon first names point out individualistic values and beliefs as being important to them. This is especially true in Scandinavia and when focusing on cultural values linked to personal autonomy and lifestyle liberty.

**Using Historical Language Data**   To gauge whether change in first name uncommonness over time is associated with cultural change in individualism, I carry out another

---

[17]Including all relevant controls, this holds for a number of survey-based indicators like the Schwartz (1994, 2004) measures of embeddedness, intellectual and affective autonomy, Van de Vliert's (2011) measure of in-group favoritism, and a measure of social tightness from Uz (2015). It also holds when looking at linguistic characteristics like pronoun drop and subject prominence, which have been shown to reflect collectivistic and individualistic cultures (Kashima & Kashima, 1998; Abdurazokzoda and Davis, 2014; Meyer-Schwarzenberger, 2015). The set of indicators is inspired by what is used in Buggle (2017).

[18]Specifically the measure is based on attitudes to homosexuality, abortion, and divorce (*lifestyle liberty*); opinions on women's right to jobs, political careers, and education (*gender equality*); the emphasis on independence and imagination over obedience as goals in child rearing (*personal autonomy*); and beliefs that people should have more say in government and local affairs (*voice of people*).

validation check using historical data for Sweden. The alternative indicator of individualism is now based on the language structure of 17 local newspapers that cover 12 Swedish districts[19] for the years 1780-1900. Using all digitized copies of these newspapers, I use the ratio of first person singular (I, me, mine) over plural (we, us, our) pronouns as an indicator of individualism. The use of different pronoun classes reflects a focus on the individual over the collective (see summary by Oyserman and Lee, 2008).

Using newspaper and time fixed effects, results in Appendix Table A.6 reveal how decadal changes in the relative pronoun use correlates with the uncommonness of first names given to cohorts born in the same district and decade that the newspaper was published. This is true when the measure of first name uncommonness is based on the 10 most popular male and female names, but not when based on the single most popular male or female name. This is expected and supports the use of the former as the baseline indicator of individualism.

## 4.4 Calculating the District Level Cultural Shock of Migration

If migrants self-select based in part on individualistic traits, then emigration must be associated with a change in the level of individualism in the migrant-sending population. I can calculate this shock quite precisely using the data described above.

Consider a population, $N$, in which the prevalence of individualism is given by $p = \frac{P}{N}$, where $P$ denotes the number of individualists in a population. The population consists of a share, $m$, that eventually migrates and a share, $r = 1 - m$, that remains, and each group is associated with a distinct intensity of individualism given by $p^m = \frac{P^m}{mN}$ and $p^r = \frac{P^r}{(1-m)N}$. Before migration, population individualism is a weighted average of the two, $p = mp^m + (1-m)p^r$, and after migration it is just that of the remaining population, $p^{post} = p^r$. The percentage point drop in home population individualism due to emigration equals:[20]

$$p - p^{post} = \frac{m}{1-m}\left(p^m - p\right) \ , \tag{1}$$

where the stronger the intensity and selectivity of emigration ( $m > 0$ and $p^m > p$), the larger the shock.

I approximate relative emigrant individualism ($p^m - p$) as an average over emigrants in

---

[19]The newspaper titles are Aftonbladet (Stockholm), Blekingsposten (Blekinge), Bollnäs tidning (Gävleborg), Dalpilen (Dalarna), Fahlu weckoblad (Dalarna), Faluposten (Dalarna), Göteborgs weckoblad (Göteborg och Bohus), Gotlands tidning (Gotland), Jönkopingsbladet (Jönköping), Kalmar (Kalmar), Lindesbergs allehanda (Örebro), Norra Skåne (Kristianstad), Östergötlands veckoblad (Östergötland), and Wernamo tidning (Jönköping). They are available for linguistic analysis at *https://spraakbanken.gu.se/*

[20]To see this, isolate $p^r$ in the formula for pre-migration population individualism: $p^r = -\frac{1}{(1-m)}(mp^m - p)$. The difference between pre-migration and post-migration population individualism then becomes: $p - p^{post} = p - p^r = p + \frac{1}{(1-m)}(mp^m - p) = \frac{1}{(1-m)}(mp^m - p + (1-m)p) = \frac{1}{(1-m)}(mp^m - mp)$

the passenger list data relative to the people in their residence district that share the same gender and birth decade.[21] This means that the measure is unaffected by the differing age and gender structures between emigrants and population. While this is beneficial in empirical analysis of district culture, the resulting shock numbers represent conservative measures. A simple comparison of emigrant and population individualism would yield larger numbers, since a majority of the emigrants of this time consisted of young people, who due to rising individualism over time carry a stronger individualistic heritage with them than older people.[22]

I then calculate the cultural shock of emigration using additional information on aggregate rates of emigration. The shock measures are available from 1860 onwards, since these are the decades covered by the emigrant passenger lists. The cultural shocks of emigration can be calculated for any time period and aggregation level. I produce baseline measures at the district and decade level. The accumulated shock over the entire Age of Mass Migration equals a loss of individualists of 3.9%-points in Denmark, 10.1%-points in Sweden, and 13.1%-points in Norway.

Although the measure most precisely reflects shocks to the distribution of *inherited* individualism in the home population, they also approximate shocks to the distribution of actual individualism if emigrants and non-emigrants are assumed to abandon their cultural heritage at similar rates. The rate and cultural shock of emigration correlate strongly around a ratio of 0.4 (see scatter plot and histogram in Appendix Figures A.1 and A.3).[23] The fact that the calculations yield positive numbers is already evidence of emigrants constituting a selected sample of the home population. This is to be further confirmed in the empirical analysis of the paper.

# 5  Empirical Frameworks

The aim of the empirical analysis of this paper is to test the three predictions of the voluntary settlement hypothesis derived in section 3. To do this, I will study the individual migrant decision and district level cultural dynamics in the short and long run. This

---

[21]For each first name carried by a resident of a certain district, birth decade, and gender in the population census, I calculate the probability of this having been chosen by individualistic parents as the share of residents that carry this name who were born in a district, where this name was not among the ten most popular names. I then assign these probabilistic measures of individualism to each emigrant based on the population census closest to emigration year. Emigrants that migrated before the census year, are added to the census before calculating the first name uncommonness. Averaging the probability of being an individualist across individuals produces the baseline measure of individualism used in this paper, which is the share of people carrying an uncommon (non-top-ten) first name.

[22]According to passenger list data described later, the mean age at emigration was 25 years and 82% emigrated below the age of 35.

[23]Additional scatter plots in Appendix Figure A.2 shows a similar correlation with the $\left(\frac{m}{1-m}\right)$-transformed rate of emigration and a more diffuse correlation with relative migrant individualism ($p^m - p$). The evolution of the shock over time closely follows the rate of emigration as seen in Appendix Figure A.4.

section outlines the relevant empirical strategies.

## 5.1 Comparing Individual Migrants and Stayers

I test the voluntary settlement hypothesis by comparing the childhood circumstances of Scandinavian migrants and non-migrants. If, among other relevant factors inherited from home, growing up in an individualistic household is a positive predictor of migration, then this is evidence of selective migration.

To construct a dataset that allows for such an analysis, I link a subsample of the newly constructed Database of Scandinavian Migrants to the pooled population census records. I identify males that migrated during the period 1869-1920 and were observed as children (below the age of 15) in either of the population censuses. Female migrants are left out of the sample as they changed their surnames at marriage, whereby they are difficult to locate in different census records. Using a procedure that is inspired by what is standard in the literature (Abramitzky et al., 2012, 2013; Long and Ferrie, 2013; Fouka, 2018), 42% of the migrants are linked to their childhood households based on first name, last name, birth year, and district of residence. Details on the linking procedure can be found in Appendix B. Of the few variables that exist in the migrant database, balance tests in Appendix Table A.8 show that differences between the matched sample and the general population of emigrants are negligible.

Wrong links are certainly present in the sample. *False negative* links, that fail to identify an emigrant in the census, may arise due to several reasons. First, some emigrants are missing from the passenger lists data and others are recorded with insufficient personal details. Second, while the last known district of residence of emigrants generally reflects the latest permanent residence (Karadja and Prawitz, forthcoming), emigrants may still be more likely to have moved away from their childhood districts due to their proposed individualistic mindset. *False positive* links, that wrongly identify stayers as emigrants, arise if a true emigrant cannot be identified due to the reasons outline above and, at the same time, another person exists in the census with similar personal characteristics. Finally, 11% of the emigrants share the same name, birth year, and residence district with other emigrants or with more than one child in the census, whereby true migrant-census links cannot be identified. The bias caused by these errors is discussed below.

For comparison, internal out-migrants are linked to the census in which they were youngest (and below 15) based on the same characteristics except that I use their birthplace district instead of residence district. Using the scarce information on emigrant birthplace, I construct an additional linked sample for robustness that is smaller but more precise.

**Regression Model**   Using the linked migrant-childhood household sample, I estimate a model that assesses how various sets of determinants early in life impacted subsequent migration behavior:

$$M_{idty} = \alpha_p P_{idty} + \phi_d + \phi_t + \phi_y + \alpha_X X_{idty} + \varepsilon_{idty} \qquad (2)$$

Here $M_{idty} = 1$ is the future decision to emigrate abroad or out-migrate to another district by an individual $i$, born in district $d$ and decade $t$, and observed in population census year $y$. $P_{idty}$ is a dichotomous variable of having been raised in an individualistic household, measured as carrying a first name not among the ten most popular boys names of the cohort (by birth decade and birth district). Birth district, birth decade, and population census year fixed effects ($\phi_d$, $\phi_t$, $\phi_y$) are included along with a comprehensive list of control variables ($X_{idty}$), which I detail later.

A positive $\alpha_p$ implies that individualistic self-selection into migration takes place. It is estimated with bias due to imperfect linking of migrants to their childhood households. False negative links are associated with downward bias as true migrants with supposedly uncommon names are now regarded as non-migrants. False positive links are more likely to happen for migrants with common first names and are therefore also a source of downward bias on the results. The downward bias from these two types of linking error is the result of migrants and stayers being more likely to be identified as the opposite cultural type than what is predicted from the theory of migration and individualism. More worrying are the missing migrant observations due to the existence of multiple potential links. These are more likely to occur with common first names, whereby collectivistic migrants tend to be left out of the analysis, causing an upward bias. A robustness check that involves all emigrants is tailored to deal with this bias.

Identification relies on the pre-determinedness of the explanatory and control variables. Another downward bias on the estimation of $\alpha_p$ stems from exactly this fact. People with uncommon first names may hold more collectivistic traits if they have been influenced in this direction by non-family members of society, and vice versa. If the voluntary settlement hypothesis is true, then $\alpha_p$ represents a lower bound of the degree of individualistic selectivity.

## 5.2   District Level Cultural Consequences in the Short Run

This paper uses two empirical frameworks to test the predictions that emigration causes lasting cultural change towards collectivism and cultural convergence across migrant-sending districts. To uncover the short-run cultural consequences, I focus on the evolution of individualistic values across decades before and during the Age of Mass Migration. To be able to do so, I transform each historical population census into a ten-year panel, where individualism is measured across districts and decades as the uncommonness of the

first names of the corresponding birth cohorts. Since first name properties are reflective of parent and societal preferences at the time of name-giving, the individualism of the past can be approximated using the first names of different birth cohorts, even though these cohorts are observed in the same year. Appendix Figure A.5 gives an overview of the periods that each panel covers, which is 110 years on average during the period 1730-1910. This time period reflects that the earliest cohorts observed in the censuses are born in 1730. I include people born before the Age of Mass Migration in the analysis to provide a reference point to the change in cultural values, which I expect took place *during* the period of mass emigration. I combine the panels with decennial data on aggregate emigration.

**Regression Model on Cultural Change**    To investigate cultural change (Prediction 2), I run panel regressions of the form

$$p_{d,y,t} = \beta_p p_{d,y,t-j} + \sum_{j=0}^{4} \beta_{sj} s_{d,t-j} + \beta_X X_{d,y,t-j} + \sum_{j=0}^{4} \beta_{Mj} M_{d,t-j} + \phi_d + \phi_t + \phi_a + \phi_y + \varepsilon_{d,y,t} \, ,$$

(3)

where $p_{d,y,t}$ is the individualism of district $d$ in decade $t$ that is measured in population census $y$, and $s_{d,t}$ is the emigration-induced cultural shock of the same district and decade. The shocks are calculated according to equation (1) from 1860 onwards and set equal to zero otherwise.[24] Note, that a positive value reflects a negative shock to the distribution of home population individualism from emigration. Control variables consist of lagged district characteristics ($X_{d,y,t}$) and additional emigration characteristics ($M_{d,t}$). $\phi_d$, $\phi_t$, $\phi_a$ and $\phi_y$ denote district, decade, cohort age, and population census fixed effects. *Cohort age* equals the time passed between each census year and cohort birth decade, and it is included to account for the mortality and emigration of the cohorts before the year of observation.[25] The error term $\varepsilon_{d,y,t}$ captures all other time-varying, district-specific unobservable shocks to individualism. Standard errors are clustered at the district level since this is the aggregation at which emigration is measured irrespective of population census year. I include up to four lags of the cultural shock variable, which is the maximum number available in the data. They capture potential delayed effects from a shock on individualism, which are present as long as the people directly affected by shock have and raise children.

The idea behind the regression model is to compare actual cultural change in the

---

[24]This paper is only concerned with the cultural shocks associated with the Age of Mass Migration. In 1850 and before this period international migration was limited, which justifies setting the shocks equal to zero for that period. I include all eight population censuses even though two are recorded before the Age of Mass Migration took place.

[25]I also conduct a robustness check, where I adjust past individualism measures by adding the people of each district and cohort who emigrated after birth and before the census year to the population first names distribution.

migrant-sending district with the change caused by selective migration. I separately test that the immediate effect, $\beta_{s0}$, and the cumulative effect of emigration, $\sum_{j=0}^{L} \beta_{sj}$, are different from zero. Specifically, I interpret negative values of $\beta_{s0}$ as the share of the compositional shock that is transmitted to new generations within the same decade. The negative sum of lag coefficients, $\sum_{j=1}^{L} \beta_{sj}$, represents the corresponding delayed transmission. If I had instead employed a model of cultural change like Voigtländer and Voth (2012) with interactions between emigration and individualism, this interpretation of parameters would have been lost. The persistence over time of the impact of a cultural shock of emigration is given by $\beta_p$.

As discussed in section 3, the exact size of the $\beta_{sj}$-coefficients is hard to predict. Fertility and socialization efforts may differ between individualists and collectivists, and the mechanisms of cultural transmission may change in response to a changing cultural environment. Such circumstances may either strengthen or weaken the cultural impact of emigration.

**Regression Model on Cultural Convergence**   To test if emigration generates cultural convergence (Prediction 3), I estimate a model that relates the dynamics of the cultural shocks of emigration to lagged individualism and emigration history:

$$s_{d,t} = \gamma_s s_{d,t-1} + \gamma_p p_{d,y,t-1} + \gamma_m pastm_{d,t} + \gamma_{int} p_{d,y,t-1} * pastm_{d,t} + \phi' + \gamma_X X_{d,y,t-1} + \varepsilon_{d,y,t} \ , \ (4)$$

where the cultural shock of emigration in district $d$ and decade $t$, $s_{d,t}$, is regressed on its lagged value, the prevalence of individualism of the same district but previous decade, $p_{d,y,t-1}$, and the cumulative sum of past rates of emigration, $pastm_{d,t} = \sum_{j=1}^{\infty} m_{d,t-j}$. Since I am interested in the nonlinear impact of past individualism and emigration history, I include an interaction between the two. Fixed effects ($\phi'$ referring to $\phi_d$, $\phi_t$, $\phi_a$, and $\phi_y$) and lagged district controls ($X_{d,y,t-1}$) are also included.

As a consequence of the voluntary settlement hypothesis, I expect more individualistic districts to experience larger flows of emigration and thus cultural shocks, which would be supported by a positive $\gamma_p$. At the same time, I expect individualism to matter less when emigration accumulates as emigration networks abroad are formed and an experience with migration spreads at home. I therefore expect a negative $\gamma_{int}$. These mean-reverting properties of the series of cultural shocks are necessary for convergence to take place.

## 5.3   District Level Cultural Consequences in the Long Run

In the final empirical framework, I test if the Age of Mass Migration was a profound enough event to have left a long-run impact on contemporary cultural differences in Scan-

dinavia. For this, I construct a dataset with measures of contemporary individualism, historical individualism, and accumulated emigration. Historical variables are taken from the earliest population census available for each country, which are the years 1845 in Denmark, 1880 in Sweden, and 1865 in Norway. I estimate a model that compares the cultural change that took place from the year of the census and until today with the sum of cultural shocks of emigration that occurred in between these years. This analysis also serves as an out-of-sample test of the short-run findings.

The regression model on long-run cultural change is given by:

$$pcont_d = \rho_1 phist_{dy} + \rho_2 cs_{dy} + \phi_y + \rho_{HX} HX_{dy} + \rho_{CM} CM_{dy} + \varepsilon_d \,, \tag{5}$$

where subscripts $d$ and $y$ refer to subnational districts and census years. Contemporary individualism, $pcont_d$, is measured as the share of newborns in 2015 that were given uncommon first names.[26] Historical individualism, $phist_{dy}$, is obtained from the early historical population census and measured as the share of children below the age of ten that were given uncommon first names. Using the first names of children yields more precise measures of individualism, because measurement error caused by mortality and emigration between birth and census year is limited.[27] The main dependent variable, $cs_{dy} = \sum_{j=y}^{1920} s_{dj}$, is the cumulative sum of cultural emigration shocks that occurred after the census was recorded, $y$, and before 1920, which was the end of the Age of Mass Migration. Historical district ($HX_{dy}$) and cumulative emigration ($CM_d$) controls are added alongside historical population census year fixed effects ($\phi_y$). Note that the model is of a cross-section type since only one census is included for each country.

A negative $\rho_2$ represents the share of the cumulative cultural shocks that have persisted until today, here over a period of up to 170 years, and a positive $\rho_1$ indicates that cross-district differences in individualism have persisted over time.

# 6  P1: Evidence of Voluntary Settlement

In this section I test the voluntary settlement hypothesis (Hypothesis 1), which states that people with stronger individualistic values were more likely to emigrate from Scandinavia during the Age of Mass Migration. Initial evidence of individualistic self-selection is given in Figure 3. Here the intensity of individualism among emigrants is compared to that of their home populations. Each dot represents a cohort gouped by gender, birth decade, and district of residence. The darker the grey color, the bigger a share of the cohort population ended up emigrating. The vast majority of emigrating cohorts had a share of

---

[26]Uncommon first names refer to first names not belonging to the ten (in Norway and Denmark) or one (in Sweden due to lack of data) most popular male or female first names of the district.

[27]Contrary to the regressions on short run dynamics, I cannot control for this kind of mortality and emigration with cross-section data.

individualists among them that was higher than their home populations as evidenced by their position above the 45 degree line.

Figure 3: Prevalence of Individualism across Cohorts
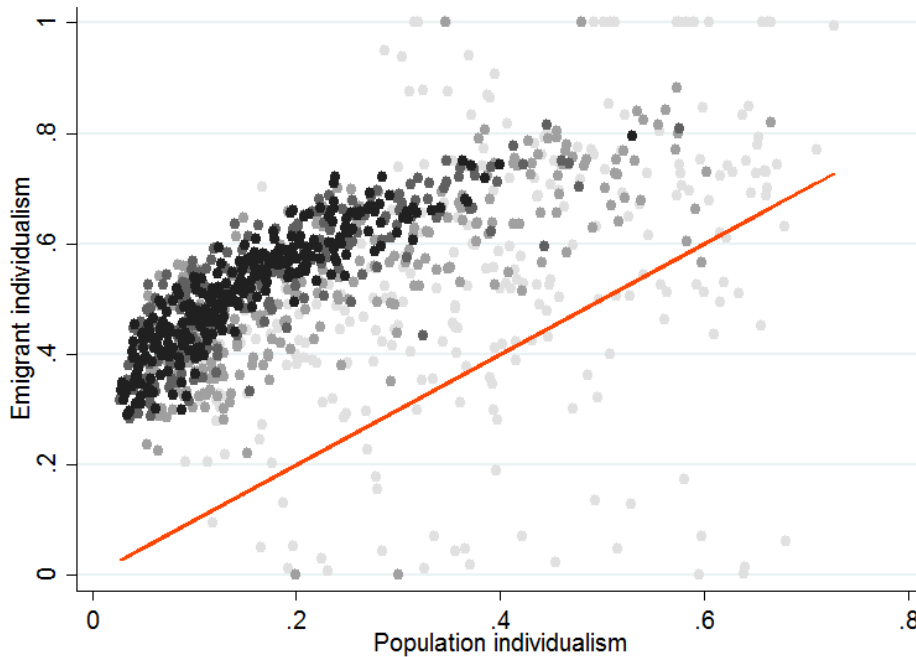


Notes: Average emigrant and initial population individualism across cohorts (by gender, birth decade, and residence district) from the pooled historical population census records. Individualism is measured as the share of the population or emigrant cohort that did not carry one of the ten most popular male or female first names of their corresponding birth cohort (see section 4.4 for further description). The darker the gray color the bigger a share of the cohort population ended up emigrating. A position above the red 45 degree line indicates that emigrating cohorts had a share of individualists among them that was higher than their home populations.

## 6.1   Individual Level Results

I proceed to investigate if the link between individualism and emigration depicted in Figure 3 was driven by self-selection or other confounding factors. While the main focus is on international emigration, which clearly dominated the migration flows of the time, internal migrants are included in the analysis as well. One would expect internal out-migrants to be more weakly associated with individualistic traits than international emigrants, because the social cost of migration would increase with moving distance. The results confirm this.

Table 1 reports the baseline OLS results on the impact of inherited individualism on the decision to migrate. In columns (1) to (4), emigrants and out-migrants are grouped together. The impact of individualism conditional only on the set of fixed effects is shown in column (1). In column (2), individual child controls are added. These include age, age squared, number of brothers, birth order among brothers, and surname uncommonness along with indicators of having migrated internally before, residing with parents, being

related to the household head (to proxy social status within the household), and being employed. I do not control for gender as the entire sample is male. Accounting for sibling structure is inspired by Abramitzky et al. (2013) and reflects the probability of future inheritance of assets such as land, which shapes one's economic prospects. As mentioned above, the uncommonness of a surname captures distinctiveness in ancestry. Additional cultural characteristics of the household are controlled for in column (3). These include indicators of whether the first name of the child is of Biblical or of Scandinavian origin, capturing religiosity, and national identity respectively. It also includes an indicator of carrying a first name that is identical to an older family member. Finally, economic characteristics of the household are added in column (4). Besides controlling for household head characteristics (age, gender, employment status, nationality, and previous migration), I also include indicators of the household being engaged in agriculture, relying on poor relief, or owning assets such as land or buildings. The total number of servants working in the household is added as a proxy of wealth and so are indicators of the household head being engaged with non-manual or skilled work (calculated from reported HISCLASS codes).

Table 1: Individualistic Selection of Scandinavian Migrants

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. Variable | | | Migration dummy | | Emigration | Out-migration |
| Inherited individualism | 0.024*** | 0.022*** | 0.019*** | 0.018*** | 0.009*** | 0.008*** |
| (first name uncommonness) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Controls: | | | | | | |
| Individual | N | Y | Y | Y | Y | Y |
| Cultural (household) | N | N | Y | Y | Y | Y |
| Economic (household) | N | N | N | Y | Y | Y |
| Observations | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 |
| R-squared | 0.04 | 0.05 | 0.05 | 0.05 | 0.02 | 0.06 |
| Mean of dep. var. | 0.0726 | 0.0726 | 0.0726 | 0.0726 | 0.0370 | 0.0360 |
| St.dev. of dep. var. | 0.259 | 0.259 | 0.259 | 0.259 | 0.189 | 0.186 |

Notes: OLS estimations. The unit of analysis is an individual male that were below the age of 15 in any of the historical population censuses. Migration dummies indicating whether the individual eventually migrated is regressed on inherited individualism (measured as carrying a first name that is not among the ten most popular). All regressions include birth district, birth cohort, and population census fixed effects. Controls are added to capture relevant child and household characteristics. Robust standard errors are shown in the parentheses with the following significance levels: *** p< 0.01, ** p< 0.05, * p< 0.1.

The coefficient on individualism remains stable and significant throughout the first four columns. Taking the coefficients at face value, a child that experiences an upbringing based on individualistic values is 2%-points more likely to migrate than a child with a collectivistic upbringing. In columns (5) and (6), the decision to migrate internationally

or internally is analyzed separately. The lower estimation coefficient on individualism in both regressions can be explained by the fact that the other migrant type is now treated as non-migrants. As expected, the act of emigration is more strongly associated with individualism than out-migration is, although the coefficients are not significantly different from each other. This is despite a relative downward bias on the emigration coefficient due to migrant-household links being less precise than for the internal migrants, where information on birthplace is used in the linking procedure. Results on all control variables are provided in the Appendix Table A.12.

**Robustness** The estimated coefficient on individualism is substantially larger when using a probit estimator in Appendix Table A.13. Here being raised in an individualistic household makes you 15-18%-points more likely to migrate. This suggests that the degree of selectivity is underestimated by the OLS regressions. The baseline results are robust to controlling for the characteristics of the father rather than the household head in a smaller sample in Appendix Table A.14, to using three alternative measures of the uncommonness of first names (based on full names, phonetically spelled names, and a continuous first name popularity share) in Appendix Table A.15, and when using the smaller sample where migrants are linked based on birthplace rather than residence in Appendix Table A.16.[28]

The biggest concern with the strategy of linking individual migrants to their childhood households is that migrants with more unusual names are easier to locate in the population census records. A robustness check is therefore carried out, where all male and female migrants that were below the age of 15 in any of the population censuses are linked based on the same characteristics as above, except the surname. Here, I do not seek true links as I am only interested in obtaining the uncommonness of the first names of migrants and non-migrants relative to their birth cohort and district of residence. Instead of using the individual level control variables directly, I average them across individuals of a cohort that share the same first name. When birth decade, district of residence, and population census year fixed effects are included, it is the variation in first name uncommonness within this cohort that is analyzed. Using this method, all migrants that were below the age of 15 in a census are accounted for, including female migrants. The results in Appendix Table A.17 confirm that in this broader sample of migrants, inherited individualism is a positive determinant in the decision to migrate. The estimates are slightly larger than the baseline results, indicating that the upward bias from identifying migrants with more uncommon names is smaller than the downward bias of not being able to identify all emigrants in the baseline analysis (see discussion in section 5.1).

---

[28] Here, the coefficient on individualism is larger for out-migration, which can be explained by the fact that a lot fewer emigrants are accounted for than in the baseline estimation due to missing birthplace information, which causes a downward bias.

**Corroborating Evidence on Migrant Behavior in the US**    The act of migration can be characterized in a number of directions. A natural extension of the hypothesis of individualistic self-selection concerns the behavior of migrants after arriving at their destination. Here one would expect that less individualistic migrants would tend to stick to each other. Analyzing Scandinavian migrants in the United States 1900 and 1910 population censuses in Appendix Table A.18 reveals that migrants of a more individualistic mindset were more likely to speak English, less likely to (after arrival to the US) marry someone of own nationality, less likely to settle in states with higher concentrations of people of own nationality, and less likely to choose Scandinavian sounding first names for children born in the US.[29,30] These results corroborate my findings on selective migration from Scandinavia as they confirm expected emigrant behavior once reaching destination.

# 7    P2: Emigration and District Level Cultural Change

## 7.1    Dynamic Panel Results

In the above section I confirmed the presence of individualistic self-selection among Scandinavian emigrants. At the district level, emigration is consequently associated with immediate shocks to the prevalence of individualism in migrant-sending population. These shocks only imply subsequent cultural change if individuals' attitudes are influenced by the attitudes of their parents or other role models. Here I test if this prediction (Prediction 2) may explain observed patterns of decennial cultural change across districts.

Table 2 presents the results of estimating model (3) on the pooled set of decennial panels that cover the period 1730-1910. Individualism is measured across districts and decades as the uncommonness of the first names of the corresponding birth cohorts. All regressions include district, decade, cohort age, and census year fixed effects. I include the average age, size, gender ratio, and average number of first names of the lagged cohorts as baseline controls. The panels contain no unit roots (p-value of Phillips-Perron test is 0.000).

---

[29]Swedish, Norwegian, and Danish born migrants are identified in the 1900 5% and 1910 1% NAPP samples for the United States, which also provide information on the year of immigration. The individualism indicator of these migrants is given by the uncommonness of their first name as it appears for the same birth cohort in the Swedish or Norwegian 1900 or the Danish 1880 population census. As I do not have information on birth place other than country of the immigrants, their first name uncommonness represent national measures. The set of control variables used in the regressions is similar but different from the rest of the empirical analysis due to differences in underlying data availability.

[30]The measure of how Scandinavian a first name sounds is calculated as the share of people in the US population of this nationality that carry this name (Fryer & Levitt, 2004). The Swedish, Norwegian, or Danish distinctiveness of a first name is not directly comparable to the first name uncommonness measure, as a name that is common in Scandinavia may not be especially distinctive in the US. All regressions control for years spent in the US, which as expected is associated with stronger assimilation outcomes. Appendix Table A.19 furthermore shows that the results are robust to adding state and county fixed effects.

The first regression includes just the lagged level of individualism, and the second adds the cultural shock of emigration that took place during the present decade. According to the estimates in column (2), a shock predicts a significant drop in the individualism that is passed on to new generations. The coefficient on lagged individualism is positive and below one, which indicates that cross-district cultural differences significantly persist but also converge over time. The fact that the persistence coefficient increases slightly in the second regression suggests that past individualism is positively correlated with the emigration shock. In light of individualistic self-selection this is an expected result and supports the idea of emigration contributing positively to the cross-district convergence of culture.

In column (3), I control for effects on individualism of emigration in general by adding the pure rate of emigration, the $\left(\frac{m}{1-m}\right)$-transformed rate of emigration, and relative emigrant individualism as control variables. From here, I add lags of the emigration shock variable (and emigration controls) to the regression in columns (4)-(7). The maximum number of lags available in the data is four. The significant lag coefficients indicate a beyond-decade cultural impact of emigration, supporting the expectation that the results of some mechanisms take time to realize.

To get a sense of the overall impact of an emigration-related cultural shock, the last rows of Table 2 report two related measures. The first is the cumulative impact, which equals the sum of all shock coefficients and reflects the total share of a shock that is transmitted to new generations at some point in time from occurrence and 40 years ahead. In the final specification this is measured as 330% of the size of the shock. In the second measure, I consider the impact that remains 50 years after the shock has occurred.[31] Over this time horizon all estimated effects are realized and have started to weaken due to culture not being perfectly persistent. In the final specification this implied impact equals 150% of the size of the initial shock. Appendix Figure A.6 maps the implied impact of a shock at several points in time after it has occurred. After 40 years, when all effects are realized, the impact slowly diminishes, and after 100 years an imprint of 35.5% of the shock remains. How this actually applies to the persistence of the cultural shocks of emigration over the 20th century will be explored in Section 7.2.

The large impact of initial shocks suggested by the estimates could be due to emigration changing the general mechanisms of cultural transmission or to differences in transmission between collectivists and individualists. It could also be due to the fact that the shock variable reflect conservative measures (see discussion in section 4.4). Of course, the exact size of the reported coefficients should be treated with some caution. What is important for this paper is that emigration appears to have been associated with a change in culture towards collectivism as a direct consequence of selective migration. The

---

[31] Using the notation from model (3) this is calculated as $\beta_{s0} * \beta_p^5 + \beta_{s1} * \beta_p^4 + \beta_{s2} * \beta_p^3 + \beta_{s3} * \beta_p^2 + \beta_{s4} * \beta_p$.

## Table 2: Emigration and Cultural Change 1730-1910

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Dep. variable | | | Individualism (first name uncommonness) | | | | |
| Lagged individualism | 0.235*** | 0.237*** | 0.234*** | 0.232*** | 0.428*** | 0.557*** | 0.750*** |
| | (0.037) | (0.037) | (0.037) | (0.038) | (0.053) | (0.101) | (0.039) |
| Cultural shock of emigration | | -0.354* | -0.785** | -0.560* | -0.736** | -0.961*** | -0.897*** |
| | | (0.208) | (0.333) | (0.287) | (0.322) | (0.334) | (0.317) |
| - first lag | | | | -1.152* | -0.911* | -0.877* | -0.626 |
| | | | | (0.620) | (0.509) | (0.496) | (0.420) |
| - second lag | | | | | -1.121** | -1.150*** | -0.690* |
| | | | | | (0.477) | (0.422) | (0.379) |
| - third lag | | | | | | -0.421 | -0.094 |
| | | | | | | (0.260) | (0.267) |
| - fourth lag | | | | | | | -0.985** |
| | | | | | | | (0.487) |
| | | | | | | | |
| Emigration controls | N | N | Y | Y | Y | Y | Y |
| Observations | 1,381 | 1,381 | 1,381 | 1,381 | 1,240 | 1,100 | 960 |
| Districts | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Census years | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| R-squared | 0.86 | 0.86 | 0.86 | 0.86 | 0.88 | 0.92 | 0.94 |
| | | | | | | | |
| Mean of dep. var. | 0.447 | 0.447 | 0.447 | 0.447 | 0.463 | 0.472 | 0.480 |
| St.dev. of dep. var. | 0.155 | 0.155 | 0.155 | 0.155 | 0.146 | 0.146 | 0.146 |
| Sum of shock lags | | -0.354 | -0.785 | -1.712 | -2.768 | -3.409 | -3.292 |
| *(p-value)* | | *0.095* | *0.022* | *0.041* | *0.020* | *0.006* | *0.005* |
| Implied impact after 50 years | | 0.000 | -0.001 | -0.004 | -0.129 | -0.465 | -1.495 |
| *(p-value)* | | *0.271* | *0.246* | *0.216* | *0.083* | *0.086* | *0.012* |

Notes: Fixed effects estimates from regressing individualism on its lagged value and the cultural shock directly associated with emigration. Data consists of the pooled historical population censuses and cover the years 1730-1910. Individualism is measured as the first name uncommonness in each population census of a cohort born in a specific decade and district. Section 3.2.2 provides an explanation of the cultural shock variable. All regressions account for district, decade, lag, and census year fixed effects as well as baseline controls (cohort number, gender, age, and number of first names). Emigration controls are the pure and $(m/(1-))$-transformed rate of emigration, relative emigrant individualism, and their lags. P-values on the sum of shock lags and the 50-year implied impact are calculated according to the delta-method. The panels contain no unit roots (p-value of Phillips-Perron test is 0.000). Robust standard errors, clustered at the district level, are shown in the parentheses with the following significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

effects on the culture passed along to new generations were realized over more decades, and they seem to have persisted over a fair amount of time. The initial cultural impact was likely strengthened by changes or cultural differences in the mechanisms underlying cultural transmission following the discussion in section 5.2.

**Robustness** Results from a number of robustness checks on the preferred specification in column (7), Table 2, are shown in Appendix Table A.20. First, I use the Arellano and Bond (1991) GMM estimator to deal with potential Nickell bias, stemming from the relatively small number of time periods compared to districts.[32] Second, I restrict the

---

[32]I use the second, third, and fourth lag of individualism along with four collapsed lags of the cultural shocks and emigration controls as internal instruments. The employed specification pass both over

sample to include only balanced 100-year panels or decades from 1860 onwards (which is the period for which measures on cultural shocks of emigration exist). Third, I control for characteristics that reflect ancestral diversity (surname uncommonness), religiosity (Biblical first name origin), or national identity (first names of Scandinavian origin), both of the district population and the emigrants.[33],[34] Fourth, I adjust past individualism measures by adding the people of each district and cohort who emigrated after birth and before the census year to the population first names distribution. Finally, I employ alternative Danish emigration numbers in the shock calculations and as control variables, because of the known errors in the original Danish passenger list data (see section 4.2 for discussion). All robustness checks yield comparable and significant results.

In Appendix Table A.21 I use alternative measures of the uncommonness of first names. These are a Gini coefficient of all first names, one minus the average first name popularity share, a non-top-10 first name share based on phonetically spelled names, and one based on full first names (see a motivation for these variables in section 4.3). Overall, the results are robust to using these measures – also when calculating relative migrant individualism.[35]

### 7.1.1  A Counterfactual Path of Scandinavian Individualism

The above analysis considered the dynamic impact of a single cultural shock of emigration. In this section I provide illustrative examples to understand how much repeated cultural shocks of emigration changed the cultural landscape in the migrant-sending districts. Based on the actual levels of individualism observed in the dynamic panels, I construct two counterfactual cultural time paths of individualism that subtracts the impact of all cultural shocks of emigration associated with the Age of Mass Migration.

Figure 4 illustrates the cross-district average of observed and counterfactual individualism based on the most recent population census of each country.[36] In decades where actual individualism is not observed, I set it equal to the level of the latest decade. The first counterfactual measure of individualism is calculated by adding the cumulative sum of all time and district specific shocks to observed district level individualism. Here I imagine that culture is perfectly persistent and cultural shocks fully transmitted to new generations. The second counterfactual measure is based on the estimation results of the

---

identification and second order correlation tests.

[33] For emigrants, calculated as the shocks to the distribution of surname uncommonness, Biblical origin and Scandinavian origin first names.

[34] Note that the majority of individual level control variables used in section 6.1 are not included here, because calculating them at the cohort level does not reflect circumstances at the time of birth of these cohorts.

[35] Only the measure based on the full list of recorded first names is associated with insignificant results. This should not be considered a source of worry since the practice of recording multiple first names differed across space, causing substantial measurement error. Using full first names in 6.1 worked well as a robustness check as district level fixed effects were controlled for.

[36] Similar figures are obtained with the other available population censuses (see Appendix Figure A.7).

Figure 4: Observed and Predicted Country Level Average Individualism 1820-2010
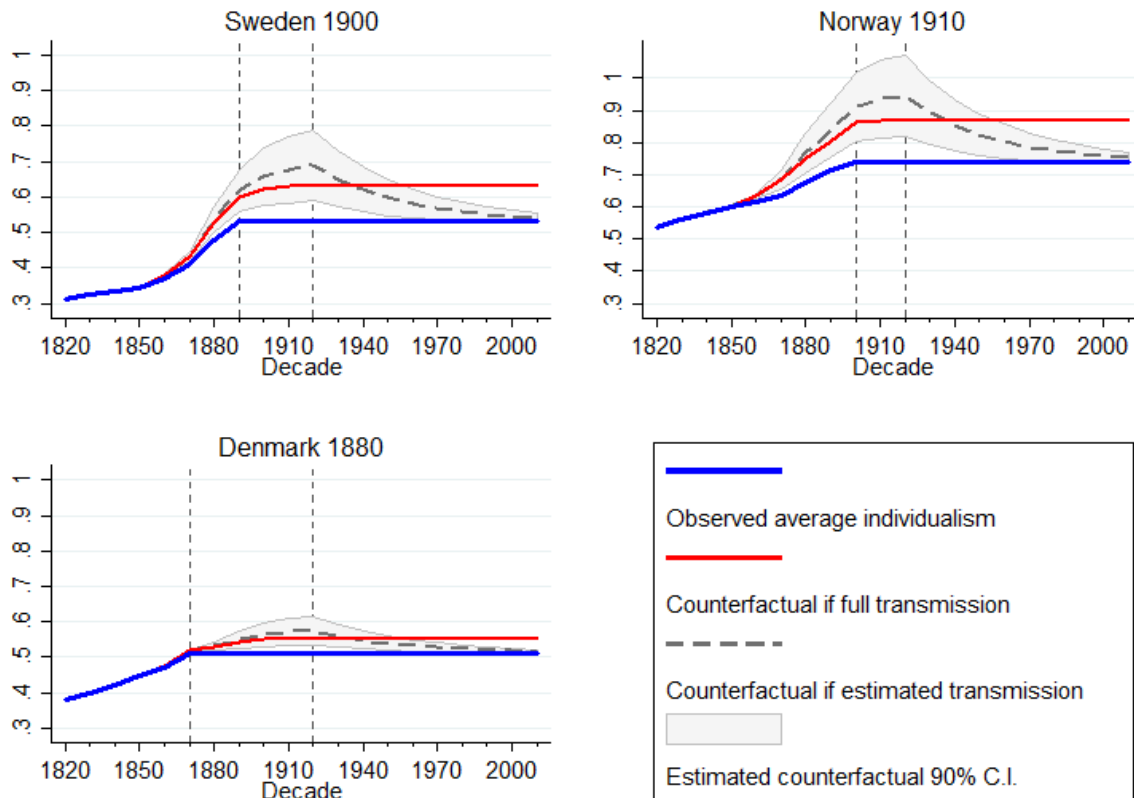


Notes: The evolution of individualism had the Age of Mass Migration not taken place. All measures are first calculated at the district level and then averaged across districts, weighted by the size of their year-1900 populations. Observed individualism equals the first name uncommonness of cohorts born in different decades and districts that are observed in the most recent population census of each country. In decades not covered by the census, individualism is set equal to level of the latest decade observed. The first counterfactual series of individualism (red line) equals the sum of observed individualism and the accumulated cultural shocks of emigration. The other measure of counterfactual individualism (gray line) equals the sum of observed individualism and the estimated cultural shock impact from the dynamic analysis in section 7.1. The associated 90 pct. confidence interval is calculated according to the delta method. The first and second vertical lines represent the latest decade observed in the population census and the latest decade of the Age of Mass Migration, 1920.

baseline dynamic regression in Table 2 column (7). Using the estimated persistence and district-specific shock (including its lags) coefficients, I predict the total impact of the cumulative sum of shocks. I add this to the observed series of individualism and calculate the associated 90% confidence intervals. These counterfactual measures represent the average level of individualism that are likely to have been reality, had the Age of Mass Migration not taken place.

Several observations are worth mentioning in light of the revealed actual and counterfactual patterns of individualism. First, one observes a general trend of rising individualism over the period, which is consistent with accounts for other countries (Macfarlane 1991, Lieberson and Lynn 2003, Twenge et al. 2010). Second, the level of individualism would have been considerably higher by the end of the Age of Mass Migration in 1920, had emigration not taken place. Taking the numbers at face value, individualism would

have been between 19.0% and 20.3% higher on average in Sweden, 17.8% and 27.9% in Norway, and 7.6% and 12.5% in Denmark, depending on the measure considered. Again, the estimated impact was more intense than the proportional impact expected if culture was transmitted perfectly from one generation to the next.

Discounting the estimated impact each decade from 1920 to 2010 with the cultural persistence coefficient yields an estimated (but statistically insignificant) long-run impact of around 12% of the cumulative sum of shocks. The next empirical analysis tests this observation using actual data on long-run cultural evolution. This is relevant since the dynamic panel estimates, which underlie the results of this subsection, may not adequately represent circumstances of time periods, including the entire $20^{th}$ century, not covered by the panel data.

## 7.2   Long-Run Results

Now I turn to look at whether the Age of Mass Migration was an event profound enough to have left a long-run impact on contemporary cultural differences in Scandinavia. Table 3 reports the results from estimating model (5) on long-run cross-section data based on the earliest census for each country. All regressions include historical census fixed effects, and baseline district controls (population size, child cohort size, and the average number of child first names). Standard errors are clustered at the district level.

The unconditional impact of historical individualism is assessed in column (1), and the cumulative cultural shock of emigration is added in column (2). Column (3) controls for the main elements of the shock calculations (the rate of emigration, the $\frac{m}{1-m}$-transformed rate of emigration, and relative emigrant individualism) to ensure that it is in fact via its initial compositional impact that emigration matters for long run cultural change. Geographical characteristics comprising district area, coastal distance, latitude, and longitude are controlled for in column (3). From here additional cultural characteristics of the historical population (ancestry diversity, religiosity, and nationalism) are controlled for in column (5). Another group of controls that capture demographic and economic circumstances of the district is added in column (6).[37] Finally, column (7) shows the results from adding the compositional shocks of emigration to population ancestry diversity, religiosity, and nationalism. The control variables are the same as in the dynamic analysis above, except for the addition of geographic, economic, and demographic characteristics.

First, the evidence weakly confirms that cross-district differences in individualism persisted over a period of between 130 and 170 years. Second, and more interestingly,

---

[37]These are the share of the population that live on a farm, rely on poor relief, live in a household that own assets such as land or buildings, are employed with non-manual work, or are employed with skilled non-manual or manual work (the latter two based on HISCLASS codes). On top thereof is the average number of servants employed in the household, and average fertility. Taking account of these economic and demographic characteristics results in a list of control variables that more precisely mirrors that of the individual selection analysis.

Table 3: Emigration and Long-Run Cultural Change (19th century - 2015)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Dep. var. | | | Contemporary individualism (first name uncommonness) | | | | |
| Historical individualism | 0.041* | 0.044** | 0.038* | 0.027* | 0.036* | 0.033* | 0.025 |
| | (0.021) | (0.021) | (0.021) | (0.015) | (0.019) | (0.019) | (0.023) |
| Cum. cultural shock of emigration | | -0.018* | -0.395*** | -0.379*** | -0.434*** | -0.490*** | -0.444** |
| | | (0.010) | (0.098) | (0.132) | (0.158) | (0.126) | (0.173) |
| Controls: | | | | | | | |
| District - baseline | Y | Y | Y | Y | Y | Y | Y |
| Emigration - numbers and individualism | N | N | Y | Y | Y | Y | Y |
| District - geography | N | N | N | Y | Y | Y | Y |
| District - add. cultural traits | N | N | N | N | Y | Y | Y |
| Emigration - add. cultural traits | N | N | N | N | N | Y | Y |
| District - economic and demographic | N | N | N | N | N | N | Y |
| Observations | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Districts | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Census years | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| R-squared | 0.98 | 0.98 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 |

OLS regressions on the long-run persistence of individualistic values and the influence from the cumulative emigration that took place during the Age of Mass Migration. The unit of observation is a district observed in a given historical population census. Contemporary individualism is measured as the share of newborns born in a district in year 2015 that received an uncommon (non-top-10 for Norway and Denmark, non-top-1 for Sweden) first name. Historical Individualism is calculated as the share of children carrying an uncommon (non-top-10) first name in the given census year. The cumulative emigration shock is the sum of all decennial cultural shocks between the census year and the end of the Age of Mass Migration in 1920. All regressions include country and historical census year fixed effects. Historical controls are added sequentially and described in the main text. Mean and S.D. of the dependent variable is 0.917 and 0.059. Robust standard errors clustered at the district level are shown in the parentheses with the following significance levels: *** p< 0.01, ** p< 0.05, * p< 0.1.

the compositional shocks accumulated over the Age of Mass Migration pushed long-run cultural development in a collectivistic direction, predicting levels of individualism that are lower by an amount corresponding to around 40% of the sum of cultural shocks. This measure is larger and statistically more significant than the implied effects of the dynamic panel analysis, indicating considerable comparative effects.[38] The cultural shock coefficient becomes particularly large and significant when controlling for the actual rate of emigration and relative emigrant individualism in column (3). This could suggest that emigration changed the cultural environment beyond the direct compositional impact to favor more individualism. Such effects could be due to indirect influences of the United States via communications with migrants or return migrants.

**Robustness** To check the robustness of the long-run results, I first repeat the regressions on the pooled set of population censuses. While the above considered only the earliest census for each country, a subnational district is now represented with multiple historical values of individualism – one for each population census. The results in Appendix Ta-

---

[38]According to Appendix Table A.6, an impact of 8% of the shock would remain after 150 years, which is the average time passed since occurrence of the shocks considered in this analysis.

ble A.22 document a statistically significant persistence of collectivistic traits, which is expected as the average time period between census year and today is now shorter. The impact of the cumulative shock is estimated to be less than half the size as in the results above, but it is still significant and sizable.

In Appendix Table A.23 I repeat the regressions based on the earliest census for each country, but this time using indicators of contemporary individualism from the World and European Values Surveys. These are the same indicators that I used in the validation tests in section 4.3. Results show that the historical level of individualism as well as the cumulative compositional shock of emigration impacted contemporary cross-differences of overall 'emancipative values' and in particular the emphasis on autonomy as expected. The fact that the estimated persistence parameters are now significant suggests that the measurement error contemporary individualism based on the uncommonness of the first names (which stems from this being based on the single most popular first names in Sweden) causes noise in the estimation.

Finally, the results also hold when using alternative sources of Danish emigration data as is clear in Appendix Table A.24.

# 8 P3: Emigration and Cultural Convergence across Districts

Up to now, I have documented that selective emigration caused a reduction in the prevalence of individualists in migrant-sending populations, which persisted over time as new generations adopted the traits of the altered cultural landscape. To conclude the empirical analysis of this paper, I investigate the validity of the final prediction outlined in section 3 that selective emigration caused asymmetric cultural change, which led to cross-district cultural convergence (Prediction 3).

To begin with, I illustrate how the predicted total impact of emigration affected the variation in individualism across districts. Figure 5 shows the country level coefficient of variation of observed individualism and the two counterfactual measures of individualism that I described in section 7.1.1. Cross-district convergence appears to have taken place, since the variation in observed individualism is falling over time.[39] Variation in counterfactual individualism is higher, indicating that cultural change caused by selective emigration impacted subnational districts asymmetrically and contributed significantly to the observed convergence. Again the estimated impact weakens over time due to imperfect cultural persistence, but a comparative impact is likely still present today following the results on long-run cultural change above.[40]

---

[39] Cross-district variation in individualism is even lower today. The coefficient of variation of contemporary first name uncommonness is 0.014 in Norway, 0.004 in Sweden, and 0.187 in Denmark.

[40] Again, I obtain similar numbers if I use the other available population censuses in Appendix Figure

Figure 5: Observed and Predicted Country Level CV of Individualism 1820-2010
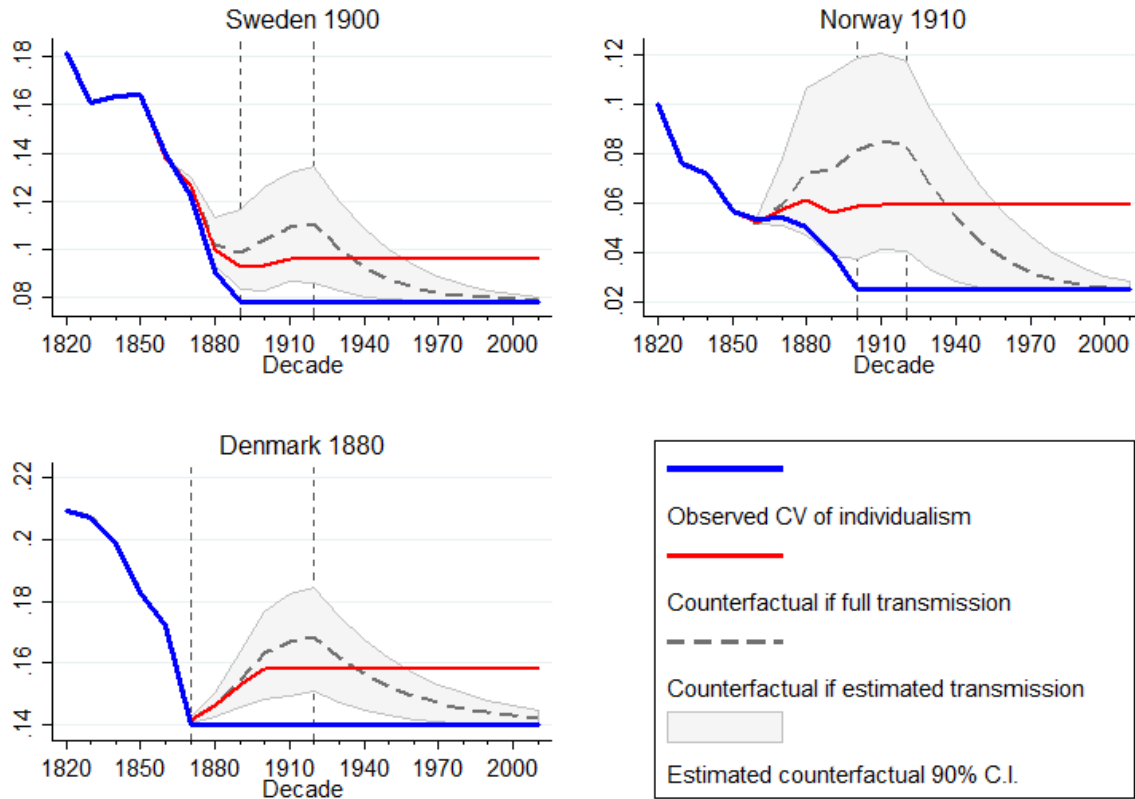
Notes: The coefficient of variation of individualism had the Age of Mass Migration not taken place. All measures are first calculated at the district level and then coefficients of variation are calculated across districts, weighted by the size of their year-1900 populations. Observed individualism equals the first name uncommonness of cohorts born in different decades and districts that are observed in the most recent population census of each country. In decades not covered by the census, individualism is set equal to level of the latest decade observed. The first counterfactual series of individualism (red line) equals the sum of observed individualism and the accumulated cultural shocks of emigration. The other measure of counterfactual individualism (gray line) equals the sum of observed individualism and the estimated cultural shock impact from the dynamic analysis in section 7.1. The associated 90 pct. confidence interval is calculated according to the delta method. The first and second vertical lines represent the latest decade observed in the population census and the latest decade of the Age of Mass Migration, 1920.

## 8.1 Dynamic Panel Results

To shed light on the mechanism underlying the contribution of emigration to cultural convergence, I study the determinants of the cultural shocks caused by selective migration. The results of estimating model (4) on the dynamic panels of culture and migration during the period 1730-1910 are shown in Table 4. In column (1), I include only the lagged level of cultural shocks, baseline district controls (average age, cohort size, gender ratio, and average number of first names of the lagged cohorts), and fixed effects for districts, decades, cohort ages, and population census years. I control for the past level of individualism in column (2) and the cumulative sum of past emigration rates in column (3). In column (4), I interact past individualism and past emigration to identify non-linear effects, and in column (5) all other variables are interacted by past individualism.

A.8.

Table 4: Cultural Shock Dynamics, 1730-1910

|  | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dep. variable | | | Cultural shock of emigration | | |
| Lagged shock | 0.581*** | 0.581*** | 0.073 | 0.069 | -1.622*** |
|  | (0.043) | (0.043) | (0.057) | (0.054) | (0.328) |
| Lagged individualism |  | 0.004 | 0.004* | 0.006** | 0.038** |
|  |  | (0.003) | (0.002) | (0.002) | (0.017) |
| Past emigration |  |  | 0.100*** | 0.219*** | 0.433*** |
|  |  |  | (0.011) | (0.036) | (0.040) |
| x lag individualism |  |  |  | -0.186*** | -0.544*** |
|  |  |  |  | (0.060) | (0.068) |
| Observations | 1,381 | 1,381 | 1,381 | 1,381 | 1,381 |
| Districts | 50 | 50 | 50 | 50 | 50 |
| Census years | 8 | 8 | 8 | 8 | 8 |
| R-squared | 0.74 | 0.74 | 0.79 | 0.80 | 0.82 |

Notes: Fixed effects estimates on the determinants of the cultural shocks of emigration. The unit of observation is a district and decade, observed in a historical population census. All specifications use the pooled historical population censuses and add fixed effects for districts, decades, lags, and census years. Baseline controls include cohort number, gender, age, and number of first names. Past emigration is the cumulative sum of past rates of emigration. The mean and S.D. of the dependent variable is 0.005 and 0.012. Robust standard errors, clustered at the district level, are shown in the parentheses with the following significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Past individualism is associated with larger cultural shocks of emigration, which is an expected consequence of the voluntary settlement hypothesis on selective migration. The impact of individualism decreases as emigration accumulates, which is consistent with the idea of individualism playing a smaller role in the decision to emigrate as migrant networks and the public collection of migrant experiences grow.

I repeat the relevant robustness checks from the dynamic analysis on cultural change in Appendix Table A.25 using different estimators, samples, control variables, and measures of individualism. Here I also run regressions using the pure rate of emigration or relative emigrant individualism as dependent variables (the latter without interaction effects). Results confirm the above findings, including the prediction that the gap between emigrant and population individualism (relative migrant individualism) narrows as past emigration accumulates. Considering the determinants of the cumulative sum of shocks from each census year to 1920 in a cross-section analysis in Appendix Table A.26 provides further support for the results.[41]

Taking this evidence together, a picture of individualism and emigration evolving jointly and causing cross-district convergence emerges. On the one hand districts that are initially more individualistic experienced faster change towards collectivism due to selective migration, and on the other hand this change slowed down as emigration accumulated.

---

[41] In this test, I pool all population censuses to get variation in the variable of past emigration.

# 9    Conclusion

This paper examines the joint evolution of culture and migration in Scandinavia during and since the Age of Mass Migration (1850-1920), which involved the exodus of around 25% of the entire Scandinavian populations. I find that people of an individualistic mindset were more prone to migrate than their collectivistic neighbors. Accounting for other relevant childhood characteristics, growing up in an individualistic household was associated with an increase in a subsequent likelihood of migration. Due to self-selection on individualistic traits, mass emigration caused a direct compositional change in the home population. Over the period this amounted to a loss of individualists of approximate 3.7%-points in Denmark, 9.4%-points in Sweden, and 13.6%-points in Norway.

I proceed to document that the transmission of cultural traits was sufficiently strong to ensure a lasting impact on the evolution of collectivistic traits of the initial compositional shocks of emigration. In the short run, a changed cultural composition in the migrant-sending population was transmitted to new generations. The cultural change that took place during the Age of Mass Migration was sufficiently profound to impact cross-district cultural differences in present day Scandinavia. Contemporary levels of individualism would thus have been significantly higher had emigration not occurred. The interdependent relationship between individualism and emigration provides the basis of another important result. The variation in individualism across subnational districts fell over the period 1730-1910, and evidence suggests that the asymmetric cultural change caused by selective migration contributed significantly to this trend.

The potential societal implications of the emigration-driven cultural change are of great importance. The period of the Age of Mass Migration was characterized by industrialization, urbanization, and democratization in Scandinavia. Individualism was generally on the rise, in part due to these developments, but it seems conceivable that the collectivistic turn caused by emigration played a role in subsequent institutional developments. While economic freedom is high in contemporary Scandinavia, the region is known for its priority of social cohesion and collective insurance. This is particularly clear when contrasting the Scandinavian welfare model with American liberal capitalism.[42] Future research may identify the impact of culture on these developments.

This study may form the basis for future research in several directions. With the construction of valid historical cultural indicators at the individual and society level, cultural change and persistence can be investigated over long periods of time and underlying mechanisms identified in detail. For example, Knudsen (2018) studies the relevance of different

---

[42]Today the Scandinavian countries are among the very top on the World Bank "Ease of Doing Business" ranking. The countries are characterized by low levels of corruption and the protection of human rights. At the same time, tax rates are among the highest in the world, and they spend on providing a large number of services. Labor unions are particularly strong and collective bargaining of wages is the norm. A famous characterization of Scandinavian culture is the "Law of Jante" which teaches people not think that they are anything special. See discussions in Acemoglu et al. (2012) and Krugman (2015).

channels of cultural transmission of collectivism at the individual child level across more than four million historical European families. Other research may study the impact of cultural traits and values on individual behavior and aggregate outcomes. Thus, further light may be shed on how culture at the same time evolves in and shapes social and economic structures of society.

# References

[1] Abramitzky, R., Boustan, L. P., and Eriksson, K. 2012. Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration. *American Economic Review*, **102**(5): 1832-56.

[2] Abramitzky, R., Boustan, L. P., and Eriksson, K. 2013. Have the poor always been less likely to migrate? Evidence from inheritance practices during the age of mass migration. *Journal of Development Economics*, **102**(1), 2-14.

[3] Abramitzky, R. and Boustan, L. P. 2017. Immigration in American economic history. *Journal of Economic Literature*, **55**(4), 1311-45.

[4] Acemoglu, D., Robinson, J. A., and Verdier, T. 2012. Can't We All Be More Like Scandinavians? Asymmetric growth and institutions in an interdependent world (No. w18441). *National Bureau of Economic Research.*

[5] Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. 2003. Fractionalization. *Journal of Economic Growth*, **8**(2): 155– 194.

[6] Alesina, A., & Giuliano, P. 2015. Culture and institutions. *Journal of Economic Literature*, 53(4), 898-944.

[7] Andersson, D., Karadja, M., and Prawitz, E. 2017. Mass migration, cheap labor, and innovation. *Working Paper.*

[8] Arellano, M., and Bond, S. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, **58**(2), 277-297.

[9] Ashraf, Q., and Galor, O. 2013. The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development. *American Economic Review*, **103**(1): 1-46.

[10] Barry, H., Child, I. L., and Bacon, M.K. 1959. Relation of child training to subsistence economy. *American Anthropologist*, **61**(1), pp.51-63.

[11] Bazzi, S., Fiszbein, M. and Gebresilasse, M. M. 2018. Frontier Culture: The Roots and Persistence of Rugged Individualism in the United States. *Working Paper.*

[12] Becker, S. O., and Pascali, L. (forthcoming). Religion, division of labor and conflict: Anti-semitism in german regions over 600 years. *American Economic Review.*

[13] Bender, H. 2007. Kilderne på nettet til udvandringen fra Danmark. *Working paper.*

[14] Bentzen, J. S. (forthcoming). Acts of God? religiosity and natural disasters across subnational world districts. *Economic Journal.*

[15] Bhugra, D., & Becker, M. A. 2005. Migration, cultural bereavement and cultural identity. *World Psychiatry*, 4(1), 18-24.

[16] Bisin, A., and Verdier, T. 1998. On the cultural transmission of preferences for social status. *Journal of Public Economics*, **70**(1), 75-97.

[17] Bisin, A., and Verdier, T. 2000. "Beyond the melting pot": cultural transmission, marriage, and the evolution of ethnic and religious traits. T*he Quarterly Journal of Economics*, **115**(3), 955-988.

[18] Bisin, A., and Verdier, T. 2001. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, **97**(2), 298-319.

[19] Bloothooft, G. and Groot, L. 2008. Name clustering on the basis of parental preferences. *Names,* **56**(1):111–163

[20] Bozeman, B. 2007. *Public values and public interest: Counterbalancing economic individualism.* Georgetown University Press.

[21] Buggle, J. C. 2017. Growing Collectivism: Irrigation, Group Conformity and Technological Divergence. *Working Paper.*

[22] Cavalli-Sforza, L. L., and Feldman, M. W. 1981. *Cultural transmission and evolution: a quantitative approach.* Princeton University Press.

[23] Cavalli-Sforza L. L., Menozzi, P., and Piazza, A. 1994. *The history and geography of human genes*, Princeton University Press

[24] Chen, C. C., Chen, X. P., & Meindl, J. R. 1998. How can cooperation be fostered? The cultural effects of individualism-collectivism. *Academy of management review*, 23(2), 285-304.

[25] Chowdhury, S., Sutter, M., and Zimmermann, K. 2018. Evaluating intergenerational persistence of economic preferences: A large scale experiment with families in Bangladesh. *Working Paper.*

[26] Darwin, C. 1859. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* John Murray, London

[27] Davis, L.S. and Abdurazokzoda, F. 2016. Language, culture and institutions: Evidence from a new linguistic dataset. *Journal of Comparative Economics*, **44**(3): 541-561.

[28] Docquier, F., and Rapoport, H. 2012. Globalization, brain drain, and development. *Journal of Economic Literature*, **50**(3), 681-730.

[29] Dohmen, T., Falk, A., Huffman, D., and Sunde, U. 2011. The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies*, **79**(2), 645-677.

[30] Eisenbruch, Maurice. 1991. "From post-traumatic stress disorder to cultural bereavement: Diagnosis of Southeast Asian refugees." *Social Science and Medicine* (1991): 673-680.

[31] Emery, H. B., 2013. "What's In a Name?" American Parents' Search for the Perfect Baby Name. *Doctoral dissertation, University of California, Berkeley.*

[32] van Everdingen, Y. M., & Waarts, E. 2003. The effect of national culture on the adoption of innovations. *Marketing Letters*, 14(3), 217-232

[33] Fouka, V. 2018. Backlash: The unintended effects of language prohibition in US schools after World War I. *Working Paper.*

[34] Fryer, R.G. and Levitt, S.D., 2004. The causes and consequences of distinctively black names. *The Quarterly Journal of Economics*, **119**(3), pp.767-805.

[35] Galor, O., and Ömer Ö. (2016). The Agricultural Origins of Time Preference. *American Economic Review*, **106**(10): 3064-3103.

[36] Galor, O., Özak, Ö., and Sarid, A. 2017. Geographical Origins and Economic Consequences of Language Structures. *Working Paper.*

[37] Galor, O., Özak, Ö., and Sarid, A. 2018. Geographical Origins of Language Structures. *Working Paper.*

[38] Giuliano, P. and N. Nunn, Understanding Cultural Persistence and Change. *Working Paper.*

[39] Gorodnichenko, Y., and Roland, G. 2011. Individualism, innovation, and long-run growth. *Proceedings of the National Academy of Sciences*, **108**(4): 21316-21319.

[40] Gorodnichenko, Y., and Roland, G. 2011. Which dimensions of culture matter for long-run growth?. *American Economic Review*, **101**(3), 492-98.

[41] Gorodnichenko, Y., and Roland, G. 2017. Culture, institutions, and the wealth of nations. *Review of Economics and Statistics*, **99**(3), 402-416.

[42] Greif, A. 1994. Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *Journal of Political Economy*, **102**(5), 912-950.

[43] Hatton, T. J., and Williamson, J. G. 1992. *What drove the mass migrations from Europe in the late nineteenth century?.*

[44] Heine, S. J., Ruby, M. B., 2010. Cultural psychology. *Wiley Interdisciplinary Reviews: Cognitive Science,* **1** (2), 254–266.

[45] Henrich, J., Heine, S. J., & Norenzayan, A. 2010. The weirdest people in the world?. *Behavioral and brain sciences*, 33(2-3), 61-83.

[46] Hofstede, G. (2001), *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, 2nd ed. Sage, Thousand Oaks, CA.

[47] Indseth, T. 2006. Forestillinger om Amerika i Norge. *Masteroppgave, University of Oslo*

[48] Jaeger, D. A., Dohmen, T., Falk, A., Huffman, D., Sunde, U., and Bonin, H. 2010. Direct evidence on risk attitudes and migration. *The Review of Economics and Statistics*, **92**(3), 684-689.

[49] Kashima, E. S., and Kashima, Y., 1998. Culture and language the case of cultural dimensionsand personal pronoun use. *Journal of Cross-Cultural Psychology* **29** (3), 461–486.

[50] Karadja, M., and Prawitz, E. (forthcoming). Exit, voice, and political change: Evidence from Swedish mass migration to the United States. *Journal of Political Economy.*

[51] Kitayama, S., Ishii, K., Imada, T., Takemura, K., and Ramaswamy, J. (2006). Voluntary settlement and the spirit of independence: Evidence from Japan's northern frontier. *Journal of personality and social psychology,* **91**(3), 369.

[52] Knudsen, A. S. B. 2018. The Intergenerational Transmission of Collectivistic Traits: Evidence Across Millions of Historical European Families. *Mimeo.*

[53] Krugman, P. 2015, October 19. Something Not Rotten. *The New York Times*, pp. A23

[54] Lieberson, S., and Lynn, F. B. 2003. Popularity as Taste. *Onoma,* **38**, 235-276.

[55] Lieberson, S. 2000, *A Matter of Taste: How Names, Fashions, and Culture Change,* Yale University Press, New Haven.

[56] Long, J., and Ferrie, J. (2013). Intergenerational occupational mobility in Great Britain and the United States since 1850. *American Economic Review,* **103**(4), 1109-37.

[57] Macfarlane, A. 1978. *The origins of English individualism: the family property and social transition.* Wiley-Blackwell.

[58] Mateos, P. 2013. *Names, Ethnicity and Populations.* Heidelberg, Germany: Springer.

[59] Meyer-Schwarzenberger, M. 2015. Indicators of Subject Prominence in 155 National Majority. *Working paper.*

[60] Minnesota Population Center. 2017. *North Atlantic Population Project: Complete Count Microdata.* Version 2.3 [dataset]. Minneapolis: Minnesota Population Center.

[61] Mokyr, J. 1992. Technological Inertia in Economic History. *The Journal of Economic History,* **52**(2): 325-338

[62] Morling, B. and Lamoreaux, M. 2008. Measuring culture outside the head: A meta-analysis of individualism–collectivism in cultural products. *Personality and Social Psychology Review.* **12**(3), 199-221.

[63] Morris, M. H., Avila, R. A., & Allen, J. 1993. Individualism and the modern corporation: Implications for innovation and entrepreneurship. *Journal of management,* 19(3), 595-612.

[64] Murrell, P. and Schmidt, M. 2011. The Coevolution of Culture and Institutions in Seventeenth Century England. *Working paper.*

[65] Nunn N. Culture and the Historical Process. *Economic History of Developing Regions.* 2012; 27 (S1) : 108-126.

[66] Semmingsen, I., and Haugen, E. I. 1978. *Norway to America: A History of the Migration.* U of Minnesota Press.

[67] Schonberg, W. B., and Murphy, D. M. 1974. The relationship between the uniqueness of a given name and personality. *The Journal of social psychology.* **93**(1), 147-148.

[68] Ogihara, Y., Fujita, H., Tominaga, H., Ishigaki, S., Kashimoto, T., Takahashi, A., Toyohara, K. and Uchida, Y. 2015 Are common names becoming less common? The rise in uniqueness and individualism in Japan. *Front. Psychol.* **6**:1490.

[69] Oishi, S. 2010. The psychology of residential mobility: Implications for the self, social relationships, and well-being. *Perspectives on Psychological Science*, **5**(1), 5-21.

[70] Oishi, S., Lun, J., and Sherman, G. D. 2007. Residential mobility, self-concept, and positive affect in social interactions. *Journal of Personality and Social Psychology*, **93**(1), 131.

[71] Olsson, O. and Paik, C. 2016. Long-run cultural divergence: Evidence from the Neolithic Revolution, *Journal of Development Economics*. **122**,(C), 197-213

[72] Oyserman D., and Lee, S. W. 2008. Does culture influence what and how we think? Effects of priming individualism and collectivism. *Psychological Bulletin*, **134**, 311-342.

[73] Schwartz, S. H. 2004. Mapping and interpreting cultural differences around the world. *International Studies in Sociology and Social Anthropology*, 43-73.

[74] Schwartz, S. H. 1994. *Beyond individualism/collectivism: New cultural dimensions of values.* Sage Publications, Inc.

[75] Sønnichsen, O. 2013. *Rejsen til Amerika.* Gyldendal.

[76] Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X. and Kitayama, S.

[77] (2014). Large-Scale Psychological Differences within China Explained by Rice versus Wheat Agriculture. *Science*, **344** (6184), 603–608.

[78] Tan, P. K. 2004. Evolving naming patterns: anthroponymics within a theory of the dynamics of non-Anglo Englishes. *World Englishes*, **23**(3), 367-384.

[79] Taylor, A. M., and Williamson, J. G. 1997. Convergence in the age of mass migration. *European Review of Economic History*, **1**(1), 27-63.

[80] Taylor, M. Z., & Wilson, S. 2012. Does culture still matter?: The effects of individualism on national innovation rates. *Journal of Business Venturing*, 27(2), 234-247

[81] Turner, F. J. 1920. *The Frontier in American History* (New York, 1920), The Significance of Sections in American History

[82] Twenge, J. M., Abebe, E. M., and Campbell, W. K. 2010) Fitting in or standing Out: Trends in American parents' choices for children's names, 1880–2007. *Social Psychological and Personality Science*, **1**(1), 19-25.

[83] Uz, I. 2015. The index of cultural tightness and looseness among 68 countries. *Journal of Cross-Cultural Psychology*, **46** (3), 319—335

[84] Van de Vliert, E. 2011. Climato-economic origins of variation in ingroup favoritism. *Journal of Cross-Cultural Psychology*, **42** (3), 494–515.

[85] Varnum, M. E., and Kitayama, S. 2010. What's in a name? Popular names are less common on frontiers. *Psychological Science*, **22**(2), 176-183.

[86] Voigtländer, N., and Voth, H. J. 2012. Persecution perpetuated: the medieval origins of anti-Semitic violence in Nazi Germany. *Quarterly Journal of Economics*, **127**(3), 1339-1392.

[87] Wagner III, J. A. 1995. Studies of individualism-collectivism: Effects on cooperation in groups. *Academy of Management journal*, 38(1), 152-173.

[88] Welzel, C. 2013. *Freedom Rising: Human Empowerment and the Quest for Emancipation*. New York: Cambridge University Press.

[89] Zweigenhaft, R. L. 1981. Unusual names and uniqueness. *The Journal of Social Psychology*, **114**(2), 297-298.

# A   Appendix: Additional Figures and Tables

Figure A.1: Rate and Cultural Shock of Emigration
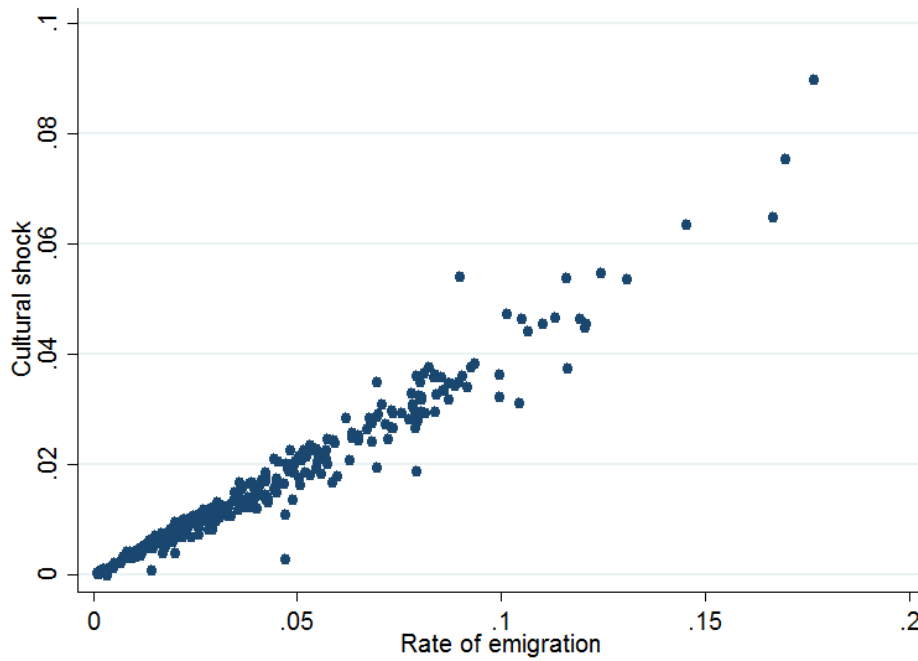


Notes: Scatter between the cultural shock and rate of emigration of each decade and district over the period 1860-1920.

Figure A.2: Rate and Cultural Shock of Emigration: Additional Measures
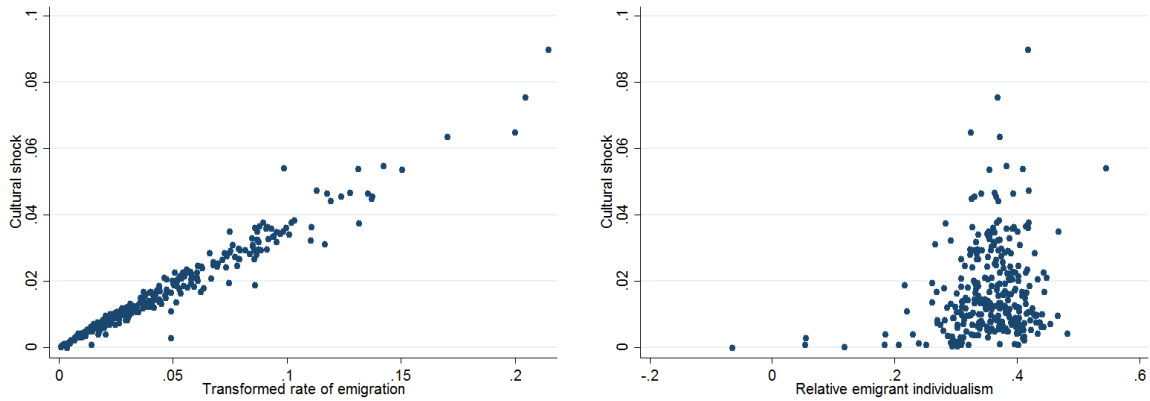


Notes: Scatter between the cultural shock of emigration and the $(m/(1-m))$-transformed rate of emigration (left) and relative emigrant individualism (right) of each decade and district over the period 1860-1920.

Figure A.3: Rate and Cultural Shock of Emigration
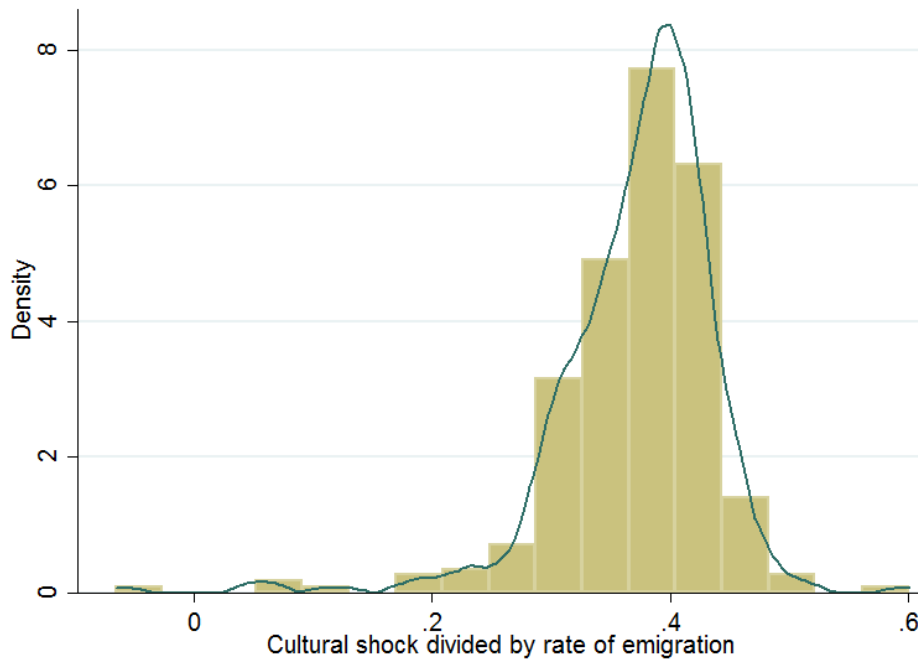


Notes: Histogram and kernel density of the cultural shock divided by the rate of emigration across districts and decades over the period 1860-1920.

Figure A.4: Decennial Rate and Cultural Shock of Emigration 1850-1920
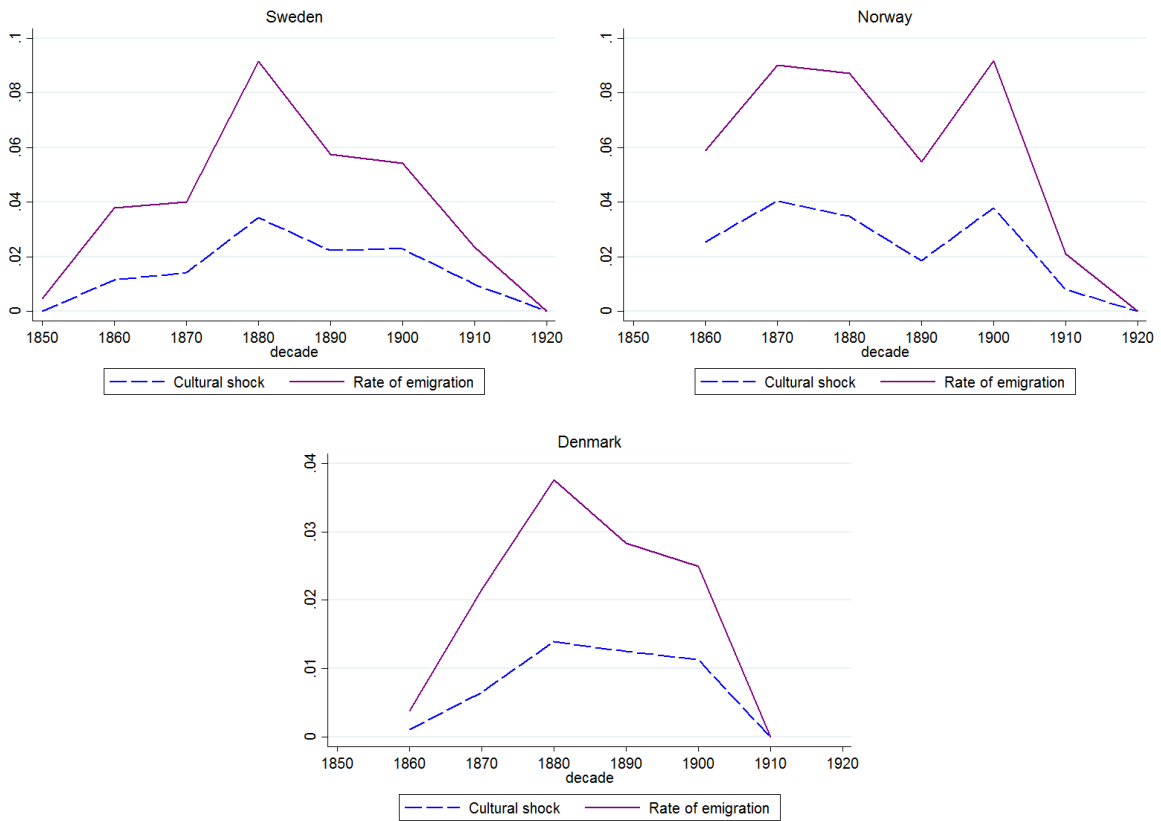


Notes: Country level cultural shocks and rates of emigration of each decade in the period 1860-1920.

Figure A.5: Panel Data Structure
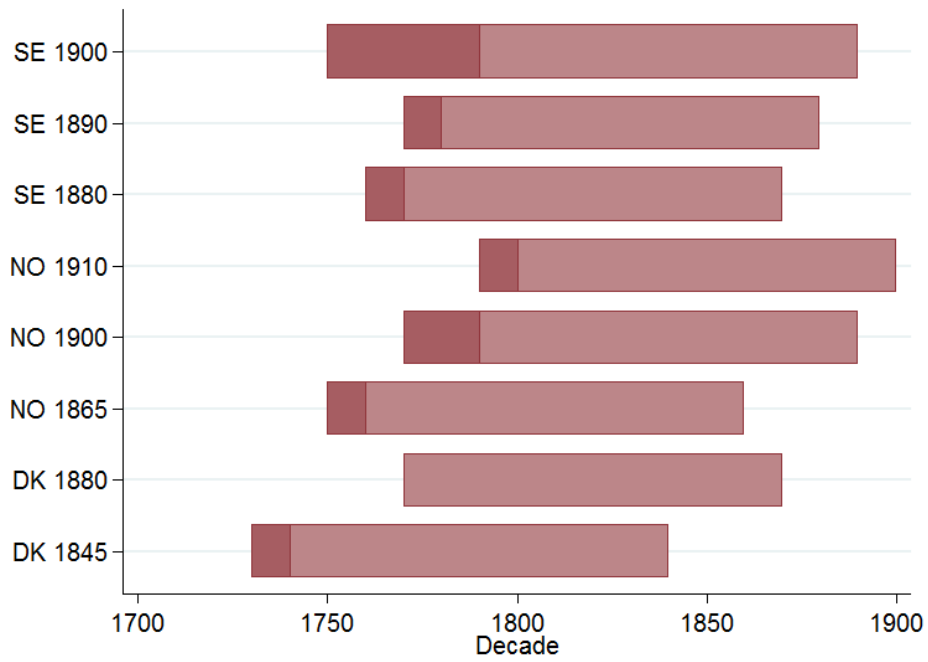


Notes: Decades covered by each historical population census once they are transmformed in to panels. The light color reflects the time coverage of the balanced panels.

Figure A.6: Cumulative Impact of a Shock
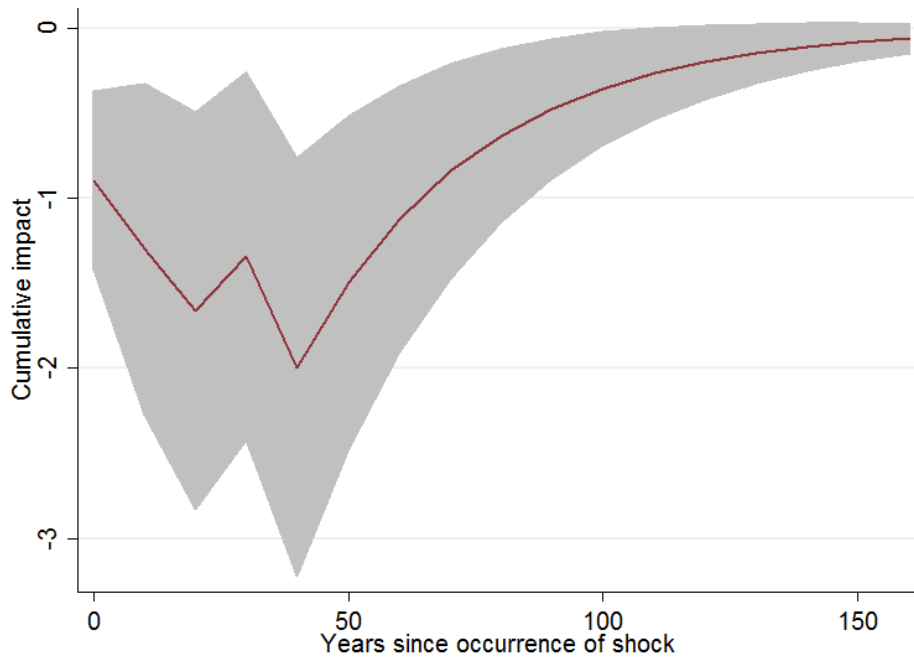


Notes: The implied cumulative impact of a shock at various times after occurrence. Calculated from the estimates in the baseline dynamic analysis. A 90 pct. confidence interval is calculated according to the delta method.

Figure A.7: Predicted Average Individualism 1820-2010
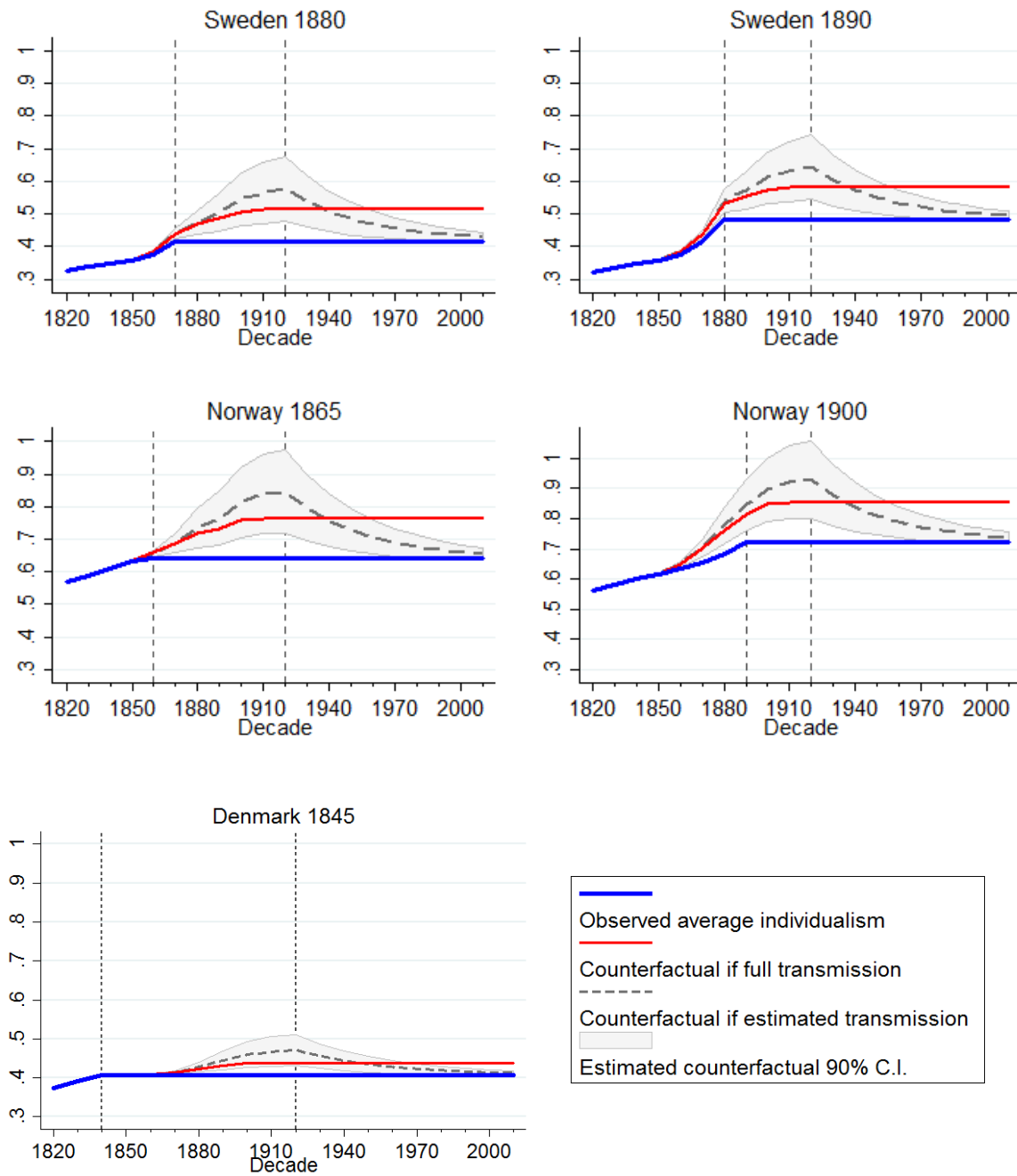
Notes: Paths of average of individualism. See section 7.1.1 for details.

Figure A.8: Predicted CV of Individualism 1820-2010
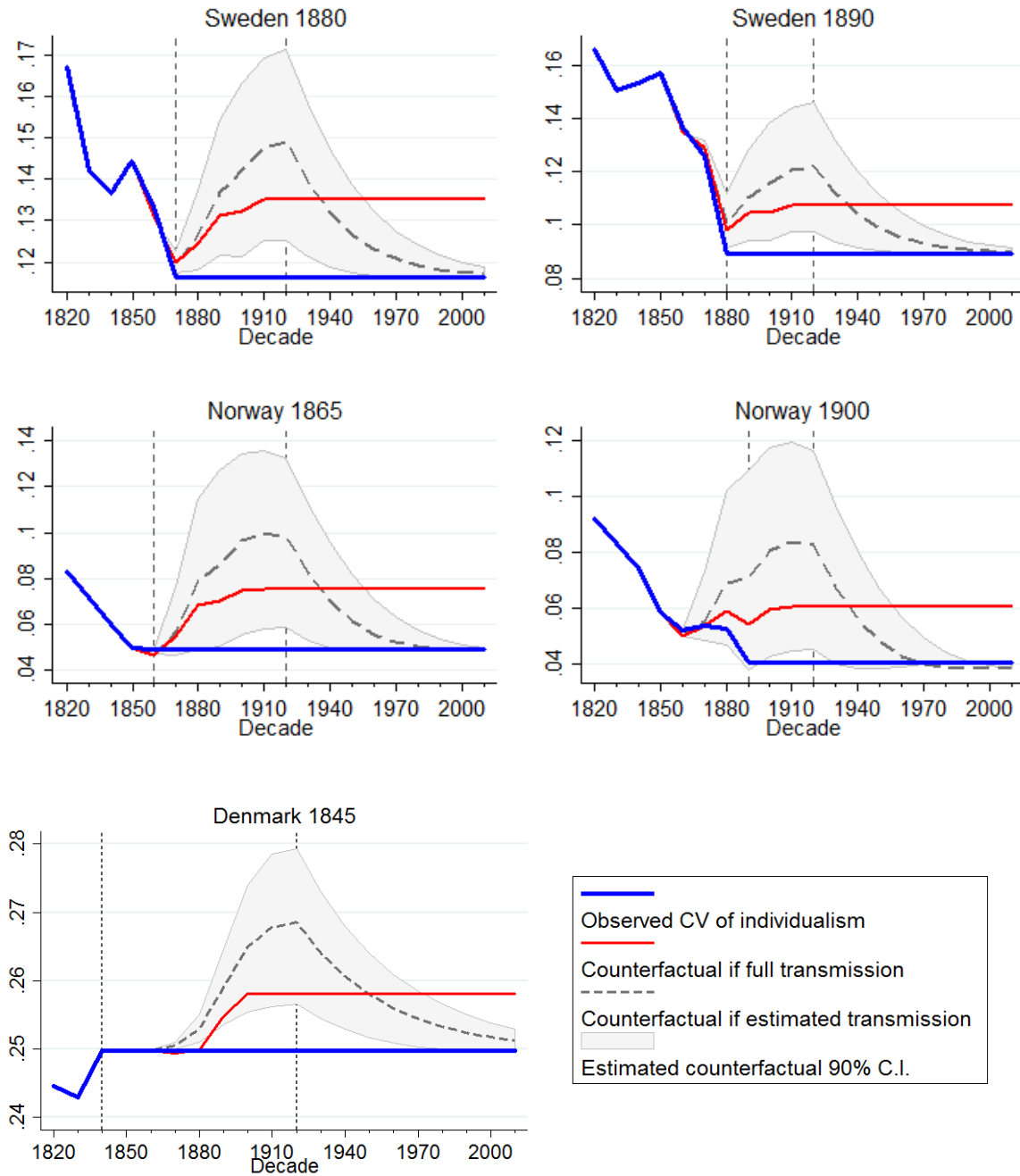
Notes: Paths of CV of individualism. See section 8 for details.

## Table A.1: Validation using Contemporary Indicators: Countries and Sources

| Country | Country level Top one | Country level Top ten | District level Top one | District level Top ten | Source |
|---|---|---|---|---|---|
| South Africa | x | x | | | Imbizo Centre |
| Argentina | x | | | | Registro Civil de la Ciudad de Buenos Aires (via lanacion.com) |
| | | | | | (Only Buenos Aires) |
| Canada | x | x | x | x | Governments of Alberta, British Columbia, Manitoba, Nova Scotia, |
| | | | | | Ontario, Québec, Saskatchewan, and Northern Territories |
| Chile | x | x | | | Chile Registro Civil |
| Colombia | x | x | | | Registraduria Nacional del Estado Civil |
| Costa Rica | x | x | | | Costa Rica Registro Civil (via crhoy.com) |
| Guatemala | x | | | | Registro Nacional de las Personas (Renap) (via soy502.com) |
| Jamaica | x | | | | Registrar General's Department (RGD) (via jamaicaobserver.com) |
| Peru | x | | | | Registro Nacional de Identificación y Estado Civil (via elcomercio.pe) |
| | | | | | (Only 3 first months of 2015) |
| Puerto Rico | x | x | | | US Social Security |
| United States | x | x | x | x | US Social Security |
| Uruguay | x | x | | | Registro Civil de Montevideo |
| | | | | | (Only Montevideo) |
| Armenia | x | x | | | The Statistical Committee of the Republic of Armenia |
| Azerbaijan | x | | | | Azerbaijan Ministry of Justice (via xezerxeber.az) |
| China | x | x | | | National Citizenship Information System, reported by Beijing News |
| | | | | | (and stored at Wikipedia) |
| Iran | x | x | | | Civil Registration (via alamto.com) |
| Israel | x | x | | | Central Bureau of Statistics Israel |
| | | | | | (Only Jewish population) |
| Japan | x | x | | | Ogihara et al. (2015) |
| Philippines | x | x | | | Philippine Statistics Authority |
| South Korea | x | x | | | Korea Supreme Court (via Wikipedia) |
| Albania | x | | | | Sipas Institutit Shqiptar (via Balkanwb.com) |
| Austria | x | x | x | x | Statistik Austria |
| Belgium | x | x | x | x | Statbelă(Algemene Directie Statistiek-Statistics Belgium) |
| Bosnia and Herzegovina | x | x | | | Federalni zavod za statistiku (official statistical office) |
| Bulgaria | x | | | | National Statistical Institute (NSI) (via nancy.cc) |
| Czech Republic | x | x | x | x | Ministry of the Interior of the Czech Republic |
| Denmark | x | x | x | x | Danmarks Statistik |
| Estonia | x | x | | | Estonia Ministry of the Interior (via nancy.cc) |
| Finland | x | x | | | Population Register Centre |
| France | x | x | x | x | Institut national de la statistique et des études économiquesă(INSEE) |
| Germany | x | x | | | blog.beliebte-vornamen.de (independent collection of city data) |
| Hungary | x | x | | | Belügyminisztérium Nyilvántartások Vezetéséért Felelos Helyettes |
| | | | | | Államtitkárság (via behindthename.com) |
| Iceland | x | x | | | Statistics Iceland |
| Ireland | x | x | x | x | The Central Statistics Office of Ireland (SCO) |
| Italy | x | x | x | x | Istat - Istituto Nazionale di Statisticaă |
| Latvia | x | x | | | Office of Citizenship and Migration Affairs (via bnn-news.com) |
| Lithuania | x | x | | | Population Register Service (via vardai.vlkk.lt) |
| Malta | x | x | | | Malta National Statistics Office (via nancy.cc) |
| | | | | | (Phonetic spellings combined) |
| Moldova | x | x | | | Moldova Civil Status Service (via noi.md) |
| Netherlands | x | x | x | x | Meertens Institut |
| Norway | x | x | x | x | StatBank Norway |
| Poland | x | x | | | Ministry of Interior and Administration (via behindthename.com) |
| Portugal | x | x | | | Instituto dos Registos e do Notariado (via behindthename.com) |
| Romania | x | x | | | Ministry of Administration and Interior (via behindthename.com) |
| Slovakia | x | x | | | Interior Ministry (via spectator.sme.sk) |
| Slovenia | x | x | | | Republic of Slovenia Statistical Office RS |
| Spain | x | x | x | x | Instituto National de Estadística |
| Sweden | x | x | x | | Statistics Sweden |
| Switzerland | x | x | x | x | Federal Statistical Office Switzerland |
| United Kingdom | x | x | x | | Office for National Statistics |
| Australia | x | x | x | x | Registers of Births, Deaths and Marriages (from NSW, Victoria, |
| | | | | | Queensland, South Australia, Western Australia, and Northern Territory) |

Notes: This table lists the countries for which contemporary first names data was collected. The sources for the collection are also noted along with secondary source websites in parentheses where relevant.

Table A.2: Validation using Contemporary Indicators: Cross-Country

|  | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dep. Variable | | Hofstede Individualism Index | | | |
| Baby name | 0.355*** | 0.350*** | 0.472*** | 0.382*** | 0.309*** |
| uncommonness | (0.093) | (0.091) | (0.092) | (0.127) | (0.094) |
| Uncommonness measure | Baseline | Baseline | Baseline | Baseline | Extended |
| Fixed Effects | None | Continent | Continent | Continent | Continent |
| Controls | N | N | Y | Y | Y |
| Add. Hofstede Controls | N | N | N | All | All |
| Observations | 40 | 40 | 40 | 37 | 40 |
| R-squared | 0.16 | 0.47 | 0.68 | 0.80 | 0.80 |

Notes: OLS regressions assessing the validity of the uncommon first names measure as an indicator of individualism. Baby name uncommonness is calculated at the country level as the share of newborns not recieving one among the most popular ten male or female names of their birth year in 2015. In column (5) baby name first name uncommonness is added for three additional countries based on the single most popular first name. These measures are compared to the Hofstede (2001) Individualism index. Controls include the log number of newborns, ethnic fractionalization, genetic diversity, and a latin language dummy. Additional Hofstede cultural variables include Power Distance, Masculinity, Uncertainty Avoidance, Long Term Orientation, and Indulgence. All available data across the globe is included in these regressions. All variables have been standardized before regression. Robust standard errors are shown in the parentheses. Significance levels are: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.3: Validation using Contemporary Indicators: Additional Country Indicators

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | | Survey Based Measures | | | | Linguistic Characteristics | | |
| Dep. Variable | Embeddedness | Affective Autonomy | Intellectual Autonomy | In-Group Favouritism | Social Tightness | Pronoun Drop (KK 1998) | Pronoun Drop (AD 2014) | Subject Prominence |
| Baby name uncommonness | -0.342* | 0.451** | 0.223 | -0.503*** | -0.479*** | -0.423*** | -0.660*** | 0.412*** |
| | (0.182) | (0.185) | (0.195) | (0.138) | (0.172) | (0.151) | (0.169) | (0.099) |
| Observations | 30 | 30 | 30 | 50 | 37 | 38 | 45 | 42 |
| R-squared | 0.27 | 0.32 | 0.27 | 0.42 | 0.52 | 0.59 | 0.47 | 0.62 |

Notes: OLS regressions. The unit of observation is a country, and baby name uncommonness is calculated as the share of newborns not recieving the most popular male or female names of their birth year in 2015. Embeddedness, Intellectual and Affective Autonomy are from Schwartz (1994, 2004), In-Group Favouritism from Van de Vliert (2011), and Social Tightness from Uz (2015). The original Pronoun Drop dummy is from Kashima and Kashima (1998) and the extended from Abdurazokzoda and Davis (2014). The Subject Prominence indicator of individualism is from Meyer-Schwarzenberger (2015). Controls include continent fixed effects, log number of newborns, ethnic fractionalization, genetic diversity, and a latin language dummy. All variables have been standardized before regression. Robust standard errors are shown in the parentheses with the following significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

## Table A.4: Validation using Contemporary Indicators: Within-Country

|  | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dep. Variable | | | Emancipative values (EVS/WVS) | | |
| Baby name uncommonness | 0.051** (0.025) | 0.981*** (0.310) | 0.049* (0.025) | 0.854*** (0.222) | 0.602*** (0.148) |
| Uncommonness measure | Baseline | Baseline | Extended | Extended | Extended |
| Fixed Effects | Country | Country | Country | Country | Country |
| Controls | Y | Y | Y | Y | Y |
| Sample | All | NO, DK | All | NO, DK, SE | SE |
| Countries | 13 | 2 | 14 | 3 | 1 |
| Observations | 55,422 | 3,210 | 58,571 | 6,301 | 6,302 |
| R-squared | 0.20 | 0.18 | 0.24 | 0.37 | 0.37 |

Notes: Notes: OLS regressions assessing the validity of the uncommon first names measure as an indicator of individualism. The unit of observation is an individual surveyed in WVS/EVS and the dependent variable "Emancipative Values" is calculated according to Welzel (2013) at the individual level. Baby name uncommonness in 2015 is here calculated for the districts in which the WVS/EVS respondents live. In columns (3)-(5) baby name first name uncommonness is added for Sweden, calculated based on the single most popular first name in each district. Controls include the log number of newborns, respondent age, age squared and gender. All variables have been standardized before regression. Robust standard errors, clustered at the district level, are shown in the parentheses. Significance levels are: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

## Table A.5: Validation using Contemporary Indicators: Additional Individual Indicators

|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Personal Autonomy | | Gender Equality | | Lifestyle Liberty | | Voice of the People | |
| Baby name uncommonness | 0.044** (0.021) | 0.570*** (0.140) | 0.015 (0.016) | 0.463 (0.284) | 0.042* (0.024) | 0.755*** (0.278) | -0.007 (0.008) | -0.002 (0.054) |
| Countries | 15 | 3 | 8 | 3 | 14 | 3 | 15 | 3 |
| Observations | 65,038 | 7,503 | 19,118 | 2,199 | 60,358 | 6,347 | 67,848 | 7,670 |
| R-squared | 0.11 | 0.12 | 0.14 | 0.07 | 0.19 | 0.27 | 0.05 | 0.02 |

Notes: OLS regressions. The unit of observation is an individual surveyed in WVS/EVS and the dependent variables are elements of the Welzel (2013) "Emancipative Values" indicator. Baby name uncommonness is here calculated for the districts in which the WVS/EVS respondents live. All countries are pooled, and where the top ten most popular names are not know, the top most popular is used in calculations. Controls include country fixed effects, log number of newborns, respondent age, age squared and gender). All variables have been standardized before regression. Robust standard errors are shown in the parentheses and clustered at the district level. Significance levels are: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

## Table A.6: Validation using Swedish newspaper language, 1780-1900

|  | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. Variable |  |  | Relative Singular Pronoun Use |  |  |  |
| Baby name | 0.433** | 0.463** | 0.455* | -0.015 | -0.016 | 0.021 |
| uncommonness | (0.213) | (0.219) | (0.234) | (0.077) | (0.076) | (0.097) |
| Uncommonness measure | Baseline | Baseline | Baseline | Non-top-1 | Non-top-2 | Non-top-3 |
| Fixed Effects | Y | Y | Y | Y | Y | Y |
| Controls | None | Paper size | Paper and pop. size | None | Paper size | Paper and pop. size |
| Observations | 178 | 178 | 178 | 178 | 178 | 178 |
| R-squared | 0.92 | 0.92 | 0.92 | 0.92 | 0.92 | 0.92 |

Notes: In this table the relative singular pronoun use, singular/(singular+plural), of a given newspaper and decade is regressed on the uncommonness of names of individuals born during the same decade. In columns (1)-(3) first name uncommonness is defined as not being among the ten most popular male or female names in a birth cohort, and in columns (4)-(6) as not being the single most popular male or female name in the cohort. Fixed effects are added for newspaper title, decade, census year, and age of cohort. Controls for the (ln) number of newspaper pages published and people born during the decade are added one at the time. Robust standard errors clustered at the district level in parentheses *** p<0.01, ** p<0.05, * p<0.1.

## Table A.7: Most Popular Child First Names by Gender and Census

| **Denmark** |  | Year 1845 |  |  |  | Year 1880 |  |  |
|---|---|---|---|---|---|---|---|---|
|  | Female |  | Male |  | Female |  | Male |  |
| Rank | Name | Name share | Name | Name share | Name | Name share | Name | Name share |
| 1 | Ane | 0.154 | Jens | 0.109 | Ane | 0.118 | Jens | 0.100 |
| 2 | Maren | 0.086 | Niels | 0.101 | Karen | 0.055 | Hans | 0.091 |
| 3 | Karen | 0.085 | Hans | 0.100 | Anna | 0.048 | Niels | 0.089 |
| 4 | Johanne | 0.041 | Peder | 0.067 | Maren | 0.048 | Anders | 0.037 |
| 5 | Anne | 0.040 | Anders | 0.045 | Marie | 0.048 | Carl | 0.034 |

| **Norway** |  | Year 1865 |  |  |  | Year 1910 |  |  |
|---|---|---|---|---|---|---|---|---|
|  | Female |  | Male |  | Female |  | Male |  |
| Rank | Name | Name share | Name | Name share | Name | Name share | Name | Name share |
| 1 | Anne | 0.063 | Ole | 0.094 | Anna | 0.043 | Ole | 0.029 |
| 2 | Anna | 0.038 | Hans | 0.054 | Astrid | 0.026 | Johan | 0.028 |
| 3 | Karen | 0.033 | Johan | 0.052 | Gudrun | 0.023 | Karl | 0.024 |
| 4 | Ingeborg | 0.028 | Peder | 0.029 | Borghild | 0.023 | Hans | 0.023 |
| 5 | Marie | 0.021 | Lars | 0.026 | Margit | 0.022 | Einar | 0.022 |

| **Sweden** |  | Year 1880 |  |  |  | Year 1900 |  |  |
|---|---|---|---|---|---|---|---|---|
|  | Female |  | Male |  | Female |  | Male |  |
| Rank | Name | Name share | Name | Name share | Name | Name share | Name | Name share |
| 1 | Anna | 0.158 | Johan | 0.123 | Anna | 0.117 | Karl | 0.132 |
| 2 | Maria | 0.045 | Carl | 0.107 | Ester | 0.038 | Johan | 0.078 |
| 3 | Emma | 0.045 | Karl | 0.094 | Elsa | 0.032 | Gustaf | 0.060 |
| 4 | Ida | 0.041 | Gustaf | 0.050 | Signe | 0.029 | Erik | 0.059 |
| 5 | Johanna | 0.037 | Anders | 0.046 | Elin | 0.028 | Nils | 0.043 |

## Table A.8: Summary Statistics in Full and Matched Emigration Database

|  | All emigrants | Matched emigrants | Difference (match-pop) |
|---|---|---|---|
| *Individual level matching* | | | |
| Emigration year | 1901.86 | 1903.019 | -1.906*** |
| Birth year | 1878.902 | 1880.586 | -2.769*** |
| Age at emigration | 22.95788 | 22.43285 | 0.863*** |
| Emigrate as child | 0.080232 | 0.080837 | -0.000994 |
| Popularity of first name among migrants | 0.028052 | 0.0265 | 0.00255*** |
| Spelling mistake | 0.878975 | 0.908337 | -0.0483*** |
| Rare name | 0.012617 | 0.006752 | 0.00964*** |
| N | 389805 | 152690 | |
| *First names matching* | | | |
| Emigration year | 1901.86 | 1901.77 | 1.439*** |
| Birth year | 1878.902 | 1878.797 | 1.682*** |
| Age at emigration | 22.95788 | 22.97304 | -0.243*** |
| Emigrate as child | 0.080232 | 0.080153 | 0.00127 |
| Popularity of first name among migrants | 0.028052 | 0.029703 | -0.0264*** |
| Spelling mistake | 0.878975 | 0.91693 | -0.607*** |
| Rare name | 0.012617 | 0.002471 | 0.162*** |
| N | 389805 | 365438 | |

Notes: Characteristics of matched and population samples in the individual linked (done based on first name, last name, gender, birth year, and district of residence) and in the aggregate linked (based on first name, gender, birth year and residence only) data.

## Table A.9: Comparing Change in First Name Carriers with Emigration of Name Carriers

|  | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. Var. | | | Change in Population Census | | | |
| Emigrants | -2.583*** | -2.575*** | -2.621*** | -0.556*** | -0.553*** | -0.574*** |
|  | (0.093) | (0.093) | (0.094) | (0.115) | (0.115) | (0.116) |
| Total Number in Census | | | | -0.162*** | -0.162*** | -0.162*** |
|  | | | | (0.007) | (0.007) | (0.007) |
| Fixed Effects | None | Census | Cohort, District, Census | None | Census | Cohort, District, Census |
| Observations | 848,907 | 848,907 | 848,907 | 848,907 | 848,907 | 848,907 |
| R-squared | 0.23 | 0.24 | 0.25 | 0.43 | 0.43 | 0.44 |

Notes: Notes: OLS regressions comparing the change from one population census to the next in the number of people born in the same cohort, residing in the same district and carrying the same first name with the number of emigrants with the same characteristics. Robust standard errors in parentheses with significance levels *** p<0.01, ** p<0.05, * p<0.1

Table A.10: Comparison of Historical Emigration Data Sets (Sweden and Norway)

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dep. Var. | | Emigrants (national accounts) | | |
| Passenger list emigrants | 1.392*** | 1.108*** | | |
|  | (0.074) | (0.073) | | |
| Passenger list emigrants | | | 1.555*** | 1.242*** |
| (with personal information) | | | (0.082) | (0.086) |
| Fixed Effects | Country | District | Country | District |
| Observations | 240 | 240 | 240 | 240 |
| R-squared | 0.73 | 0.90 | 0.72 | 0.90 |

Notes: OLS regressions comparing the absolute number of emigrants recorded each decade 1860-1910 across 40 Swedish and Norwegian subnational districts. Robust standard errors in parentheses with significance levels *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A.11: Comparison of Historical Emigration Data Sets (Denmark)

|  | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. Var. | | Emigrants (Hamborg lists) | | | Emigrants (joint lists) | |
| Passenger list emigrants | 0.023 | -0.049 | | | 1.001*** | 0.957*** |
|  | (0.022) | (0.040) | | | (0.020) | (0.049) |
| Joint lists emigrants | | | 0.051*** | -0.015 | | |
|  | | | (0.018) | (0.037) | | |
| Fixed Effects | Country | District | Country | District | Country | District |
| Observations | 60 | 60 | 60 | 60 | 60 | 60 |
| R-squared | 0.01 | 0.46 | 0.07 | 0.42 | 0.97 | 0.97 |

Notes: OLS regressions comparing the absolute number of Danish emigrants recorded in different passenger lists for each decade 1860-1910 . Robust standard errors in parentheses with significance levels *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A.12: Individualistic Selection of Scandinavian Migrants: All Controls

|  | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. Variable | | Migration dummy | | | Emigration | Out-migration |
| Individualism | 0.024*** | 0.022*** | 0.019*** | 0.018*** | 0.009*** | 0.008*** |
|  | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Age | | 0.009*** | 0.009*** | 0.009*** | 0.004*** | 0.005*** |
|  | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Age squared | | -0.000*** | -0.000*** | -0.000*** | -0.000*** | -0.000*** |
|  | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Living at home (d) | | -0.006*** | -0.005*** | -0.007*** | -0.005*** | -0.001 |
|  | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Previous internal migration (d) | | 0.078*** | 0.075*** | 0.067*** | -0.007*** | 0.074*** |
|  | | (0.001) | (0.001) | (0.001) | (0.000) | (0.001) |
| Household head (d) | | -0.002 | 0.002 | 0.004 | -0.001 | 0.004* |
|  | | (0.003) | (0.003) | (0.003) | (0.002) | (0.003) |
| Ln (number of brothers) | | 0.004*** | 0.002*** | 0.002*** | 0.001*** | 0.001*** |
|  | | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Ln (birth order) | | 0.004*** | 0.004*** | 0.005*** | 0.005*** | 0.000 |
|  | | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) |
| Work (d) | | 0.003 | 0.003 | -0.000 | 0.008*** | -0.008*** |
|  | | (0.002) | (0.002) | (0.002) | (0.001) | (0.002) |
| Related to household head (d) | | -0.001 | 0.001 | 0.001 | -0.002*** | 0.003*** |
|  | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Surname commonness | | | 0.600*** | 0.563*** | 0.252*** | 0.313*** |
|  | | | (0.004) | (0.004) | (0.003) | (0.003) |
| Family given name (d) | | | -0.011*** | -0.011*** | -0.007*** | -0.004*** |
|  | | | (0.000) | (0.000) | (0.000) | (0.000) |
| Biblical name (d) | | | 0.001 | 0.001*** | 0.000* | 0.001** |
|  | | | (0.000) | (0.000) | (0.000) | (0.000) |
| Scandinavian origin name (d) | | | 0.002*** | 0.002*** | 0.001*** | 0.001*** |
|  | | | (0.000) | (0.000) | (0.000) | (0.000) |
| Ln (household size) | | | 0.005*** | 0.004*** | 0.003*** | 0.001*** |
|  | | | (0.000) | (0.001) | (0.000) | (0.000) |
| Household head characteristics: | | | | | | |
| Age | | | | 0.000*** | 0.000** | 0.000*** |
|  | | | | (0.000) | (0.000) | (0.000) |
| Male (d) | | | | 0.002* | 0.002*** | 0.000 |
|  | | | | (0.001) | (0.001) | (0.001) |
| Foreigner (d) | | | | -0.003*** | -0.001 | -0.001* |
|  | | | | (0.001) | (0.001) | (0.001) |
| Previous internal migration (d) | | | | 0.012*** | -0.000 | 0.012*** |
|  | | | | (0.000) | (0.000) | (0.000) |
| Engaged in farming (d) | | | | -0.011*** | -0.001*** | -0.010*** |
|  | | | | (0.000) | (0.000) | (0.000) |
| Ln (number of servants) | | | | 0.004*** | -0.004*** | 0.008*** |
|  | | | | (0.000) | (0.000) | (0.000) |
| Poor relief (d) | | | | -0.011*** | -0.002** | -0.008*** |
|  | | | | (0.001) | (0.001) | (0.001) |
| Asset holder (d) | | | | -0.005*** | 0.001*** | -0.006*** |
|  | | | | (0.000) | (0.000) | (0.000) |
| Work (d) | | | | 0.002*** | 0.002*** | 0.000 |
|  | | | | (0.001) | (0.000) | (0.001) |
| Non-manual work (d) | | | | 0.011*** | -0.002*** | 0.013*** |
|  | | | | (0.001) | (0.000) | (0.000) |
| Skilled work (d) | | | | 0.002*** | 0.001*** | 0.001*** |
|  | | | | (0.000) | (0.000) | (0.000) |
| Observations | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 |
| R-squared | 0.04 | 0.05 | 0.05 | 0.05 | 0.02 | 0.06 |
| Mean of dep. var. | 0.0726 | 0.0726 | 0.0726 | 0.0726 | 0.0370 | 0.0360 |
| St.dev. of dep. var. | 0.259 | 0.259 | 0.259 | 0.259 | 0.189 | 0.186 |

Notes: The same OLS estimations as in Table 1, but here the coefficients of all controls are reported. The unit of analysis is an individual male that were below the age of 15 in any of the historical population censuses. All regressions include birth district, birth cohort, and population census fixed effects. Controls are added to capture relevant child and household characterstics. Robust standard errors are shown in the parentheses with the following significance levels: *** p< 0.01, ** p< 0.05, * p< 0.1.

## Table A.13: Individualistic Selection of Scandinavian Migrants: Probit Regressions

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. Variable | | Migration dummy | | | Emigration | Out-migration |
| Individualism | 0.182*** | 0.175*** | 0.154*** | 0.145*** | 0.127*** | 0.128*** |
| (first name uncommonness) | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) | (0.003) |
| | | | | | | |
| Controls: | | | | | | |
| Individual | N | Y | Y | Y | Y | Y |
| Cultural (household) | N | N | Y | Y | Y | Y |
| Economic (household) | N | N | N | Y | Y | Y |
| | | | | | | |
| Observations | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 2,672,646 |
| Mean of dep. var. | 0.0726 | 0.0726 | 0.0726 | 0.0726 | 0.0370 | 0.0525 |
| St.dev. of dep. var. | 0.259 | 0.259 | 0.259 | 0.259 | 0.189 | 0.223 |

Notes: Probit estimations. Specifications are the same as in Table 1. The unit of analysis is an individual male that were below the age of 15 in any of the historical population censuses. All regressions include birth district, birth cohort, and population census fixed effects. Controls are added to capture relevant child and household characterstics. Robust standard errors are shown in the parentheses with the following significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

## Table A.14: Individualistic Selection of Scandinavian Migrants: Parental Characteristics

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. Variable | | Migration dummy | | | Emigration | Out-migration |
| Individualism | 0.024*** | 0.022*** | 0.019*** | 0.018*** | 0.010*** | 0.008*** |
| (first name uncommonness) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | | | | | | |
| Controls: | | | | | | |
| Individual | N | Y | Y | Y | Y | Y |
| Cultural (household) | N | N | Y | Y | Y | Y |
| Economic (household) | N | N | N | Y | Y | Y |
| | | | | | | |
| Observations | 3,645,969 | 3,645,969 | 3,645,969 | 3,645,969 | 3,645,969 | 3,645,969 |
| R-squared | 0.04 | 0.05 | 0.05 | 0.05 | 0.02 | 0.06 |
| Mean of dep. var. | 0.0719 | 0.0719 | 0.0719 | 0.0719 | 0.0371 | 0.0352 |
| St.dev. of dep. var. | 0.258 | 0.258 | 0.258 | 0.258 | 0.189 | 0.184 |

Notes: The same OLS estimations as in Table 1, but here parent characteristics are used as household controls. The unit of analysis is an individual male that were below the age of 15 in any of the historical population censuses. All regressions include birth district, birth cohort, and population census fixed effects. Controls are added to capture relevant child and household characterstics. Robust standard errors are shown in the parentheses with the following significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.15: Individualistic Selection of Scandinavian Migrants: Alternative First Names

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| Dep. Variable | Migration | Emigration | Out-migration | Migration | Emigration | Out-migration | Migration | Emigration | Out-migration |
| Individualism (first name uncommonness) | 0.210*** (0.003) | 0.115*** (0.002) | 0.095*** (0.002) | -0.009 (0.006) | 0.013*** (0.003) | -0.022*** (0.006) | -0.021*** (0.007) | 0.009*** (0.003) | -0.030*** (0.006) |
| Names used | Continuous measure | | | Full first name | | | Phonetically spelled first name | | |
| Controls: | | | | | | | | | |
| Individual | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| Cultural (household) | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| Economic (household) | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| Observations | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 |
| R-squared | 0.06 | 0.02 | 0.06 | 0.05 | 0.02 | 0.06 | 0.05 | 0.02 | 0.06 |
| Mean of dep. var. | 0.0726 | 0.0370 | 0.0360 | 0.0726 | 0.0370 | 0.0360 | 0.0726 | 0.0370 | 0.0360 |
| St.dev. of dep. var. | 0.259 | 0.189 | 0.186 | 0.259 | 0.189 | 0.186 | 0.259 | 0.189 | 0.186 |

Notes: The same OLS estimations as in Table 1, but using alternative measures of inherited individualism. The unit of analysis is an individual male that were below the age of 15 in any of the historical population censuses. All regressions include birth district, birth cohort, and population census fixed effects. Controls are added to capture relevant child and household characterstics. Robust standard errors are shown in the parentheses with the following significance levels: *** p< 0.01, ** p< 0.05, * p< 0.1.

Table A.16: Individualistic Selection of Scandinavian Migrants: Birthplace Sample

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. Variable | | | Migration dummy | | Out-migration | Emigration |
| Individualism | 0.015*** | 0.014*** | 0.012*** | 0.011*** | 0.003*** | 0.008*** |
| (first name uncommonness) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Controls: | | | | | | |
| Individual | N | Y | Y | Y | Y | Y |
| Cultural (household) | N | N | Y | Y | Y | Y |
| Economic (household) | N | N | N | Y | Y | Y |
| Observations | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 | 3,897,243 |
| R-squared | 0.05 | 0.06 | 0.06 | 0.06 | 0.05 | 0.06 |
| Mean of dep. var. | 0.046 | 0.046 | 0.046 | 0.046 | 0.01 | 0.036 |
| St.dev. of dep. var. | 0.209 | 0.209 | 0.209 | 0.209 | 0.0991 | 0.187 |

Notes: The same OLS estimations as in Table 1, but using the birthplace-matched sample. The unit of analysis is an individual male that were below the age of 15 in any of the historical population censuses. All regressions include birth district, birth cohort, and population census fixed effects. Controls are added to capture relevant child and household characterstics. Robust standard errors are shown in the parentheses with the following significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.17: Individualistic Selection of Scandinavian Migrants: Aggregate Level

|  | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | | | | | | |
| | Migration rate among carriers of same given name: | | | | | |
| Dep. Variable | All migrants | | | | Out-migrants | Emigrants |
| | | | | | | |
| Name commonnness | 0.023*** | 0.023*** | 0.012*** | 0.009*** | 0.004*** | 0.006*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | | | | | | |
| Controls: | | | | | | |
| Individual | N | Y | Y | Y | Y | Y |
| Names related | N | N | Y | Y | Y | Y |
| Household | N | N | N | Y | Y | Y |
| | | | | | | |
| Observations | 8,493,221 | 8,493,221 | 8,493,221 | 8,493,221 | 8,493,221 | 8,493,221 |
| R-squared | 0.33 | 0.34 | 0.34 | 0.35 | 0.40 | 0.18 |
| Mean of dep. var. | 0.149 | 0.149 | 0.149 | 0.149 | 0.076 | 0.073 |
| St.dev. of dep. var. | 0.211 | 0.211 | 0.211 | 0.211 | 0.147 | 0.127 |

Notes: The same OLS estimations as in Table 1, but using the first-name-matched sample. The unit of analysis is an individual (male or female) that were below the age of 15 in any of the historical population censuses. All regressions include birth district, birth cohort, and population census fixed effects. Controls that capture relevant child and household characterstics are calculated as averages over individuals that share the same first name, birth place, and birth decade. Robust standard errors are shown in the parentheses with the following significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.18: Assimilation of Scandinavian Immigrants in the 1900 and 1910 US Samples

| Dep. var. | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| | No English spoken | | Spouse same nationality | | State concentration of own nationality | | First name ethnic distinctiveness of child born in the US | | | |
| Individualism (first name uncommonness) | -0.135*** | -0.084** | -0.198*** | -0.156*** | -0.028*** | -0.012*** | -0.236*** | -0.147*** | -0.015 | -0.015 |
| | (0.040) | (0.041) | (0.047) | (0.047) | (0.003) | (0.003) | (0.032) | (0.033) | (0.018) | (0.019) |
| Years spend in the US | -0.008*** | -0.008*** | -0.004*** | -0.004*** | 0.000*** | 0.000 | -0.001*** | -0.002*** | -0.002*** | -0.002*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| | | | | | | | | | | |
| Sample | All | All | All | All | All | All | Fathers | Fathers | Mothers | Mothers |
| Additional controls | N | Y | N | Y | N | Y | N | Y | N | Y |
| Observations | 47,758 | 47,758 | 47,758 | 47,758 | 47,758 | 47,758 | 37,335 | 37,335 | 32,371 | 32,371 |
| R-squared | 0.18 | 0.19 | 0.54 | 0.54 | 0.15 | 0.21 | 0.05 | 0.05 | 0.06 | 0.06 |
| Mean of dep. var. | 0.0773 | 0.0773 | 0.458 | 0.458 | 0.0268 | 0.0268 | 0.0561 | 0.0561 | 0.0682 | 0.0682 |
| St.dev. of dep. var. | 0.267 | 0.267 | 0.498 | 0.498 | 0.0245 | 0.0245 | 0.128 | 0.128 | 0.145 | 0.145 |

Notes: OLS regressions. Unit of observation is a Scandinavian immigrant in the Unites States 1900 and 1910 population census. Inherited individualism is measured as the share of a birth decade cohort in the 1880 Danish and the 1900 Swedish and Norwegian censuses that carried the same first name. Non-assimilation is given by the lack of English proficiency, by marrying someone of own nationality, settling in a state with a high concentration of own nationality, and by giving a child a more distinctively Scandinavian sounding name (compared to the entire US census population of 1900 and 1910). All regressions contain controls for age, age squared, a gender, marriage, marriage in the US, census as well as cohort fixed effects. In the last four columns the number of siblings, age, age squared and gender of the child is also controlled for. Additional controls include a spelling mistake proxy, first name Biblical or Scandinavian origin dummies, indicators for living on a farm or in an urban area, the Duncan Socioeconomic index, and the IPUMS occupational income score. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A.19: Assimilation of Scandinavian Immigrants in the 1900 and 1910 US Samples (Fixed Effects Reg.)

|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var. | No English spoken | | Spouse same nationality | | First name ethnic distinctiveness of child born in the US | | | |
| Individualism (first name uncommonness) | -0.067* (0.040) | -0.070* (0.040) | -0.133*** (0.047) | -0.152*** (0.047) | -0.140*** (0.034) | -0.139*** (0.034) | -0.027 (0.018) | -0.034* (0.018) |
| Years spend in the US | -0.008*** (0.000) | -0.008*** (0.000) | -0.004*** (0.000) | -0.004*** (0.000) | -0.002*** (0.000) | -0.002*** (0.000) | -0.002*** (0.000) | -0.002*** (0.000) |
| Sample | All | All | All | All | Fathers | Fathers | Mothers | Mothers |
| Fixed effects | State | County | State | County | State | County | State | County |
| Observations | 47,758 | 47,758 | 47,758 | 47,758 | 37,335 | 37,335 | 29,419 | 29,419 |
| R-squared | 0.20 | 0.20 | 0.55 | 0.55 | 0.06 | 0.06 | 0.05 | 0.05 |
| Mean of dep. var. | 0.0773 | 0.0773 | 0.458 | 0.458 | 0.0561 | 0.0561 | 0.0602 | 0.0602 |
| St.dev. of dep. var. | 0.267 | 0.267 | 0.498 | 0.498 | 0.128 | 0.128 | 0.132 | 0.132 |

Notes: OLS regressions. Unit of observation is a Scandinavian immigrant in the Unites States 1900 and 1910 population census. Inherited individualism is measured as the share of a birth decade cohort in the 1880 Danish and the 1900 Swedish and Norwegian censuses that carried the same first name. Non-assimilation is given by the lack of English proficiency, by marrying someone of own nationality, and by giving a child a more distinctively Scandinavian sounding name (compared to the entire US census population of 1900 and 1910). All regressions contain controls for age, age squared, a gender, marriage, marriage in the US, spelling mistakes, first name Biblical and Scandinavian origin, indicators for living on a farm or in an urban area, the Duncan Socioeconomic index, and the IPUMS occupational income score. Census as well as cohort fixed effects are included as well. In the last four columns the number of siblings of the child, child age, age squared and gender is also controlled for. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

63

Table A.20: Deccennial Cultural Change 1850-1910: Robustness

| | Baseline | GMM estimation | Balanced panels of 100 yrs | Sample from 1860 onwards | Add cultural control variables | Adj. For cohort emigration | Incl. add. Danish emi. sources | Danish emi. through Hamburg |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Dep. variable | Individualism (first name uncommonness) | | | | | | | |
| Lagged individualism | 0.750*** | 0.366** | 0.704*** | 0.484*** | 0.718*** | 0.730*** | 0.757*** | 0.756*** |
| | (0.039) | (0.145) | (0.028) | (0.074) | (0.029) | (0.050) | (0.039) | (0.039) |
| Cultural shock of emigration | -0.897*** | -4.840*** | -0.828*** | -0.003 | -0.796*** | -1.070*** | -0.675** | -0.789*** |
| | (0.317) | (1.590) | (0.300) | (0.672) | (0.297) | (0.314) | (0.255) | (0.254) |
| - first lag | -0.626 | -3.983*** | -0.508 | -0.371 | -0.641 | -0.211 | -0.265 | -0.292 |
| | (0.420) | (1.405) | (0.452) | (0.484) | (0.402) | (0.404) | (0.266) | (0.267) |
| - second lag | -0.690* | -5.185*** | -0.579 | -1.005** | -1.127** | -0.420 | -0.419 | -0.479 |
| | (0.379) | (1.678) | (0.393) | (0.492) | (0.464) | (0.396) | (0.284) | (0.291) |
| - third lag | -0.094 | -3.142** | -0.142 | -0.661** | -0.553 | -0.413 | -0.039 | -0.071 |
| | (0.267) | (1.235) | (0.260) | (0.328) | (0.417) | (0.310) | (0.218) | (0.222) |
| - fourth lag | -0.985** | -3.883** | -0.945 | -0.975* | -0.678 | -0.572 | -0.943** | -0.975** |
| | (0.487) | (1.531) | (0.564) | (0.562) | (0.525) | (0.422) | (0.459) | (0.465) |
| Observations | 960 | 820 | 700 | 399 | 960 | 960 | 961 | 962 |
| Districts | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Census years | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| R-squared | 0.94 | | 0.94 | 0.96 | 0.95 | 0.94 | 0.94 | 0.94 |
| Mean of dep. var. | 0.480 | 0.490 | 0.495 | 0.541 | 0.480 | 0.480 | 0.480 | 0.480 |
| St.dev. of dep. var. | 0.146 | 0.145 | 0.146 | 0.141 | 0.146 | 0.142 | 0.146 | 0.146 |
| Sum of shock lags | -3.292 | -21.034 | -3.002 | -3.014 | -3.794 | -2.685 | -2.340 | -2.607 |
| (p-value) | 0.005 | 0.002 | 0.024 | 0.111 | 0.001 | 0.030 | 0.017 | 0.009 |
| Implied impact after 50 years | -1.495 | -2.196 | -1.206 | -0.761 | -1.512 | -1.084 | -1.173 | -1.276 |
| (p-value) | 0.012 | 0.113 | 0.033 | 0.048 | 0.009 | 0.062 | 0.031 | 0.021 |
| Hansen J-test (p-value) | | 0.197 | | | | | | |
| AR(2) test (p-value) | | 0.565 | | | | | | |

Notes: Robustness check of the results in Table 2. All specifications use the pooled historical population censuses and include fixed effects for districts, decades, lags, and census years, baseline controls (cohort number, gender, age, and number of first names), and emigration controls (pure and $(m/(1-))$-transformed rate of emigration, relative emigrant individualism, and their langs). See section 6.2 for description of robustness checks. P-values on the sum of shock lags and the 50-year implied impact are calculated according to the delta-method. The panels contain no unit roots (p-value of Phillips-Perron test is 0.000). Robust standard errors, clustered at the district level, are shown in the parentheses with the following significance levels: *** p< 0.01, ** p< 0.05, * p< 0.1.

## Table A.21: Deccennial Cultural Change: Alternative Emigration Variables

| Dep. variable | Individualism (first name uncommonness) based on: | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Phonetic spelling | | Full first names | | Gini coef. | | Popularity shares | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| **Dep. variable** | | | | | | | | |
| Lagged individualism | 0.748*** | 0.751*** | 0.929*** | 0.879*** | 0.693*** | 0.693*** | 0.729*** | 0.732*** |
| | (0.044) | (0.043) | (0.069) | (0.064) | (0.033) | (0.033) | (0.034) | (0.032) |
| Cultural shock of emigration | -0.674** | -0.929** | -0.296 | -0.156 | -0.109*** | -0.109*** | -0.224*** | -0.037 |
| | (0.282) | (0.364) | (0.371) | (0.302) | (0.031) | (0.031) | (0.061) | (0.996) |
| - first lag | -0.413 | -0.317 | 0.394 | 0.220 | -0.116*** | -0.116*** | -0.231*** | 2.874*** |
| | (0.376) | (0.415) | (0.466) | (0.368) | (0.038) | (0.038) | (0.074) | (0.965) |
| - second lag | -0.546 | -0.270 | 0.874 | 0.497 | -0.072** | -0.072** | -0.136** | -0.303 |
| | (0.367) | (0.476) | (0.584) | (0.466) | (0.030) | (0.030) | (0.060) | (1.233) |
| - third lag | -0.001 | -0.100 | 0.830* | 0.357 | -0.040 | -0.040 | -0.080 | -0.507 |
| | (0.244) | (0.308) | (0.489) | (0.385) | (0.027) | (0.027) | (0.052) | (1.140) |
| - fourth lag | -0.702* | -1.003*** | -0.679 | -0.063 | -0.090** | -0.090** | -0.180** | 1.359 |
| | (0.360) | (0.317) | (0.531) | (0.445) | (0.039) | (0.039) | (0.074) | (1.072) |
| Different calculation of shock | N | Y | N | Y | N | Y | N | Y |
| Observations | 960 | 960 | 960 | 960 | 960 | 960 | 960 | 960 |
| Districts | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Census years | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| R-squared | 0.95 | 0.95 | 0.96 | 0.96 | 0.88 | 0.88 | 0.89 | 0.89 |
| Mean of dep. var. | 0.431 | 0.431 | 0.710 | 0.710 | 0.977 | 0.977 | 0.955 | 0.955 |
| St.dev. of dep. var. | 0.141 | 0.141 | 0.132 | 0.132 | 0.011 | 0.011 | 0.022 | 0.022 |
| Sum of shock lags | -2.336 | -2.619 | 1.123 | 0.856 | -0.427 | -0.427 | -0.851 | 3.386 |
| *(p-value)* | *0.022* | *0.031* | *0.538* | *0.594* | *0.000* | *0.000* | *0.000* | *0.266* |
| Implied impact after 50 years | -1.041 | -1.246 | 0.875 | 0.608 | -0.150 | -0.150 | -0.338 | 1.422 |
| *(p-value)* | *0.038* | *0.033* | *0.572* | *0.606* | *0.001* | *0.001* | *0.000* | *0.306* |

Notes: Robustness check of the results in Table 2. All specifications use the pooled historical population censuses and include fixed effects for districts, decades, lags, and census years, baseline controls (cohort number, gender, age, and number of first names), and emigration controls (pure and $(m/(1-))$-transformed rate of emigration, relative emigrant individualism, and their langs). See section 6.2 for description of robustness checks. P-values on the sum of shock lags and the 50-year implied impact are calculated according to the delta-method. The panels contain no unit roots (p-value of Phillips-Perron test is 0.000). Robust standard errors, clustered at the district level, are shown in the parentheses with the following significance levels: *** $p< 0.01$, ** $p< 0.05$, * $p< 0.1$.

Table A.22: Emigration and Long-Run Cultural Change: All Census Records

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Dep. var. | Contemporary individualism (first name uncommonness) | | | | | | |
| Historical individualism | 0.037** | 0.041** | 0.046*** | 0.036** | 0.038*** | 0.036** | 0.043*** |
| | (0.017) | (0.017) | (0.017) | (0.014) | (0.014) | (0.014) | (0.016) |
| Cum. cultural shock of emigration | | -0.025* | -0.180*** | -0.174*** | -0.182*** | -0.177*** | -0.199*** |
| | | (0.013) | (0.065) | (0.061) | (0.059) | (0.053) | (0.060) |
| Controls: | | | | | | | |
| District - baseline | Y | Y | Y | Y | Y | Y | Y |
| Emigration - numbers and individualism | N | N | Y | Y | Y | Y | Y |
| District - geography | N | N | N | Y | Y | Y | Y |
| District - add. cultural traits | N | N | N | N | Y | Y | Y |
| Emigration - add. cultural traits | N | N | N | N | N | Y | Y |
| District - economic and demographic | N | N | N | N | N | N | Y |
| Observations | 140 | 140 | 140 | 140 | 140 | 140 | 140 |
| Districts | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Census years | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| R-squared | 0.98 | 0.98 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 |

OLS regressions on the long-run persistence of individualistic tratis and the influence from the cumulative emigration that took place during the Age of Mass Migration. The unit of observation is a district observed in a given historical population census. Contemporary individualism is measured as the share of newborns born in a district in year 2015 that recieved an uncommon (non-top-10 for Norway and Denmark, non-top-1 for Sweden) first name. Historical Individualism is calculated as the share of children carrying an uncommon (non-top-10) first name in the given census year. The cumulative emigration shock is the sum of all deccennial cultural shocks between the census year and the end of the Age of Mass Migration in 1920. All regressions include country and historical census year fixed effects. Historical controls are added sequentially and described in the main text. Mean and S.D. of the dependent variable is 0.917 and 0.059. Robust standard errors clustered at the district level are shown in the parentheses with the following significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.23: Long-Run Cult. Change: World/European Values Survey Results

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| Dep. var. | Emancipative Values | | Personal Autonomy | | Gender Equality | | Lifestyle Liberty | | Voice of the People | |
| Historical individualism | 0.141*** | 0.201*** | 0.215*** | 0.217** | 0.057 | -0.182*** | 0.272*** | 0.387*** | 0.001 | 0.007 |
| | (0.038) | (0.043) | (0.046) | (0.097) | (0.100) | (0.014) | (0.098) | (0.080) | (0.003) | (0.006) |
| Cum. cultural shock of emigration | -0.069*** | -0.120 | -0.099*** | -0.828** | -0.024 | 3.089*** | -0.135** | -0.130 | 0.001 | 0.005 |
| | (0.018) | (0.257) | (0.025) | (0.396) | (0.078) | (0.648) | (0.055) | (0.536) | (0.002) | (0.038) |
| | | | | | | | | | | |
| Hist. controls: | Baseline | All | Baseline | All | Baseline | All | Baseline | All | Baseline | All |
| Observations | 6,301 | 6,301 | 7,503 | 7,503 | 2,199 | 2,199 | 6,347 | 6,347 | 7,670 | 7,670 |
| Districts | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Historical census years | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| WVS/EVS census years | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| R-squared | 0.38 | 0.38 | 0.13 | 0.13 | 0.06 | 0.08 | 0.28 | 0.29 | 0.02 | 0.02 |
| | | | | | | | | | | |
| Mean of dep. var. | 0.483 | 0.483 | 0.631 | 0.631 | 0.857 | 0.857 | 0.630 | 0.630 | 0.002 | 0.002 |
| St.dev. of dep. var. | 0.155 | 0.155 | 0.291 | 0.291 | 0.149 | 0.149 | 0.311 | 0.311 | 0.022 | 0.022 |

Robustness check on table 3 using data from the World and European Values Surveys. The unit of observation is an individual surveyed in WVS/EVS and the dependent variables are elements of the Welzel (2013) "Emancipative Values" indicator. Regressors are measured at the district level. All regressions include country, historical census year, and WVS/EVS survey year fixed effects. Individual controls are added for age, age squared, and gender. Historical controls are described in the main text. Robust standard errors clustered at the district level are shown in the parentheses with the following significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

## Table A.24: Long-Run Cult. Change: Alternative Emigration Measures

| | Incl. add. Danish emi. sources | | Danish emi. through Hamburg | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Dep. var. | Contemporary individualism (first name uncommonness) | | | |
| Historical individualism | 0.044** | 0.024 | 0.044** | 0.021 |
| | (0.021) | (0.023) | (0.021) | (0.021) |
| Cum. cultural shock of emigration | -0.017* | -0.433** | -0.018* | -0.535*** |
| | (0.010) | (0.172) | (0.010) | (0.163) |
| Hist. controls: | Baseline | All | Baseline | All |
| Observations | 50 | 50 | 50 | 50 |
| Districts | 50 | 50 | 50 | 50 |
| Historical census years | 3 | 3 | 3 | 3 |
| R-squared | 0.98 | 0.99 | 0.98 | 0.99 |

Notes: Robustness check on table 3 using alternative sources of emigration data. All regressions include country and historical census year fixed effects. Historical controls are described in the main text. Mean and S.D. of the dependent variable is 0.917 and 0.059. Robust standard errors clustered at the district level are shown in the parentheses with the following significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

## Table A.25: Decennial Cultural Shock Dynamics, 1730-1910: Robustness

| | Baseline | GMM estimation | Balanced panels of 100 yrs | Sample from 1860 onwards | Add cultural control variables | Adj. For cohort emigration | Baseline | Baseline |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Dep. variable | | | | Cultural shock of emigration | | | Emigration | Rel. emi. individualism |
| Lagged dep. var. | -1.622*** | -2.766** | -1.631*** | -2.071*** | -1.642*** | -1.475*** | -2.383*** | 0.134 |
| | (0.328) | (1.131) | (0.329) | (0.480) | (0.337) | (0.386) | (0.325) | (0.085) |
| Lagged individualism | 0.038** | 0.370*** | 0.037 | -0.779* | 0.042** | 0.050** | 0.130*** | 0.025* |
| | (0.017) | (0.130) | (0.024) | (0.460) | (0.019) | (0.020) | (0.035) | (0.013) |
| Past emigration | 0.433*** | 0.328** | 0.437*** | 0.707*** | 0.435*** | 0.436*** | 1.445*** | -0.091* |
| | (0.040) | (0.140) | (0.043) | (0.125) | (0.040) | (0.035) | (0.116) | (0.052) |
| x lag individualism | -0.544*** | -0.834*** | -0.546*** | -1.039*** | -0.547*** | -0.553*** | -2.107*** | |
| | (0.068) | (0.194) | (0.074) | (0.211) | (0.068) | (0.061) | (0.204) | |
| Observations | 1,381 | 1,240 | 1,260 | 399 | 1,381 | 1,381 | 1,381 | 1,381 |
| Districts | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Census years | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| R-squared | 0.82 | | 0.81 | 0.61 | 0.82 | 0.82 | 0.88 | 0.96 |
| Mean of dep. var. | 0.005 | 0.006 | 0.006 | 0.019 | 0.005 | 0.005 | 0.016 | 0.016 |
| St.dev. of dep. var. | 0.012 | 0.012 | 0.012 | 0.015 | 0.012 | 0.012 | 0.029 | 0.029 |
| Hansen J-test (p-value) | | 0.691 | | | | | | |
| AR(2) test (p-value) | | 0.000 | | | | | | |

Notes: Robustness check of the results in Table 4. The unit of observation is a district and decade, observed in a historical population census. All specifications use the pooled historical population censuses and fixed effects for districts, decades, lags, and census years, and baseline controls (cohort number, gender, age, and number of first names). Past emigration is the cumulative sum of past rates of emigration. Robust standard errors, clustered at the district level, are shown in the parentheses with the following significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

## Table A.26: Long-Run Cultural Shock Dynamics

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Dep. var. | | | Cumulative cultural shock of emigration | | | | |
| Individualism | 0.158 | 0.153 | 0.247** | 0.280** | 0.228* | 0.163 | 27.930** |
| | (0.097) | (0.095) | (0.105) | (0.128) | (0.117) | (0.143) | (12.311) |
| Past emigration | | 0.131*** | 0.709*** | 0.602*** | 0.725*** | 0.643** | 1.206* |
| | | (0.024) | (0.160) | (0.216) | (0.258) | (0.262) | (0.627) |
| x hist. ind. | | | -0.852*** | -0.784** | -0.948** | -0.906** | -1.733* |
| | | | (0.221) | (0.307) | (0.356) | (0.378) | (0.872) |
| Controls: | | | | | | | |
| District - baseline | Y | Y | Y | Y | Y | Y | Y |
| Emigration - numbers and individualism | N | N | N | Y | Y | Y | Y |
| District - geography | N | N | N | N | Y | Y | Y |
| District - add. cultural traits | N | N | N | N | N | Y | Y |
| District - economic and demographic | N | N | N | N | N | N | Y |
| Observations | 140 | 140 | 140 | 140 | 140 | 140 | 140 |
| Districts | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| Historical census years | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| WVS/EVS census years | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| R-squared | 0.51 | 0.52 | 0.53 | 0.54 | 0.56 | 0.65 | 0.71 |
| Mean of dep. var. | 0.084 | 0.084 | 0.084 | 0.084 | 0.084 | 0.084 | 0.084 |
| St.dev. of dep. var. | 0.091 | 0.091 | 0.091 | 0.091 | 0.091 | 0.091 | 0.091 |

Notes: OLS estimates. The unit of observation is a district observed in a historical population census. The cumulative cultural shock of emigration from census year until 1920 is regressed on the level of individualism in the census year (the share of children that carry non-top-10 first names), the sum of past emigration rates, and baseline controls (child cohort size, population size, and average number of first names). All regressions include country and historical census year fixed effects. Robust standard errors, clustered at the district level, are shown in the parentheses with the following significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

# B    Appendix: Linking Migrants with Census Records

I link male international emigrants and internal out-migrants to the population census in which they were below the age of 15. Emigrants are identified from passenger lists and out-migrants as those living outside their birth district in any of the population censuses. In case a migrant can be linked to multiple censuses, I allocate them to the one in which they were youngest. Some out-migrants are already observed in their childhood household as they were below the age of 15 in the same census in which they were recorded as out-migrants. These are noted as exact matches. I match emigrants and remaining out-migrants to their childhood household using the following procedure:

1. A sample of potential links is constructed based on the phonetic spelling (truncated to the first four letters) of first and each last name, a two-year band around the birth year, and district of residence (emigrants) or birh (out-migrants). These are binding matching criteria.

2. Migrants that share the exact characteristics with one or more individuals in the census are identified as either exact matches or failed matches respectively. In terms of district of residence/birth similarity this step concerns the village (within district) level.

3. From here follows an iterative process where each matching criteria (similarity in first name, last name, birth year, and district of residence or birth) is weakened. The birth year is allowed to differ with one or two years. Non-perfect string similarity between first and last names is allowed (with Jaro-Winkler measures of 0.95, 0.9, 0.85., 0.8, and 0.75). Matching district of residence/birth is allowed to be at a higher level than the village: Sub-district and district.