

Pushing and Pulling: A population based approach to analysing the historical determinants of internal migration

Siddartha Aradhya, Björn Eriksson, Finn Hedefalk

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

Introduction

The push and pull migration framework has long influenced the conventional understanding of migration patterns. Historical empirical works seeking to identify the forces that push and pull individuals to and from particular locations 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

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 amenities. 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).

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

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

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 commensurate with individual ability (Roy 1951; Borjas 1987)

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.

Data and descriptive statistics

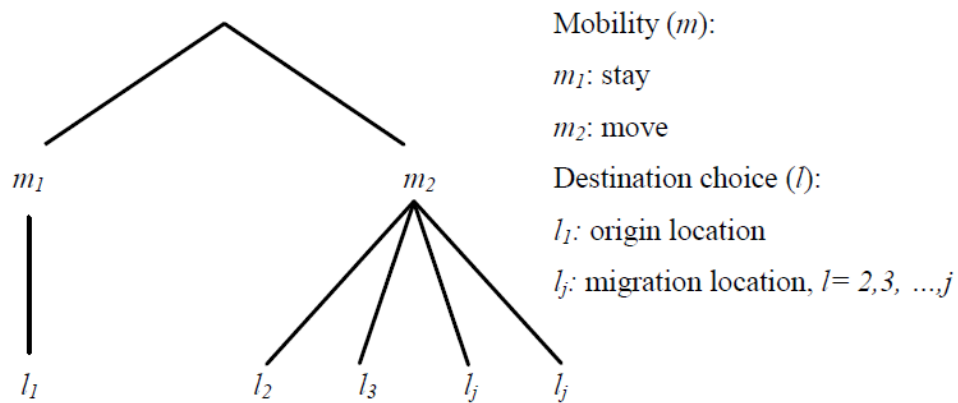
This paper exploits individual panel data and contextual level data in order to model the migration decision in a comprehensive manner. The individual level data comes from the complete Swedish censuses of 1880 and 1890. The contextual level data, which is at the county level, comes from historical official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

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. 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 293,129 individuals evenly distributed by sex.

Modelling the migration decision

The migration choice maybe 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 1. 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. 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.

Results

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 exerted the strongest push force. Likewise, wages rather than GDP per capita in the origin were the important economic determinant, indicating migration decisions are made due to individual circumstance. Wages, however, were not an important pull factor. GDP per capita, rather, 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 associated with locating in a more developed destination.

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 in our subsequent analyses. 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.