Pushing and Pulling: Determinants of migration during Sweden's industrialisation

Björn Eriksson
Centre for Economic Demography, Department of Economic History, Lund University
bjorn.eriksson@ekh.lu.se

Siddartha Aradhya
Centre for Economic Demography, Department of Economic History, Lund University
siddartha.aradhya@ekh.lu.se

Finn Hedefalk
Centre for Economic Demography, Department of Economic History, Lund University
Department of Physical Geography and Ecosystem Science, Lund University
finn.hedefalk@nateko.lu.se

Paper prepared for the Social Science History Association Annual Conference, Chicago, November 17-20, 2016.

Introduction

The push and pull migration framework has long influenced the conventional understanding of migration patterns. Historical empirical work seeking to identify the forces that push individuals from particular locations and those that pull individuals to given destinations, have relied on one of two approaches. The first relies on the estimation of gravity models of aggregate flows between regions. The second approach is to model the individual migration decision using individual level data. Both approaches have its shortcomings: aggregate analysis neglects important heterogeneities that may misrepresent true push and pull forces. By combining dissimilar subgroups, aggregate migration data may actually hide the true impacts of important determinants of migration decision. Micro approaches offer an improvement to aggregate analyses by explicitly modelling the individual migration decision, but due to data limitations, are often limited to an incomplete set of origins and possible destinations or may encompass a non-representative sample.

This paper contributes to the literature by estimating parameters for push and pull factors by modelling the complete migration decision. Using data based on individuals linked between the complete Swedish censuses of 1880 and 1890 we consider both the push factors which determined whether an individual choose to leave their origin and the pulls factors which attracted migrants to specific destinations. The analysis includes both male and female migrants and takes into account the effect of individual and family characteristics. Moreover, we consider all possible origins and destinations for internal migrants.[1]

Our focus is on late nineteenth century Sweden, a country experiencing rapid economic growth and increasing rates of internal migration. The analysis is based on a cohort of men and women born between 1860 and 1870 that transitioned into adulthood during the height of Swedish industrialisation. These men and women are first observed as children and adolescents residing with their parents in 1880, and then again ten years later in 1890. Upon leaving their parental home, most moved a relatively short distance, often to a rural area not dissimilar to that in which they were born. By doing so, these

[1] Because of data limitations, emigration, which primarily took place to the US, has been omitted from the analysis.

migrants were following a well-established pattern dating back to pre-industrial Sweden. Many did however not follow in these well-trodden tracks, instead migrating further away. We focus on these medium and long range moves by considering inter county migration and the push and pull factors which determined the migration decision and the location choice.

Push and pull factors

Although migration constitutes complex decision making on part of the prospective migrant it may be conceptualized in a straight forward manner in which each possible move is associated with certain benefits and costs. If the net return from moving is positive, migration subsequently takes place to the chosen destination (Sjastaad 1962; Lee 1966). Several aspects relating to individual characteristics, the origin, potential destinations and intervening factors all form part of the decision process and affect the probability of migration and choice of destination.

Economic conditions

Expectations about the benefits of migrating to a certain locations is formed based on anticipated income gains and other non-pecuniary amenitie(s. Two distinct reasons for why migrants realise a monetary return may be considered. First, migrants may earn more as a result of wage levels being higher at their chosen destination than in the origin. Historically, this has been shown to be an important explanation for migration from mainly agricultural and rural areas with plentiful labour and low wages to urban and industrialized areas where labour was in demand and wages accordingly higher (Boyer 1997; Boyer and Hatton 1997). Secondly, differences in earnings reflect differences in occupational structure between the destination and origin of migrants. By moving to a location with better prospects for upward occupational mobility anticipated earnings increase as a result (Sjaastad 1962; Long 2005).[2]

[2] These expectations should be adjusted in order to account for differences in the probability of employment (Harris and Todaro 1970) we have yet to find an indicator of regional unemployment differences.

An associated literature have emphasised the importance of migration for eroding regional wage differences and explaining historical income convergence both between and within countries (Boyer and Hatton 1997; Taylor and Williamson 1997). For Sweden, recent work have shown that aggregate migration flows was an important factor in driving regional convergence in Sweden, in particular in the period leading up to 1910 (Enflo, Lundh and Prado 2014; Enflo and Roses 2015).

Intervening factors

A wealth of evidence show that distance has a strong negative effect on migration, with more remote locations being consistently less attractive destinations. In terms of costs, the distance between two locations is a proxy for both upfront monetary costs associated with a particular move, and the psychological cost implied by the separation from amenities in the origin such as friends and family (Sjaastad 1962; Schwartz 1973). Apart from affecting the cost side of the migration decision, distance also captures differences in the information available about a given location. With distance, the uncertainty about conditions in a location thus increases.

A number of regional and individual characteristics may serve to mediate the effect of distance. At the regional level, access to transportation and communication infrastructure such as roads, railways and postal services serves to lower the cost associated with distances between locations. At the individual level, networks of friends and family may serve to decrease uncertainty and lower the psychological cost of a move.

Individual characteristics

Apart from general conditions associated with a certain location, individual characteristics affects the decision of whether to migrate or not. Expectations, ability, benefits, costs and resources are all characteristics that vary between individuals and simultaneously determine the incidence of migration and the return thereof. Migration is as a result a highly endogenous process undertaken by a certain groups and individuals, each differently selected depending on individual characteristics and circumstances. If costs are important the expectation is selection of the most able, ambitious and

entrepreneurial part of the population who are better able to recoup costs in the form of substantial returns (Lee 1966). Similarly, costs may affect selection if cost is a negative function of ability, the able being, in Chiswick's (1999) words "more efficient in migration". Upfront migration costs may also serve as a more direct barrier by preventing the financially constrained from moving. Even when costs are fixed, as in the case of a train or boat ticket, migration is still relatively more expensive for the less able because fewer hours of work are required on the part of the more able to cover expenses associated with a move. Selection may also be negative if there are regional differences in terms of returns to skills which will result in opposing migrant streams of skilled and unskilled migrants drawn to location in which the returns to skills are consummate with individual ability (Roy 1951; Borjas 1987)

Migration and industrialisation

As the nineteenth century was drawing to an end, Ravenstein (1889:288) concluded that 'an increase in the means of locomotion and a development of manufactures and commerce have led to an increase of migration. In fact you need only seek out those provinces of a country within which migration is proceeding most actively, and you will either find yourself in the great centres of human industry, or in a part of the country whose resources have only recently become available. Migration means life and progress; a sedentary population stagnation'. The changes in migration patterns described by Ravenstein have been argued to constitute an essential component of the modernization process (Zelinsky 1971:222), conspicuous enough to mark a transition from a pre-industrial to a modern migration pattern (Parish 1973; Pryor 1975). The decline of traditional trades and the emergence of new occupations, faster and cheaper forms of transportation, unbalanced regional development and increasing population growth were some of the factors that served to facilitate this transition (Parish 1973).

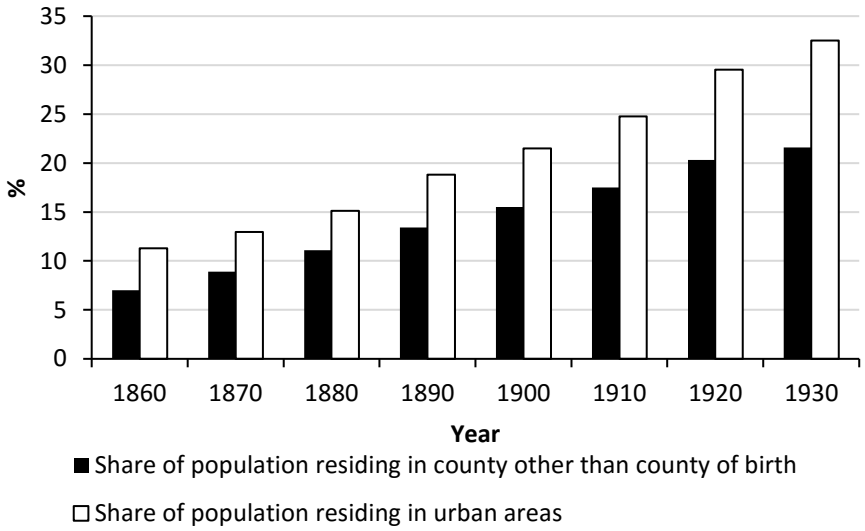
This shift does not imply that pre-industrial societies were geographically immobile. Agricultural societies were in fact characterised by high rates of geographic mobility. This is true for Sweden (Gaunt 1977:195, Dribe 2000:5-6, Dribe 2003), and a number of other European countries (Moch 1992; Jackson & Moch 1989), such as Germany

(Hochstadt 1983), England (Patten 1976, Clark 1979) and Scotland (Devine 1979). Migration was to a large extent driven by frequent short distance moves between rural areas by young people working as servants (Eriksson and Rogers 1978; Dribe & Lundh 2005). Migration was tied to the seasons and was the means by which the young earned their keep and acquired skills in agricultural work (Moch 1992: 61).

What sets modern geographic mobility apart from the pre-industrial agrarian case is thus not primarily the incidence of migration but rather its form. In contrast to pre-industrial Sweden, geographic mobility during the second half of the nineteenth century display a markedly different pattern (Brändström, Sundin & Tedebrand 2012). The shift in migration was underpinned by greater access to transportation. Real improvements to the infrastructure began with the canal construction boom between 1780 and 1830 which connected the hinterland of central Sweden with the Baltic and North Seas. This was followed by the opening of stagecoach lines connecting Stockholm with Gothenburg and Scania in the 1830s. The building of canals and stagecoach lines primarily served to lower the cost of transporting bulky goods and improving the flow of information through the post (Schön 2010:101-102). The real breakthrough in terms of affordable passenger transport followed the construction of railways which commenced in 1855. The arrival and expansion of the railway markedly reduced both costs and transit times for passengers. As a result the pace of urbanization began to accelerate as migrants started flocking to towns connected to the railway (Berger and Enflo 2013).

Migration was increasingly taking place over longer distances in response to relatively higher wages in industrialising regions (The Institute for Social Sciences 1941:42; Jörberg 1972: 348). As can be garnered from Figure 1, only 7% of the population in 1860 resided in a county different from their county of birth; by 1900 however, the share of county migrants had more than doubled to 15.5% and continued to rise over the following decades (also see Thomas 1941:28). This resulted in most towns growing at a vigorous pace, but the three largest cities, Stockholm, Gothenburg and Malmö grew particularly fast (Statistiska Centralbyrån 1969). The inflow of migrants from the countryside resulted in a doubling of the urban share of the population between 1860 and 1900 (see Figure 1).

Figure 1. Migration and urbanisation, 1860-1930.



Source: Sveriges Officiella Statistik 1930: 84*; Statistics Sweden 1999: 42.

This transition was not unique to Sweden, but was reflected in many other European countries (Baines 1985; Moch 1992; van der Woude 1992). With the advent of industrialisation the population of Europe had entered a new phase of geographic mobility. As a result of falling transportation costs, a declining agricultural sector and new opportunities outside of farming, the nature of migration was changing. Instead of making a living in agriculture, increasing numbers of people were leaving the countryside in favour of cities.

As long as certain administrative requirements were met there were no legal obstacles to internal migration in Sweden during the nineteenth century. Before moving, a prospective migrant was required to notify the ministers of both the home and destination parish. Permission to settle in a new parish was given as long as it was not suspected that the migrant would have difficulty supporting him- or herself. Refusal of permission to move was exceedingly rare, with less than 1 per cent of applications denied (Eriksson & Rogers 1978:180-181). One institutional barrier to migration did however exist, the Servants Act, which in particular hampered migration for agricultural workers. The act mandated that yearly employment contracts for farmhands and maids must begin on the 1st of November and run until the 24th of October the following year.

This resulted in little down time between employment contracts, which made it difficult for farm workers to find employment anywhere beyond the vicinity of their last place of work (Lundh 1999:61; Lundh 2003). Early and mid-nineteenth century migration was also limited by the late arrival of the railway on which construction did not commence until 1855.

Data and descriptive statistics

This paper exploits individual panel data and contextual level data in order to model the migration decision in a comprehensive manner. The individual level data comes from the complete Swedish censuses of 1880 and 1890. The contextual level data, which is a county level, has been sourced historical official statistics and constructed regional wage and GDP series.

Linked census data

The censuses are the most comprehensive source of individual level data for Sweden around the turn of the century. The Swedish censuses differ from the U.S. and British censuses by not being the product of a census taking done by enumerators actually visiting and counting the populace. Instead, with one exception, that of the city of Stockholm, the Swedish censuses were the result of a compilation of excerpts from continuous parish registers which were kept by the Swedish Lutheran church and maintained by the parish priest. For Stockholm, the source of the census were excerpts from the Roteman register, an administrative register supervised by Mantalsnämnden (The population and tax registration board), which replaced the Church registers in 1878 in order to cope with the rapidly growing, and increasingly mobile, population of Stockholm at the end of the nineteenth century (Geschwind & Fogelvik 2000:207-208).

Because the Swedish census is no more than an excerpt from a continuous and consistent source rather than a recreation of a population register as in the case of the U.S. and British censuses, the quality of the raw data is comparatively better. Any errors resulting from the misreporting by the enumerated or recording mistakes by enumerators may thus be largely discounted. Moreover, because Swedes were entered

into the parish books at the time of christening and not removed until time of death or emigration, the under-enumeration of the population as whole and specific groups is less of a problem than what is normally expected from historical censuses.

The analysis relies on a new panel sample which has been created by linking individuals between the 1890 and 1880 Swedish complete count censuses. The linking process relies on exact comparisons of sex, birth place and birth year, and probabilistic matching of names for identifying and linking individuals between the censuses. Importantly, and uniquely, women appear with their maiden name, even after marriage, in the Swedish censuses. This enabled women to be linked to nearly the same extent as men between the two censuses. For a thorough discussion of the linking process see Eriksson (2015; 2016)

From the linked sample a sub-sample of men and women that were born between 1860 and 1870 were selected. We further restrict the analytical sample to those that resided with their father in 1880 in order to collect information about social status. One significant group has by necessity been excluded from the analysis: emigrants. The reason is that in order to be linked between the censuses, an individual had to reside in Sweden in both 1880 and 1890. Anyone emigrating out of Sweden between the two time points was thus lost in the linking process. After restricting our sample according to the above criteria, we are left with 293129 individuals evenly distributed by sex.

As a result of the linking we have information about the location of each individual in 1880 and ten years later in 1890. Although we have information on the parish level, we use Sweden's 24 counties as the geographic level of division in order to make data collection and modelling feasible. We include a number of individual characteristics associated with the probability of migration. We use HISCO coded occupations and the HISCLASSS scheme to classify the social status of an individual's father into one of four categories; white collar (HISCLASS 1-5), farmer (HISCLASS 8 with a HISCO code that corresponds to the occupation of farmer), skilled (HISCLASS 6-8) and low skilled (HISCLASS 9-12) (see van Leeuwen et al. 2002 and van Leeuwen et al. 2011). To account for the effect of previous migration experiences we coded any individual residing in a county in 1880 which was different from his or her county of birth as a

previous migrant. The migrant status of an individual's mother or father was coded in the same manner.

Table 1. Descriptive statistics of individual sample

	All	Non-migrants	Migrants
Sex	0.51	0.50	0.52
Age	13.88	13.85	14.15
	(2.77)	(2.77)	(2.70)
Urban resident	0.10	0.09	0.17
Father social status			
White collar	0.09	0.08	0.17
Farmer	0.66	0.68	0.42
Skilled	0.20	0.19	0.33
Low skilled	0.05	0.05	0.08
Migration history			
Previous migrant	0.05	0.04	0.15
Mother migrant	0.18	0.16	0.33
Father migrant	0.16	0.14	0.31
Migrant 1880-90	0.10		
No of observation	293129	265134	27995

Note: Reported statistic is mean. Standard deviations for continuous variables within parentheses.

Sources: see text.

Table 1 presents the descriptive statistics for the complete sample and by migrant status. All variables refer to characteristics observed in 1880. Three things are apparent from the descriptive statistics. The first is the low incidence of migration for the sons and daughters of farmers. The second is the higher mobility of urban residents. The final point to note is the higher mobility of individuals with either a personal, or family history of migration.

County level data

We use a number of sources to collect data on county specific variables. All data refers to the year 1880. Although it is likely that conditions in both origin and destination would have changed between the time of observation in 1880 and the year in which an individual choose to migrate, we are limited in this regard because we do not know the exact time at which migration took place. In addition, even if we had information about

the exact time point of migration, most of the county level characteristics lack annual data.

To approximate the economic conditions in an individual's origin and possible destinations we use data on the real wage level and GDP per capita for each county. The real wage data comes from Dribe (2008) and is based on nominal day wages in agriculture (except for Stockholm county, where the wages are those of unskilled labourers) which have been deflated by the county specific prices of a food basket. Wages are thus to some degree adjusted by regional differences in cost of living. The measure of GDP per capita for each county comes from Enflo, Henning and Schön (2014). Although both variables captures the economic conditions in a specific county, each measure also reflect different aspects of economic conditions. Since the real wage index is based on the wages of day labourers, it is likely to primarily capture differences in living conditions of low skilled labour. In contrast, GDP per capita is a broader measure, reflecting differences between counties in terms of economic development. To ease the interpretation and comparison of the estimates of real wages and GDP we use standardised variables (mean = 0, s.d.= 1) when estimating our migration models.

The degree of urbanisation of each county has been calculated using the complete census of 1880 and is reported as the share of the population in each county that resided in an urban area. The area of each county comes from BISOS I (1881) and is measure in Swedish square miles (approximately equal to 106.89 km^2). Information about access to transportation has been collected from official statistics. The data on railways can be found in BISOS L (1881) and the data on roads in BISOS H (1884). Both are measured as kilometres of road/railway per Swedish square mile. The data on roads only include public roads, while the railroad data includes both public and privately owned railways. Two variables, distance and the historical migration stream, are constructed based on specific information about the origin and a possible destination and are thus different for each unique combination of county of origin and potential destination. Distance has been calculated as the number of kilometres between the centroid of the origin of each county. Similarly, the migration stream variable has been calculated for each county by

Table 2. County level variables

County code	County	Real wage index	GDP per capita	Urban population share	Roads (km per Swedish mile^2)	Railways (km per Swedish mile^2)	Area (Swedish mile^2)
1	Stockholms län	52.7	607.4	56.2	33.7	2.4	68.2
3	Uppsala län	55.8	347.6	16.4	36.2	5.9	46.5
4	Södermanlands län	52.9	335.2	11.4	38.2	4.1	59.7
5	Östergötlands län	44.6	337.8	15.4	37.9	2.9	96.9
6	Jönköpings län	55.6	315.0	10.4	52.8	3.5	100.7
7	Kronobergs län	65.4	258.6	2.9	52.7	3.4	87.1
8	Kalmar län	49.8	282.0	10.4	30.9	3.2	100.6
9	Gotlands län	55.5	322.1	12.7	49.3	2.1	27.3
10	Blekinge län	45.3	299.5	19.1	39.5	2.7	26.4
11	Kristianstads län	68.7	279.4	5.7	62.2	4.5	56.8
12	Malmöhus län	47.1	403.3	24.1	70.9	8.0	41.9
13	Hallands län	58.4	276.4	11.2	53.7	1.8	43.1
14	Göteborg och Bohus län	38.2	381.1	33.7	32.7	0.8	44.3
15	Älvsborgs län	47.5	240.4	5.7	37.8	4.2	112.2
16	Skaraborgs län	51.9	295.1	6.9	46.6	5.9	74.4
17	Värmlands län	36.0	266.8	5.9	24.1	3.1	166.5
18	Örebro län	46.4	283.4	9.1	28.2	7.4	79.4
19	Västmanlands län	44.7	365.8	13.9	32.0	5.6	59.5
20	Kopparbergs län	48.5	315.1	4.9	10.3	1.3	258.9
21	Gävleborgs län	61.1	475.1	17.4	14.8	2.2	168.2
22	Västernorrlands län	71.3	392.9	8.6	10.4	0.7	215.8
23	Jämtlands län	85.4	415.8	3.4	3.5	0.5	444.4
24	Västerbottens län	78.3	313.7	3.5	5.2	0.0	497.4
25	Norrbottens län	71.4	407.4	7.2	0.2	0.0	913.9

Sources: see text

identifying all individual living outside their county of birth according to the 1880 census, and then calculating the share residing in each possible destination. The real wage index

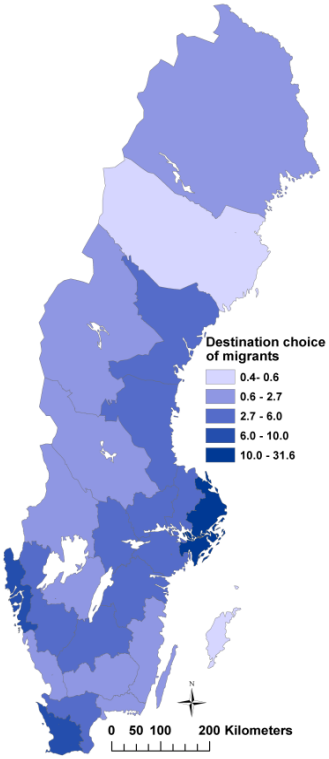
The county level characteristics are presented in table 2. In terms of economic conditions there is considerable variation between counties. The real wage index is more than twice as high in the best paying region compared to the lowest paying. GDP per capita displays a similar pattern in terms of the variation in regional development. Wages and GDP are, as may be expected, positively correlated (0.18) although not very strongly. Differences in road and railway coverage are also pronounced. The northern counties in particular lacked transport infrastructure compared to central and Southern Sweden.

The differences in real wages and GDP are reflected in the incidence of migration at the county level and in the destination choice of migrants. Figure 2 displays maps of the regional variation in real wages and GDP together with rate of outmigration and the destination choice of migrants. The rate of outmigration for each county has been calculated as the share that choose to leave the county between 1880 and 1890 using the linked census sample defined above. The destination choice has been calculated as the share of all county migrants between 1880 and 1890 that choose to settle in a specific county. The urbanised counties of Stockholm, Malmöhus and Göteborg was clearly the favoured destination of the migrants in our sample. The migrants to these counties were primarily drawn from neighbouring counties and the central counties in southern Sweden. Outmigration from the northern counties was low, a pattern which is consistent with the relatively higher wages and high level of GDP in the north. Although the maps provides clear descriptive evidence between economic conditions and migration, these correlations tell us little about their magnitude and importance relative to other factors which affected individuals migration decisions. This is the question to which we turn next.

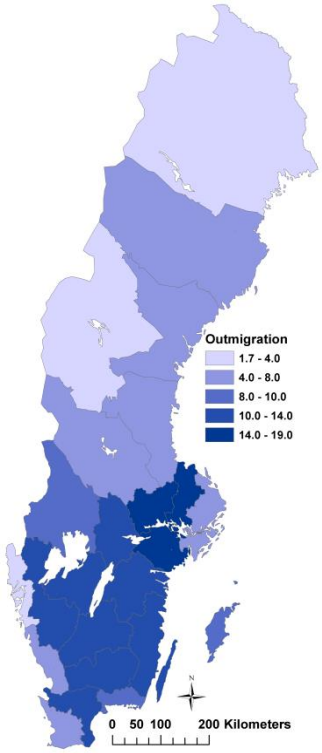
Modelling the migration decision

The migration choice maybe be conceptualized as a decision tree with two levels (see figure 1). The top level entails the choice of migrating or remaining in the place of origin. The second choice, which is conditional on migration, concerns the choice of which destination to move to.

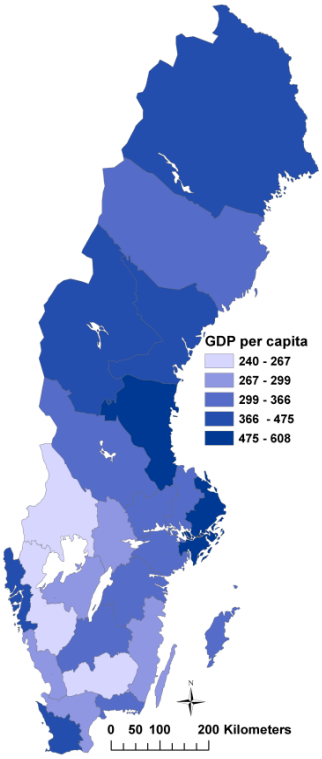
Figure 2. Migration and economic conditions



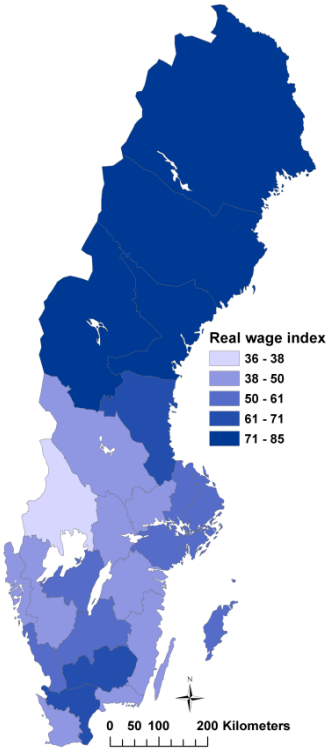
(a) Destination choice of migrants, 1880-90 (%)



(b) Outmigration, 1880-90 (%)

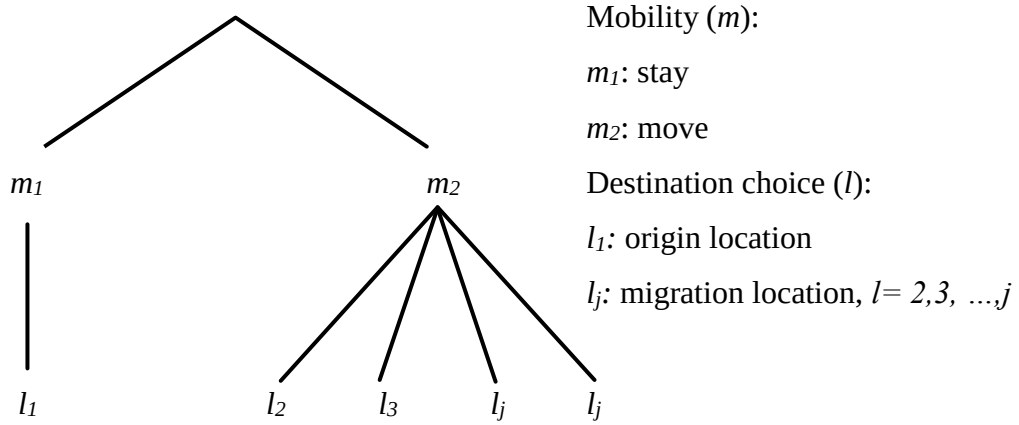


(d) GDP per capita, 1880



(d) Real wage index, 1880

Figure 3. Two tier nested structure of migration and destination choice



To account for the fact that destination choice is nested within the migration branch of the decision tree, thus making destination choice conditional on migration, we employ the utility maximizing nested logit model developed by McFadden (1978; 1981, 1984). The nested logit is a less restrictive alternative to the multinomial logit model since the independence of irrelevant alternatives (IIA) assumption is relaxed. The IIA assumption requires the response elasticities of choices to be equal. In our case, the use of a multinomial logit model and the associated IIA assumption would mean that the introduction of a new destination would require an equal decrease the predicted probability of choosing a specific destination location and remaining in the origin. By nesting the destination choice within the migration decision, the IIA assumption only needs to hold between the choice to migrate or not or between choosing a specific location, and not across the two choice sets. Moreover, the model allows us to simultaneously assess both the push and pull factors which affect migration. By using this approach, push factors are evaluated in a binomial migration choice model and pull factors are assessed in a multinomial destination choice model. Although the choice is conceptualized as sequential in the decision tree, the nested logit does not impose a sequential decision process. Or in other words, the model does not imply the unrealistic assumption that an individual first chooses whether to migrate or not, and only thereafter considers which destination to choose. The two equations are estimated simultaneously using maximum likelihood.

Results

The results from the regression analysis are to be interpreted in the two levels as specified in the decision tree. The top panel in Table 3 presents the results corresponding to the first level decision, the push factors that are associated with the decision to relocate. In this case, individual characteristics, as well as the conditions in the origin county influence the decision to migrate from or stay in the origin county. The lower panel in table 3 displays the impact of the pull factors, the characteristics of the destination counties that influenced the destination choice. In order to understand the economic mechanisms that drive the migration decision, we estimate 3 model specifications. Model one only contains a measure of wages in the origin and destination counties in order to understand whether migration decisions are driven by labor market conditions. Model two, on the other hand, contains GDP per capita of the origin and destination counties, to understand the relationship between the broader macro-economic conditions and migration decisions. Finally, Model 3 contains wages and GDP per capita. The final model specification is used to test whether different economic factors exert a push or pull force on migration decisions. In general, the results are entirely consistent with theoretical expectations.

Push factors:

The top panel in table 3 present the results of the individual and contextual push factors, the first level decision. The results indicate that females and individuals residing in urban origins experience higher odds of choosing to migrate across each of the models. Additionally, a one year increase in age is associated with roughly a 5.5 percentage point increase in the odds of choosing to migrate. The magnitude of the coefficient for age is large, because the sample is restricted to individuals that are 10 to 20 years of age in 1880, with the higher age groups being the most likely to migrate. Father's social status also seems to exert an important push force. Children of farmers display a roughly 58 percentage point lower odds of choosing to relocate, while those of white collar workers display a 13 percentage point higher odds compared to those with skilled fathers.

The most important individual level push factor, however, is having previously migrated or having a parent that previously migrated. Individuals who were residing in a county in 1880 that was different than their birth location experienced approximately a 74 percentage point higher odds of migrating again.

Table 3. Odds ratio estimates from nested logit models of migration and destination choice

	1	2	3
Panel A: Migration Choice			
Sex	1.056 ***	1.055 ***	1.055 ***
	(0.014)	(0.014)	(0.014)
Age	1.055 ***	1.055 ***	1.055 ***
	(0.002)	(0.002)	(0.002)
Urban resident	1.222 ***	1.224 ***	1.224 ***
	(0.025)	(0.025)	(0.025)
Father social status			
White collar	1.136 ***	1.131 ***	1.133 ***
	(0.024)	(0.024)	(0.024)
Farmer	0.420 ***	0.417 ***	0.418 ***
	(0.007)	(0.007)	(0.007)
Skilled	ref.	ref.	ref.
Low skilled	0.930 ***	0.933 ***	0.934 ***
	(0.024)	(0.024)	(0.024)
Migration history			
Previous migrant	1.741 ***	1.744 ***	1.743 ***
	(0.041)	(0.041)	(0.041)
Mother migrant	1.514 ***	1.513 ***	1.514 ***
	(0.025)	(0.025)	(0.025)
Father migrant	1.554 ***	1.555 ***	1.556 ***
	(0.027)	(0.027)	(0.027)
County level variables			
Wages	0.968 ***		0.964 ***
	(0.009)		(0.011)
GDP per capita		0.940 ***	0.983
		(0.017)	(0.022)
Urban share	0.982 ***	0.989 ***	0.985 ***
	(0.001)	(0.002)	(0.002)
log area	0.957 **	1.011	1.004
	(0.019)	(0.021)	(0.021)
Road network	1.004 ***	1.003 ***	1.003 ***
	(0.001)	(0.001)	(0.001)
Railway network	1.004	1.022 ***	1.016 ***
	(0.004)	(0.004)	(0.005)
Constant	7.235 ***	3.769 ***	4.171 ***
	(0.955)	(0.550)	(0.626)
Panel B: Destination choice			
County level variables			
Wages	1.038 ***		0.996
	(0.003)		(0.003)
GDP per capita		1.100 ***	1.103 ***
		(0.005)	(0.005)
Urban share	1.009 ***	1.002 ***	1.002 ***
	(0.000)	(0.000)	(0.000)
Distance	0.999 ***	0.999 ***	0.999 ***
	(0.000)	(0.000)	(0.000)
Historical migration stream	1.059 ***	1.056 ***	1.056 ***
	(0.001)	(0.001)	(0.001)
No. of observations	293129	293129	293129
No of iterations	41	38	42
χ^2	16539.099	16541.828	16525.295

Note: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Similarly, having a parent that previously migrated increased the odds of migrating for that individual by approximately 50 percentage points. The mechanisms for these associations are potentially two-fold. First, individuals that have previously migrated may have existing networks outside of their county of residence. Second, the costs associated with migration are reduced with every subsequent act of migration as an individual develops migration specific human capital.

Next we turn to the county-level push factors. The results indicate that individuals are less likely to relocate from counties with better economic prospects. Specifically, a one standard deviation increase in the real wage index in the origin county is associated with approximately a 3.5 percentage point decrease in the odds of migrating in models 1 and 3. In model 2, we replace the real wage index with GDP per capita and find that a one standard deviation increase is associated with a 6 percentage point decrease in the odds of resettling. In model 3, however, GDP per capita does not appear to play a role in the first level decision. Rather, the real wage index seems to be an influential factor. This indicates that wages are a more important push factor than the general economic condition of the origin county.

The results further suggest that individuals are less likely to move away from urban origin locations, and are more likely to migrate when there is extensive transportation infrastructure in their origin. The magnitude of these effects are rather small, but remain robust across model specifications. In terms of push factors, it seems as though the most important drivers can be found at the individual level, with previous migration experience being particularly important.

Pull factors:

The estimated pull factors are found in the bottom panel of table 3. The characteristics of each destination county in the year 1880 are used to estimate the parameters of the pull factors. The covariates are measured in the year 1880 so as to ensure that they represent the conditions prior to the year of migration, and are not themselves an outcome of migration flows between 1880 and 1890. The results again corroborate a priori expectations.

The estimates suggest that a distance of 100 kilometers between the origin and a given destination decreases the odds of selecting that destination by one percentage point. This is expected as the emotional and financial cost of migration and distance are positively

correlated. Additionally, the migration flows tend to be pulled to more urbanized destinations, with urban share of a given county slightly increases the likelihood of selecting that destination.

The dynamics of network or chain migration seem to play an important role in determining the destination location. Although we are unable to identify the channel through which this relationship exists, the results indicate that a 10 percent increase in the origin migrant stock leads to an approximately 50 percentage point increase in the odds of selecting that given location. Network theories of migration postulate that existing migrant networks decrease the cost of migration by providing access to housing or labor market opportunities.

The final pull factor tested in our models are the economic conditions in each given destination. Similar to the push factors, we first tested a model that included only an unskilled wage index. According to model 1, a one standard deviation increase in the index was associated with a 3.8 percentage point increase in the odds of selecting a given destination. This indicates that the wage conditions in a destination are an important pull factor. In model 2, we replaced the index with a GDP per capita measure for each destination location. In this model, the results again show that a one standard deviation increase in GDP per capita increased the odds of selecting a given destination by 10 percentage points. This model suggests that broader macroeconomic conditions may be an important pull factor. The final model included both, the wage index as well as GDP per capita. This model specification allows us to test whether GDP per capita or wages were both important factors. The results here suggest that wages are not an important pull factor, rather general economic conditions drove destination selection. The magnitude of the estimate for GDP per capita remains unchanged from model 2, while the estimate for the wage index are no longer statistically significant. This finding is particularly important as it suggests that the economic mechanisms that push individuals to migrate differ from the factors that pull migrants to a give destination. Based on these results, we argue that wages are an important push factor while general economic prospects are an important pull factor. This may be due to the fact that migrants select a particular destination based on its broader societal prospects. Higher GDP per capita may indicate a better standard of living, or more opportunities for the individual and one's family.

Concluding remarks

The results in this paper constitutes a first attempt at modelling push and pull factors using a micro level population based sample which considers all possible origins and destinations. The results clearly show that economic push and pull factors drove migration decisions during the period. Although macro level push forces seemed to play a role, individual level factors dominated. Specifically, having a previous migration experience or having a parent that migrated was the strongest push effect. Likewise, wages rather than GDP per capita in the origin was the important economic determinant, indicating migration decisions are made due to individual circumstance. Wages, however, were not an important pull factor. GDP per capita was the most important determinant of destination location. This finding is consistent with the idea that migration decisions are made with imperfect information. Individuals may be pushed by wages while they are pulled by GDP per capita simply because it is easier to judge the attractiveness of a given destination by its macro conditions. Individuals are unable to know what their specific outcomes may be in a given destination, but can better judge the positive externalities of locating in a more developed destinations.

The implications of these are important to the literature as they are consistent with existing theories of migration. The validity of theoretical push and pull factors are uncompromised when comprehensively modelling the decision process. More work remains to be done in order to identify the mechanisms through which some of these factors are operating. We will additionally account for regional health pull factors in our subsequent analysis as this should help further explain the migration decision.

It is important to note that this paper does not explicitly test individual level pull factors, which theoretically must exist, but we intend to address this at a later time. One important individual level pull factor is the existence of networks. Although our results indicate that networks are an important pull factor, it is unclear as to whether this relationship is dominated purely by migration flows be concentrated in the larger cities. We intend to disentangle this effect further by capturing networks at a more detailed level in the future.

References

Sjastaad, L.A. (1962). The Costs and Returns of Human Migration. *Journal of Political Economy*, 70: 80-93.

Lee, E.S. (1966) A Theory of Migration, *Demography*, 3: 47-57.

Long, J. (2005). Rural-Urban Migration and Socioeconomic Mobility in Victorian Britain. *The Journal of Economic History*, 65: 1-35.

Harris, J.R. and Todaro, M.P. (1970). Migration, Unemployment and Development: A Two-Sector Analysis, *American Economic Review*, 60: 126–142

Boyer, G.R. (1997). Labour migration in southern and eastern England, *European Review of Economic History*, 1: 191-215

Boyer, G.R. and Hatton, T.J. (1997). Migration and labour market integration late nineteenth-century England and Wales, *Economic History Review*, 50: 697-734.

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

Enflo, K. and Roses, J.R. (2015). Coping with regional inequality in Sweden: structural change, migrations, and policy, 1860-2000. The Economic History Review, 1: 191-217.

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

Schwartz, A. (1973). Interpreting the effect of distance on migration, *Journal of Political Economy*, 81: 1153-1169.

Chiswick, B. (1999). Are immigrants favorably self-selected? *The American Economic Review*, 89: 181-185.

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

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3: 135-146.

Ravenstein, E.G. (1889). The Laws of Migration: Second Paper. *Journal of the Royal Statistical Society*, 52: 241-305.

Zelinsky, W. (1971). The Hypothesis of the Mobility Transition. *Geographical Review*, 61: 219-49.

Parish, W.L. (1973). Internal Migration and Modernization: The European Case. *Economic Development and Cultural Change*, 21: 591-609.

Pryor, R.J. (1975). Migration and the process of modernization. pp. 23-38, in Kosinski, L.A. & Prothero, R.M. *People on the Move*. London: Methuen & Co.

Gaunt, D. (1977). Pre-industrial Economy and Population Structure. *Scandinavian Journal of History*, 2: 183-210.

Dribe, M. (2000). *Leaving home in a peasant society. Economic fluctuations, household dynamics and youth migration in southern Sweden, 1829-1866.* Södertälje: Almqvist & Wiksell International.

Dribe, M. (2003). Migration of rural families in 19th century southern Sweden. A longitudinal analysis of local migration patterns. *The History of the Family*, 8: 247-65.

Moch, L.P. (1992). *Moving Europeans: Migration in Western Europe since 1650*. Bloomington: Indiana University Press.

Jackson, J.H. and Moch, L.P. (1989). Migration and the Social History of Modern Europe. *Historical Methods*, 22: 27-36.

Clark, P. (1979). Migration in England during the Late Seventeenth and Early Eighteenth Centuries. *Past & Present*, 83: 57-90.

Devine, T.M. (1979). Temporary Migration and the Scottish Highlands in the Nineteenth Century. *The Economic History Review*. 32: 344-59.

Hochstadt, S. (1983). Migration in Preindustrial Germany. *Central European History*, 16: 195-224.

Patten, J. (1976). Patterns of migration and movement of labour of three preindustrial East Anglian towns. *Journal of Historical Geography*. 2: 111-29.

Eriksson, I. and Rogers, J. (1978). *Rural Labor and Population Change*. Stockholm: Almqvist & Wiksell.

Dribe, M. and Lundh, C. (2005). People on the move: determinants of servant Migration in nineteenth-century Sweden. *Continuity and Change*, 20: 53-91.

Brändström, A., Sundin, J. and Tedebrand, L-G. (2012). Two Cities: Urban Migration and Settlement in Nineteenth-century Sweden. *The History of the Family*, 5: 415-29.

Berger, T. and Enflo, K. (2013). Locomotives of local growth: the short- and long-term impact of railroads in Sweden, *EHES Working Papers in Economic History*, 42.

Schön, L. (2010). *Sweden's Road to Modernity: An Economic History*, Stockholm: SNS.

Jörberg, L. (1972). *A History of Prices in Sweden 1732-1914.* Vol. 2, Lund: CWK Gleerup.

The Institute for Social Sciences, Stockholm University, (1941). *Population Movements and Industrialization*. London: P.S. King & Son.

Thomas, D.S. (1941). *Social and Economic Aspects of the Swedish Population Movement, 1750-1933.* New York: The Macmillan Company.

Statistiska Centralbyrån (1969). *Historisk statistik för Sverige*, 1, Stockholm.

Baines, D. (1985). *Migration in a Mature Economy: Emigration and Internal Migration in England and Wales 1861-1900*. Cambridge: University Press.

van der Woude, A., Hayami, A. and Vries, J. (1990). *Urbanization in History.* Oxford: Clarendon Press.

Lundh, C. (1999). Servant Migration in Sweden in the Early Nineteenth Century. *Journal of Family History*, 24: 53-74.

Lundh, C. (2003). Life Cycle Servants in Nineteenth Century Sweden – Norms and Practice. *Lund Papers in Economic History*, No. 84.

Geschwind, A. and Fogelvik, S. (2000). The Stockholm Historical Database. In P.K. Hall, R. McCaa, G. Thorvaldsen (eds.), *Handbook of International Historical Microdata for Population Research*, pp. 207 – 231. Minneapolis: Minnesota Population Center.

Eriksson, B. (2015). Dynamic Decades: A micro perspective on late nineteenth century Sweden, Lund: Media-Tryck.

Eriksson, B. (2016) *False positives and faulty estimates: Linked census data and bias to estimates of social mobility*, unpublished manuscript.

van Leeuwen, M. H. D. and Maas, I. (2011). *HISCLASS. A historical international social class scheme.* Leuven: Leuven University Press.

van Leeuwen, M. H. D., Maas, I., and Miles, A. (2002). *HISCO. Historical international standard classification of occupations*. Leuven: Leuven University Press.

Dribe, M. (2008). Demand and supply factors in the fertility transition: A county level analysis of age-specific marital fertility in Sweden 1880-1930. *European Review of Economic History*, 13: 65-94.

Enflo K, Henning M. and Schön L. (2014). Swedish regional GDP 1855-2000: Estimations and general trends in the Swedish regional system. *Research in Economic History*, 30: 47-89.

Bidrag till Sveriges Officiella Statistik. H) Kungl. Maj:ts befallnings-havandes fem-årsberättelser 1976-1880. (1884)

Bidrag till Sveriges Officiella Statistik. I) Telegrafväsendet 1880. (1881)

Bidrag till Sveriges Officiella Statistik. L) Statens Järnvägstrafik 1880. (1881)

McFadden, D. (1978). Modelling the choice of residential location, (pp.75-96) in A. Karlquist, L. Lundquist, F. Snickars, and J.W. Weibull (Eds.), *Spatial interaction theory and planning models.* Amsterdam: North Holland.

McFadden, D. (1981). Econometric models of probabilistic choice, (pp. 198-272) in C.F. Manski and D, McFAdden (Eds.), *Structural analysis of discrete data with econometric applications.* Cambridge: MIT Press.