Universitat Autònoma de Barcelona

Faculty of Economics and Business

Bachelor's Degree in Economics

AGENT-BASED SIMULATIONS ON CATALAN INTERPROVINCIAL MIGRATIONS

Authored by

ALEEM K. SAEED

Director: Xavier Vilà

Dept. of Economics and Economic History,
Area of Fundamentals of Economic Analysis. UAB

Barcelona, May 30, 2022

Acknowledgments

Special thanks to my family, Xavier Vilà, Javier Fernandez-Blanco, Anna Klabunde, and Barend Ruesink Bueno for helping, guiding, and inspiring me throughout this process.

Abstract

This work contemplates the phenomenon of migration related to economic reasons. Firstly, the foundations of a mathematical model are laid down, which have key departure points that mold the approach for doing the rest of the thesis. Secondly, a statistical model on the probability of Catalan interprovincial migrations is displayed. Finally, a NetLogo model is used to simulate the Catalan interprovincial migrations, more specifically, return migrations throughout the period 2008-2020.

Sinopsis

Este trabajo contempla el fenómeno de la migración relacionado con motivos económicos. En primer lugar, se establecen las bases de un modelo matemático, que tiene puntos de partida claves que moldearon el enfoque para hacer el resto de la tesis. En segundo lugar, se muestra un modelo estadístico sobre la probabilidad de las migraciones interprovinciales catalanas. Finalmente, se utiliza un modelo hecho con NetLogo para simular las migraciones interprovinciales catalanas, más concretamente, las migraciones de retorno a lo largo del periodo 2008-2020.

Contents

1. Agent-Based Simulations on Catalan Interprovincial Migrations

In general terms, in human migration, a migrant is an individual either going through the process of moving, has moved within a State, or across an international border away from their earlier habitual place of residence. This definition of migration holds regardless of the migrant's legal status, whether they are moving voluntarily or not, the causes for their movement, and the length of their stay (Migration, 2022).

In specific terms, the diverse types of human migrations include:

- External migration: migrating to a different state, country, or continent.
- Internal migration: migration within a state, country, or continent.
- Emigration: leaving the place of origin to set up in another one.
- Immigration: an outsider settling down in destiny is an immigrant.
- Return migration: moving back from destiny to origin.
- Seasonal migration: moving with the seasons, for seasonal labor or due to climate conditions (Human Migration, 2022).

This work considers and substantiates certain factors that influence human decision-making when it comes to moving. This topic involves assumptions related to demography, migration, and policies. These assumptions are initially portrayed in the thesis in a purely indicative mathematical model that lays the thesis foundation in a compressed manner at a city level. Then, these assumptions are studied empirically. The empirical findings are used to solidify a set of hypotheses of an Agent-Based Model and can finally be applied in the NetLogo Circular Migration Model (NCMM). Given the inputs, the NetLogo model reflects the corresponding output, which naturally differs from case to case. In short, the thesis has three models: an indicative mathematical model (IMM), a statistical model, and an Agent-Based (NetLogo) model.

This paragraph dives straight into basic concepts that orbit the thesis. One of these is demography, which covers birthrates, death rates, migrations, and more to illustrate the changing structure of the human population given a delimited territory (Demography, 2022). The next highlight is migration tied with depopulation. For example, in Spain, depopulation is a widespread phenomenon affecting rural areas, municipalities, capitals, intermediate cities, and more than half of the provincial capitals. Depopulation is one of the demographic challenges that Spain has to face (Bandrés & Azón, 2021). This thesis studies and models a part of the Spanish current demographic situation, namely the Catalan provinces and their recent historical demographic changes.

The thesis tacitly considers the vegetative balance as it scoops data of the Catalan population using yearly figures. Still, as stated above, the thesis focuses on the social and economic side of the equation that leads migrants to make migration decisions at an individual level. Hence, for example, larger salaries and higher vacancies in a specific city logically attract a larger active population as it creates demand for it.

More specifically, this thesis aims at studying migration decisions by focusing on the social and economic parts, leaving other variables as given. While part of the focus is on creating a statistical model that supports the NetLogo one, the choice of the NetLogo model comes from a mathematical model molded by the author. The key inspirations of the mathematical model are its assumptions. For example, the mathematical model poses topics like differences in salaries and other factors such as the number of vacancies available. It also considers government policy differences amongst the cities, such as taxes and subsidies that can influence an individual’s decision to change cities. The mathematical model also covers aspects such as improvements in technology that increase productivity and, therefore, the salary, increasing the utility that it would bring to a worker. In addition, economic shocks can also be entertained. These additions are suggestions for continuing the thesis for future work.

Although a reductionist comparison, the mathematical model, which involves two cities, is not much different than a game of tennis, in which the ball represents parts of the population, and each half of the tennis court represents a city, where the players would be the variables and parameters that determine the changes.

Hypothetically, let us imagine for some reason that an agent such as a government intervenes by lowering taxes or granting aid packages to a specific city or province, e.g., subsidizing part of the salary for the firms, the cost of creating vacancies. This process produces an artificial rise in the chances of relocating to the city that receives governmental aid. In short, with the given indicative model, we play with the utility of the worker, distribution of wealth, and expansionary fiscal policies.

Note that the thesis is done in terms of average salaries during a period instead of wages and that monetary policies are not covered in this thesis.

2. A Theoretical Macroeconomic Model on Unemployment and Depopulation at the City Level

The making of the mathematical model itself is inspired by the on-site taught macroeconomics subject's workload, coursed by the author at the faculty of economics and business of the UAB during the academic year 2020-21. The following is the mathematical model upon which is based the choice of the NetLogo model:

Consider a one-period economy made up of two cities.

There is a mass equal to one consisting of risk-neutral workers, all unemployed at birth.

Firms have free entry to establish in any city to maximize profits.

The productivity of firms $n \in \{1,2\}$ is represented by y_n.

We assume that $y_1 \neq y_2$.

There are four stages:

- In the first stage, the initial corresponding allocation of the population between the two cities is expressed by β, which is predetermined and is thus exogenous.
- In the second stage, in each city, the matching between vacancies and unemployed takes place with respectively assigned indicator probabilities $p(\theta_n)$, where θ_n denotes the size of the degree of the narrowness of the labor market of n city, which is the relation between workers and vacancies.
- Production and consumption occur during the third stage.
- In the possible fourth stage, a "shock" can occur that makes a small or large part of the population return to stage one, that is, to get unemployed, with the new β adjusted to it. For those affected, this "shock" determines the new assignment of city and employment, where both the number of

vacancies and unemployed are adjusted to obtain the correct probabilities, so the parameter θ_n is adjusted marginally, emphasizing the variation of the unemployed, as explained later.

We assume that the living costs depend on and vary according to the parameter β.

Next, we assume throughout the stages that the number of inhabitants remains constant and allocated between the two cities. In addition, because the proposed model focuses on strictly economic aspects, we control birth and death. Consequently, this allows us to control for the population's average aging since these are factors that can contribute to depopulation, but do not necessarily have an economic meaning unless we try to find some correlations. For this reason, we assume that the percentage of deaths is equal to the percentage of births and that they are balanced in a way that, over time, the active population also remains relatively constant percentage-wise.

The utility received by an individual in one city is not the same as in the other. Therefore, as long as there are vacancies, the individual can try to obtain employment in the other city. For example, the higher the productivity in a city, the higher the salaries. Therefore, the higher the utility this city could bring to the individual due to higher expected salaries. We assume that the individual becomes unemployed before emigrating to look for work in the other city.

T_n is a tax that is charged in percentages proportional to the salary in case the government wants to intervene to change the population in a specific city (case exposed in section 1). Although it appears to be a fixed tax, it is indeed proportional to the salary: $T_n = w_n * x_n$, where w_n is the salary and x_n the tax rate $\in [0,1]$. Subsequently, the government would distribute the amount collected between the cities through their respective salaries, being $s_1 = (1 - s_2)$, according to the objectives to be met. In the scenario where the government is not applying its policy, T_n and s_n are 0.

We assume that workers cannot produce anything on their own, nor do they obtain monetary aid while unemployed and that in all cases, the unemployed individual is welcomed by the state via non-monetary aid for the essential necessities of living, ridding them of having to bear the costs of living in the city, for example, food and rent. Hence, this aid is not covered in the model as it is not monetary but a direct material provision to cover basic needs (housing, food, etc.). We can also assume, for example, that this frees you from paying costs such as rent when you are unemployed, postponing them. This argument is especially valid in the face of shocks and unemployment caused by natural disasters. In short, it is an assumption that allows us not to eliminate the active population and drive the unemployed out of the cities, more specifically those in search of work, which in the model represents all those of working age.

On a separate note, the functions p and q shown below are standard. In summary, the U_n and π respectively denote the worker's utility and costs of creating vacancies, respectively; the $p(\theta_n)$ and $q(\theta_n)$ are the indicator probabilities, where θ_n denotes the size of the degree of the narrowness of the labor market of n city, which is the relation between workers and vacancies; the w_n are the sum of salaries firms pay to workers; l_n is the amount of occupied job positions per firm; k represents the bargaining power of the worker over the firm's productivity, which is y_n; c_n represents the cost of living in the city n; if deemed necessary, T_n is the tax, that is, $T_n = w_n * x_n$, and is charged in percentages proportional to the salary represented by x_n; s_1 and s_2 represent the distribution of the amount collected by the government between the cities via the relation $s_1 = (1 - s_2)$.

More specifically, if p and q return 1 for a specific city, it means that the match with said city is successful, and 0 for the other, which means the opposite. The (θ_n), that is, the market size, plays the main role in determining the matches, hence the reference. U_n is the utility of the worker and π are the costs of creating vacancies. Note that a budget constraint for the worker is not included here in the model's layout, but can be implemented.

- The Individual Workers Utility: $U_n = p(\theta_n) \cdot \left(\frac{w_n}{l_n} \cdot (1 + s_n) - T_n - c_n\right)$
- Free Entry Condition for Companies: $\pi = q(\theta_n) \cdot (y_n - w_n \cdot (1 + s_n))$
- Salaries: $w_n \cdot (1 + s_n) = k \cdot y_n$
- Population and Its Variation:

In the following, P represents the total population of the two cities; P^{AP} represents the fraction of active population between the two cities; β is the parameter that endogenizes the population changes (it is initially predetermined); b represents the percentage of births over P while d is the percentage of deaths over P. Therefore, since we assume for our case that $b = d$, the vegetative balance is 0.

$$P = P\beta + P(1 - \beta)$$

$$P = P(1 + b) - Pd \ \ where\ b = d\ (hence\ P = P\beta + P(1 - \beta)\ is\ condition\ enough)$$

$$P^{AP} = \alpha P\beta + \alpha P(1 - \beta)$$

$$t_1(\theta_2^*, \theta_1^*) = \alpha P\beta \pm \sum_{i=1}^{m} \min(v_{2i}, u_{1i}) * e(p(\theta_n^{\pm}))$$

$$t_2(\theta_2^*, \theta_1^*) = \alpha P(1 - \beta) \pm \sum_{i=1}^{m} \min(v_{1i}, u_{2i}) * e(p(\theta_n^{\pm}))$$

Hence, $t_1 = P\beta$ and $t_2 = P(1 - \beta)$ are the population of cities 1 and 2 respectively, which change and depend on the size of the degree of the narrowness of the labor market of the two cities. They change specifically by their vacancies (v_n) and the unemployed (u_n), where $\theta_n = \frac{v_n}{u_n}$ and θ_n is adjusted marginally. This is due to the possible entry of a new individual to the destiny city because their match has happened, so there would be a free vacancy in the city of origin if the individual worked at origin, or one less unemployed. It is vice versa in the other city according to the situation of the individual.

The following formula displays the variation amongst cities, where for each case, each city goes either through addition or subtraction of its total number of vacancies and unemployed: $\theta_n^{\pm} = \frac{v_n \pm v_n}{u_n \pm u_n}$. Therefore, the marginal variation in v_n and/or u_n can be 0 depending on the specific case. For example, the sum denotes the variation between the populations of the two cities according to the minimum between vacancies in city 2 and unemployed in city 1 subject to e, where $e \in \{0,1\}$, adding only when $e = 1$, that is if there is a successful match, while the same is subtracted in the other city.

Competitive Equilibrium:

The equilibrium consists of the size of the degree of the narrowness of the labor market θ_1^* and θ_2^*, the salaries w_1^* and w_2^* and government policy $\{T_n, s_n\}$ such that, given the parameters $\{\pi, y_1, y_2, k\}$, the following conditions are met:

1. Given θ_n^* and w_n^* for $n = \{1,2\}$, the inhabitants/workers initially make the optimal decision to locate, which is defined in β.

2. Free entry condition: $\pi = q(\theta_n^*) \cdot (y_n - w_n^*)$ in the city $n = \{1,2\}$.

3. Individuals salary: $\frac{w_n^*}{l_n} = ky_n$, where k is the bargaining power of workers.

4. The government budget is balanced: $T_n^* = s \cdot w_n^*$

5. The population varies and balances according to the size of the degree of the narrowness of the labor markets and can, for example, be prevented from reaching critical levels through government intervention: $t_1(\theta_1^*, \theta_2^*) = 1 - t_2(\theta_2^*, \theta_1^*)$.

3. Catalan Interprovincial Migration Model, 2008-2020

The Catalan Interprovincial Migration Model's (CIMM) structure departure point is Santillana's (1981) model about internal interprovincial migrations in Spain and their economic determinants. Although an old reference, the base structure of Santillana's model still holds well and is chosen for the effectiveness and simplicity conveyed when describing migration fluxes and determining their causes.

Thus, the CIMM is similar in structure to the Spanish interprovincial migrations model. The main difference is that while Santillana's model studies the Spanish case, the CIMM studies the Catalan case. The general objective of the modeling is to determine the relevant causes and their weight in the Catalan interprovincial migration case. This process also serves to aid the NetLogo model empirically.

3.1. Structure of the CIMM

As for the model's assumption criteria, we assume that the migrants base their moving decisions on rationality, which implies that they are searching for a better economic deal. In other words, the migrant is looking to maximize their net income during their life cycle.

Since the geographical units for studying are the four Catalan provinces, namely Barcelona, Girona, Lleida, and Tarragona, the potential migrant can pick any from the other three provinces for migration. Hence, in this case, the total interprovincial fluxes are $4x3 = 12$. The annual migrant flows data and the province population at the end of each year determines the dependent variable.

Now, given that there is a trend in males to migrate more than females on average, to have higher salaries than females on average and the fact that males are both amongst their category and in total more employed in percentages than their female counterparts, it would be fair to make a case study for each case. Since males migrate more and they have more incentives to migrate than females, it is reasonable to assume that they base their migration decisions on maximizing their net income during their life cycle. But, although it is indeed possible to analyze the migrant flows by gender and age groups, the only case study done is the total migration flows regardless of gender and age, although the population data is of the working age. The reasoning behind the decision to study only the total cases is related to:

- "New economics of migration theory." The theory states that not just individuals, but also social units such as households make migration decisions (Yaukey, Anderton, & Lundquist, 2015), meaning that, in some cases, the migration of a male also comes with a female. In other words, many migrations can be family migrations, thus making many males and females move together.
- Given the data from the Spanish National Institute of Statistics (INE), one can extract information and appreciate that the quantity of migrant flows per gender is, in most cases, similar to the difference in employment rate by gender across all four provinces. The employment rates do not exceed the 60% to 40% male to female ratio in all annual migrant fluxes with more than 500 individuals in total.
- In the Catalan case, the flux data shows that the gender gap in migrant flows lowers the higher the flows are. Hence, we find an inversely proportional relationship between the migrant flows per gender and the relative number of flows. In other words, in the relatively higher migration cases, which are also the most relevant ones, the gender gap in migration is less. This observation reinforces the household migration theory and is consistent with the given employment rate, where the male gender is among itself more employed than its female counterpart.
- Ultimately, the flux data by gender shows that although Catalan males are always a little more prone to migration, the difference is consistent and not deemed relevant enough for making a case study for each group in the Catalan case. Hence, groups are not the subject of study.

Now that we have a study group and its migrant flows, we proceed to define our dependent variable. Three categories can compose the dependent variable: the migrant fluxes from province i to province j (M_{ij}), their origin province's working-age population P_i, and their destiny provinces working-age population (P_j). As the working-age population of each province is different and varies each year, it is essential to normalize the interprovincial migrant fluxes according to the population sizes. To normalize, we use the expression M_{ij}/P_i as it is the proportion between migrant fluxes and population of origin for each case, which is the migration probability. With M_{ij}/P_i we avoid the heteroskedasticity derived from the possible proportionality between the population size and error term (Santillana, 1981). We refer to our dependent variable as ProbMig, which is our short for migration probability.

Next, we choose independent variables that reflect the conditioning factors of the benefits and costs of emigration. As for the costs and risks of migration, the destiny province unemployment rate (dUnemp.), the stock of migrants (Stock.ij), and the distance in km between the capital of each province (Dist.ij) reflect both the monetary and non-monetary costs of migration. Specifically, the unemployment rate represents the risk of migration assuming that employment is the main reason for migration.

Other essential explanatory variables that express the cost and benefit relationship are the salary of origin (oSal.i), the salary of destiny (dSal.j), origin employment rate (oEmpl.), destiny employment rate (dEmpl.), and origin unemployment rate (oUnemp.), GDP of origin in terms of people of working age (oGDPWA) and GDP of destiny in terms of people of working age (dGDPWA).

3.2. Hypotheses

The origin (oSal.i) and destiny salaries (dSal.j) reflect the potential profits; hence they must be introduced as our explanatory variables. The hypothesis is that migrants move from a relatively lower average salary province to a higher one.

The emigrant Stock present at destiny (Stock.ij) is the number of people born in province i who now reside in province j. The stock present at destiny represents the migrant's network potential and it acts as the best possible source for diminishing risks that come from migration.

The distance between province i and j (Dist.ij) is an approximate variable that reflects costs of transportation, psychological, and information costs. Hence, the hypothesis related to distance is that the higher the distance, the lower the migration fluxes, an inversely proportional relation.

The origin (oEmpl.) and destiny employment rate (dEmpl.) reflect market size. The hypothesis is that a migrant is more prone to move towards relatively larger market sizes.

As for the origin (oUnemp.) and destiny employment rate (dUnemp.), the origin unemployment reflects proneness to migrate to look for work elsewhere, while the destiny unemployment represents the migration risk taken when moving for work as jobs are not guaranteed.

The origin (oGDPWA) and destiny GDP per person of working age (dGDPWA) reflect the relative stack of the pie per person in working age between origin and destiny provinces in nominal values. A relatively higher output per person represents a higher quality of life as it incorporates topics like productivity and efficiency, which, for example, can vary with technological advances. Hence, the hypothesis is that a higher prospect of GDP per person attracts a higher number of potential workers.

3.3. Data Sources and Treatments

This subsection lays out the sources of data used for the thesis and its treatments to obtain practical input data and information from the database.

3.3.1. Real Salaries

Both the origin (oSal.i) and destiny salaries (dSal.j) for each province are from the Spanish social security database. The problem with the provincial salaries is that they are only available from 2018 onwards for both genders. It is essential to find further data on salaries as accurately as possible with the correct relative proportions between each province. Although this was not possible at the time of the writing of this thesis, it is still possible to estimate the salary for 2008-2017. For this, an abductive research method is opted for. The method goes as follows:

- Firstly, we need the relative proportions of nominal salaries for each province for at least one year since the relative proportions are more important than the actual numbers, both for the statistical and NetLogo model. The Spanish social security database does offer that accurately and per gender for recent years. To apply the abductive method, we use the average of the year 2020 salary data as a reference as they appear to maintain the usual relative amount even after the hit of an inhibiting pandemic.
- Secondly, it is necessary to have the average salary of the autonomous community of Catalonia for the years 2008-2020. This data is accessible from INE.
- Thirdly, we need the amount of working population for each year. These figures are obtainable via INE.

After gathering the data, we proceed by putting them together in a table per year, province, and gender in the following manner:

- For each year, province, and gender we assign the corresponding proportion of workers population of all Catalonia, where the sum of all Catalonia's working population for a given year is equal to one.

- Now we assign the 2020 relative salaries to each corresponding province and gender of the year 2020.

After laying out the data together in the table, we can proceed to estimate the salaries with the most realistic proportion possible:

- We begin by adjusting the sum of the average 2020 salary per province and gender. This adjustment involves the salaries of each category multiplied by the value given by the following division: $\frac{Catalonia's\ 2020\ average\ salary}{Catalonia's\ 2020\ average\ salary\ according\ to\ sum}$. This way, the departure point for the successive adjustments is adjusted one to one.
- Given that the process is chronologically reversed, our next year is 2019. For 2019 we multiply all subsets of salaries by the following value: $\frac{Catalonia's\ 2019\ average\ salary}{Catalonia's\ 2020\ average\ salary}$. We do this process consecutively til we estimate salaries until 2008.

Note that the salaries obtainable from the abductive method can now be used as input for the statistical salaries once they adjust from nominal to real salaries. For this, we use the formula $\frac{Nominal\ wage\ in\ a\ year}{CPI\ in\ a\ year} \cdot 100 = Real\ Wage$, where CPI stands for the consumer price index. The annual average CPI data of each province is accessible through INE's database. We also need the total average salaries per province instead of gender and province. This can be calculated since we know the amount of working population for each gender for the years 2008-2020. The model's input real salaries for modeling are approximate values that are the average real purchasing power of an individual per province regardless of gender.

3.3.2. Migration Fluxes, Population, and Emigrant Stock

The data on annual interprovincial migration fluxes and population is obtainable from INE. While the migration flux is the total of each year, the population data used for our model is the population at the very end of each year. The data available for the first day of a year is our approximate for the very end of the previous year.

The emigrant Stock present at destiny (Stock.ij) approximates the average of the fluxes from origin to destiny between the years 2008-2020, with 12 fluxes in total. This discounts for the decrease in relationship capital, a concept discussed in section 5.5.12.

3.3.3. Distance

The distance between province i and j (Dist.ij) is the distance between the capital of each province. Hence, the same value repeats twice, making a total of $2x6$.

3.3.4. Employment and Unemployment Rate

The origin employment rate (oEmpl.), destiny employment rate (dEmpl.), origin employment rate (oUnemp.), and destiny unemployment rate (dUnemp.) are extractable from INE.

3.3.5. GDP per Person of Working Age

The origin and destiny gross GDP at market prices are obtainable by INE. We obtain our origin (oGDPWA) and destiny (dGDPWA) GDPs per person of working age by dividing the gross GDP at market prices by the corresponding amount of working-age people. This GDP is our measure instead of GDP per capita as it is more relevant to the study group.

3.4. Weaknesses of the Statistical Model

We simplify and study the case of circular migration in the NetLogo model. The concept of circular migration is simple: it refers to the type of migration in which the individual travels between origin and destiny, where the destiny can range from cities, provinces, countries, and more, all in search of work. The problem is that the aim is to make an accurate migration model in the case of Catalonia. This aim would not be possible given the lack of official empirical data and surveys on returning migration and migration history for Catalan interprovincial migrants. However, there is an approximate way to solve this for the NetLogo model which is discussed later on.

The treated migrant flow data does not include the exact age of the migrants in smaller subgroups, but they are all part of the working-age population.

The statistical model should ideally enrich with data related to GDP per person of working age for each province when put together with the salaries at once. But this is not the case as these two together do not fit well when modeling, given the model's inherent structure of comparing origin and destiny. It is also because the GDP and salaries have similar trends in our case study, which leads to more problems related to multicollinearity. Multicollinearity makes the regression unstable and inflates the standard errors. Hence, multicollinearity is tested for and confirmed via the correlation matrix between all independent variables, where the correlation between cor(oSali, oGDPWA) and cor(dSalj, dGDPWA) is positive and very high. Multicollinearity is also found through a Variance Inflation Factor (VIF) analysis, in which VIF > 5 for both the salary and GDP variables when put together. The previous is not the case when GDP and salaries are separate. Hence, to solve multicollinearity, it is opted to regress the GDP and salaries separately. RStudio performs the previous analyses.

3.5. Results

a) OLS done via RStudio using 156 observations. The dependent variable is ProbMig.

$$\text{Log-Log Linear Model: } \ln(\text{ProbMig})=\beta_0-\beta_1\cdot\ln(\text{oSali})+\beta_2\cdot\ln(\text{dSalj})+\beta_3\cdot\ln(\text{Stockij})-\beta_4\cdot\ln(\text{Distij})-\beta_5\cdot\ln(\text{oEmpl})+\beta_6\cdot\ln(\text{dEmpl})-\beta_7\cdot\ln(\text{oUnemp})+\beta_8\cdot\ln(\text{dUnemp})$$

```
Residuals:
     Min       1Q   Median       3Q      Max 
-0.47661 -0.11567  0.00543  0.10900  0.45368 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)    7.82840    5.87811   1.332   0.1850    
log(oSal.i)   -7.84009    0.32391 -24.205  < 2e-16 ***
log(dSal.j)    6.81388    0.31795  21.431  < 2e-16 ***
log(Stock.ij)  0.47153    0.02403  19.626  < 2e-16 ***
log(Dist.ij)  -0.69257    0.06824 -10.149  < 2e-16 ***
log(oEmpl.)   -0.47741    0.60577  -0.788   0.4319    
log(dEmpl.)    4.70864    0.60457   7.788 1.12e-12 ***
log(oUnemp.)  -0.25496    0.13295  -1.918   0.0571 .  
log(dUnemp.)   0.90903    0.13334   6.817 2.23e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1812 on 147 degrees of freedom
Multiple R-squared:  0.9742,	Adjusted R-squared:  0.9728 
F-statistic: 693.4 on 8 and 147 DF,  p-value: < 2.2e-16
```

For both models a) and b) the 156 observations consist of 13 years x 12 migration fluxes. Although the relationship between our dependent and independent variables is linear, log-log regressions are opted for as they give us elasticities. In contrast, the log-log elasticity coefficients are much more convenient for interpretation than the linear-linear regression ones because the linear ones are too small for meaningful interpretation.

Recapitulating, the previous table displays our OLS regression output, displaying the elasticities. This means that for each 1% increase in an independent variable, the dependent variable increases by the value of the independent variable β coefficient. At a 5% level of significance, all the variables are significantly different from zero except for oEmpl and dUnemp as they have p-values higher than 0,05.

The R-squared is 0,9742, meaning that 97,42% of the variation in the probability of migrating is explainable by the origin and destiny salary, emigrant stock at the destination, the distance between the capital of origin and destiny, origin employment and unemployment, destiny employment and unemployment.

The empirical results are similar to the expected ones. Given the table, the following hypotheses are not refutable:

- For starters, we can confirm that the difference in origin and destiny salaries are factors for migration. When studying the origin and destiny salary relation, we can see the value of the origin salary coefficient is approximately the same in absolute values but is negative. Given that the dependent variable represents the probability of migration, this relation expresses that people are more prone to remain at home rather than emigrate for work. But relatively speaking, individuals consider migration the higher the salary prospect is. This is inferable since origin salary has a negative sign while destiny salary has a positive sign.
- The stock variable includes even the 2^{nd} and 3^{rd}+ connections that have one thing in common with the migrant: sharing the origin and destiny. The table shows that the higher the number of possible connections amongst those that share a flux throughout the years, the higher the probability of migration given that the variable Stockij's coefficient is positive.
- We can interpret from the table that the higher the distance, the less prone individuals are to do interprovincial migrations.
- Given the p-values of oEmpl and dEmpl, we see that the origin employment may not influence migration decisions much as it is not statistically significant. On the other hand, the destiny employment rate matters much more when migrating.

b) OLS done using 156 observations via RStudio. The dependent variable is ProbMig.

Log-Log Linear Model: $\ln(ProbMig)=\beta_0-\beta_1\cdot\ln(oGDPWA)+\beta_2\cdot\ln(dGDPWA)+\beta_3\cdot\ln(Stockij)-\beta_4\cdot\ln(Distij)-\beta_5\cdot\ln(oEmpl)+\beta_6\cdot\ln(dEmpl)-\beta_7\cdot\ln(oUnemp)+\beta_8\cdot\ln(dUnemp)$

```
Residuals:
     Min       1Q   Median       3Q      Max 
-0.82508 -0.21403  0.01271  0.24414  0.74067 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)    21.1444    18.6529   1.134   0.2588    
log(oGDPWA)   -12.3109     1.0044 -12.257  < 2e-16 ***
log(dGDPWA)    10.0501     0.9933  10.118  < 2e-16 ***
log(Stock.ij)   0.4669     0.0413  11.306  < 2e-16 ***
log(Dist.ij)   -0.7381     0.1259  -5.862 2.89e-08 ***
log(oEmpl.)    -5.1067     1.0998  -4.644 7.53e-06 ***
log(dEmpl.)     8.7739     1.0859   8.079 2.18e-13 ***
log(oUnemp.)   -0.5280     0.2431  -2.172   0.0315 *  
log(dUnemp.)    1.1972     0.2438   4.911 2.39e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3334 on 147 degrees of freedom
Multiple R-squared:  0.9126,	Adjusted R-squared:  0.9078 
F-statistic: 191.9 on 8 and 147 DF,  p-value: < 2.2e-16
```

Note that model a) offers a better fit to the data given that it has a higher log-likelihood-ratio test value of 1067,47 over the 972,357 value of model b).

We can observe that using the origin and destiny GDPs as our main twin variables has similar effects as using the salary twin variables. The effects amplify in model b) as our significance is also enhanced. Every variable is significant at 1% except for oUnemp, which is significant at 5%. Note that our R-squared has diminished from 97,42% to 91,26%.

More specifically, it is observable for GDP per person of working age that the migrant behavior patterns are similar to the salary ones when making migration decisions. If we interpret the oGDPWA as the amount that shows how meaningful and productive the work is, the migrant cares about both the origin and destiny

GDPs as both are statistically significant. In short, our initial hypothesis for GDPWA is not refutable.

On the other hand, the following is not as expected for both models:

- The case of oUnemp and dUnemp, which are the origin and destiny unemployment rates, is similar to the employment case. Although the origin unemployment is not statistically significant at a 5% significance in the first model, it is the second one. According to our model, migrants are prone to travel where there is currently more unemployment given that dUnempl's coefficient is positive. Depending on the case study, this phenomenon could happen because more vacancies might be waiting to be filled where there is more unemployment. One argument against the previous statement is that, for example, unemployment can lower in times of recession as the main breadwinner of a household may suddenly find themselves out of a job. This issue may force other family members to find jobs to compensate for the loss in household income, even if they may not pay as well. Hence, the effects of unemployment can depend on circumstances and remain inconclusive in general terms.

4. Agent-Based Modelling

In simple terms, an Agent-Based Model (ABM) is a computational model that simulates actions and interactions between agents to understand better the behavior of a system, and the parameters that govern it and to study hypothetical cases (Agent-Based Model, 2022).

4.1. What is NetLogo and Why Use It

Designed by Uri Wilensky, NetLogo is a programming language and an integrated development environment for ABM. The simplicity of NetLogo paired with its extensive model library for fields such as economics, physics, biology and so make it a go-to tool for broader types of audiences. At the same time, it serves as an introduction to the programming world and the logic behind it to the inexperienced (Agent-Based Model, 2022).

Given NetLogo's nature and the fact that it is open source and relatively easy to use than some other programs, it is one of the best academic tools for analyzing topics that combine demography and economics, such as our topic.

4.2. Decisions in Circular Migration

Given that in the statistical model we must assume that migrations are permanent, we use the NetLogo model to study the case in which they are not permanent, namely, the circular migrations.

I have used Anna Klabunde's economic migration model and adapted the input data, parameters, and some parts of the code to analyze three circular migration cases from the given 12 interprovincial fluxes. To check further details on the model, please consult Bibliography.

5. The NetLogo Model

The model has two types of agents: firms and workers. They are spread randomly on a grid. The heterogeneity among the workers comes from their time-specific savings and home preference parameter.

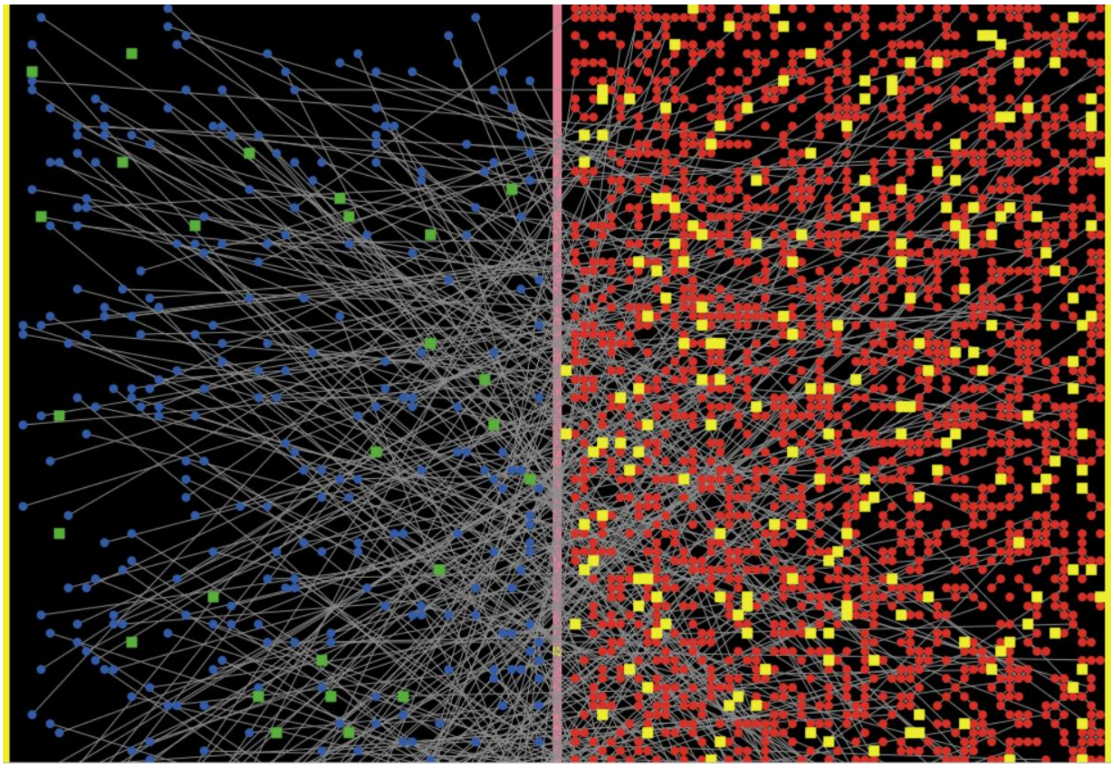

Figure 1. Model's initial setup with Girona as home province. The original NetLogo 4.1.2. model is adapted to version 6.2.2. The original model's code is available at https://www.comses.net/codebases/3893/releases/1.2.0/.

The grid in Figure 1 represents two provinces. One is the host province, which is the province that receives the migrants and is always the one to the left of the middle vertical line, which is the border between provinces. The home province is on the right side. The red and blue circles are workers. Workers can move between the two provinces, but firms, which are squares, cannot move across provinces. Our host province (Barcelona) is a higher productivity province, while the home provinces for each case, namely Girona, Lleida, and Tarragona, have lower productivity.

5.1. Fundamental Concepts

To start the model, we click on the setup button. Clicking randomly generates a home and host province separated by a border. After this, workers generate in the home province equal to the slider number, for example, 3152, which is the case where Girona is the home province. Then, a percentage equal to "100 – percent at home" of workers move to the host province, which for Girona's case is equal to 12.5%. Each worker is assigned to a firm and becomes heterogeneous by two things, having their own savings parameters randomly assigned each year via input data. Workers also have different home preferences. The workers in the host province create links with other workers in a radius. The radius is the value of the neighborhood size slider. The workers present in their home province create links in their square-shaped neighborhoods, also known as Moore neighborhoods. As for firms, the number of firms per province is the amount initially assigned by their respective sliders "number-firms-red" for the home province and "number-firms-blue" for the host province. Firms are placed in random spots and set random initial salaries.

5.2. Description of the Processes That Follow the Startup

The link between workers that are not immediate neighbors weakens every period by 2% and only fully recovers through closer physical contacts on the grid. These contacts are a result of migrations.

Workers consume each period part of their salary from the previous period and save a part of their wealth that corresponds to their personal savings rate per period. These savings add to their wealth. In contrast, the potential workers without a salary share a minimum consumption.

Through their network neighbors, workers calculate their expected salary in the host province for the current period. The expected salary is the average salary of their network neighbors, not of everyone in the province:

$$w_{exp,i,t} = \frac{1}{N}\sum_{n=1}^{N} w_{n,t}$$

where salaries are expressed with w and $n = 1, \ldots, N$ are the network neighbors of a worker at time t that are at the host province.

After computing salaries, the next step involves the decision to emigrate. The decision has three sub-steps:

1) During the first sub-step, migrants compare whether their expected salary is higher than their current salary and if their wealth exceeds their moving costs.
2) If the previous is true for some migrants, they proceed to the second sub-step, which involves calculating their moving probability. Their migration probability consists of the following:
a. A baseline migration probability (p_0), which can be determined by having data on failed migrations, but we lack data on the matter.
b. The difference between expected and current salary if it is their first migration ($p_{1,i}$). If it is not their first time migrating, they already know what real difference to expect.
c. The amount of stock at the host province of one worker that is present at home. In other words, the worker's network in the host province (p_2). This is similar to our Stockij independent variable in the statistical model. The probability p_2 is zero if a migrant at home has no stock in the host province. In practice, if migrants have links with host neighbors, incoming migrants choose to go where the neighbor with the highest salary is. Otherwise, migrants go to a random spot in the host province.
d. The worker's individual home preference is inversely proportional to the migration probability. This makes probability $p_{3,i}$ have a negative sign (- $p_{3,i}$).
The parameter that would adjust the migration probability per age is $p_{4,i}$, but it is not accounted for since our dataset includes the entire active population.
Workers decide to migrate if the randomly drawn number between (0,1) is smaller than their migration probability.

3) The last step involves accounting for the distance between home and the host province's capital, which in contrast with real-life migrations, reflects overcoming non-monetary costs of establishing at the host province, e.g., psychological costs, finding a new spot to stay...

In summary, the probability that worker i migrates at time t subject to their wealth K being higher than costs m_1 is the following:

$$p_{i,t}\,(migrate|K_{i,t} > m_1\,, w_{exp,i\,,t} > w_{i,t}\,) = p_0 + p_{1,i}\left(w_{exp,i\,,t} - w_{i,t}\,\right) + p_2 N_i - p_{3,i}$$

where $K_{i,t}$ expresses the worker's wealth during time t and m_1 the migration costs.

After migrants overcome these steps, they secure a spot in the host province, albeit with a lower wealth due to moving costs.

Note that before moving, migrants become unemployed and go through frictional unemployment. As for migrants present at home and host that do not migrate but are also unemployed, they move towards their network neighbor with the highest salary in the same province. If they do not have links, they move in random directions in their province until they find employment. In other words, individuals move until they end up on the same patch (square) as a firm.

Migrants can also return to their homes. For returning migrants, the process is partially analogous to the migration one. The differences are the following:

1) Although the home preference probability is positive when returning, the ties to home are relative to the age of links, meaning that as time passes without migrating back home, links with home get weaker. This concept incorporates the effects of stocks at home.
2) The distance does not matter here, i.e. for when going back home.

In summary:

$$q_{i,t}\,(return|K_{i,t} > m_2) = q_0 + \sum_{r=1}^{R} \frac{q_1}{a_{r,t}}$$

where $q_{i,t}$ expresses the probability to return of a worker at time t, which is dependent on whether their wealth $K_{i,t}$ is higher than their migration cost m_2, q_0 is the baseline return probability, q_1 is the parameter ties-to-home-province, $r = 1, \dots, R$ expresses the worker's network in the home province, $a_{r,t}$ expresses the age of a link.

Once the model runs all the ticks, which are the periods separated in years, it generates output for each period. Our modified model is similar to the original. The difference is that some things in Klabunde's original case study about Mexican migrants in the US do not apply in our model, for example, age and border controls.

Note that it is possible to add the original population of both the host and home province at once. But it is opted to analyze specific streams of migrations and the total population with home province as the origin. This approach allows us to compare our migration and return reference data with the simulated ones.

5.3. NetLogo Model's Input Data

5.3.1. Salaries and CPI

The input salary for both the home and host provinces is the annual nominal salaries obtained using the abductive method. Given that these sets of salaries are statistically significant and give a better R-squared than the empirically determined GDPWA, they are valid for use. Furthermore, relativeness is more important than the actual values of the salaries, which is an aspect that we account for.

The input salaries for the model are nominal because the model accounts for inflation using the CPI. The annual CPI per province is obtainable via INE.

5.3.2. Stocks

The stock data represents the percentage of people born in origin and currently present at destiny. We use the IDESCAT data on the population on the 1st of January per place of birth. For consistency, the data of the first day of each year is our approximate for the very end of the previous one as this is the best source and approximation found. The data on stocks is accessible via the Statistical Institute of Catalonia (IDESCAT). Since the number of people born in the home province i and currently residing in the host province j is our equivalent to Stockij, the formula to obtain the stock is: $\frac{Stockij}{Home\ Province\ Active\ Population + Stockij}$.

5.3.3. Migration and Return Probability

For migration probability, the standard migration flux probability seen in section 5.2. is used.

As for the return probability, it is not as straightforward. This is because there is no official data available on interprovincial circular migration to compare with the simulated return rates. Due to this, the abductive method is opted for once more. It involves the following:

1) We take the annual data that is our "Stockij" and calculate the raw amount of the variation year by year starting from the years 2007-2008. For reasons explained further ahead, negative values emerging from this are not considered.
2) Next, we subtract the amount of home to destiny province annual flux by the previous variation. This amount shows us how many of them did not stay, in other words, the ones that would leave and have a high chance of returning right back to their homes, giving rise to circular migration. The previous process can be refined for future studies by multiplying the population by the percentage that represents the autochthonous population to exclude all foreign migrants. This is because our Stockij includes those born in a Catalan province i that are residing in province j. Further refining is not considered for this analysis as we already refine for the working-age population.
3) Finally, we use the following formula for each year to calculate the return probability, which summarizes the previous steps:

$$\frac{(Home\ to\ Destiny\ Province\ Annual\ Fluxij) - Variation\ in\ Stockij}{Home\ Province\ Active\ Population + Stockij}$$

In step two, the negative Ci values do not factor in since it is not logical to have negative return rates since we differentiate between migration rate and return migration rate. To clarify further, the following is a numerical example for the case where Girona (GIR) is the home province and Barcelona (BCN) the host:

A	B	C	D	E	F	G	H
Province	GIR	**STEP 1**	Origen/Home	GIR	.	.	.
2007	85899	.	Destiny/Host	BCN	**STEP 2: Ei - Ci**	Home Province Active Population + Stockij	**STEP 3: Fi / Gi**
2008	87107	1208	2008	6128	4920	701027	0,007
2009	87625	518	2009	6464	5946	704149	0,0084
2010	88500	875	2010	6560	5685	707434	0,008

Table 1. The red numbers represent years; the A – I classify columns; the blue words represent descriptions; data source: INE and IDESCAT.

5.4. Stylized Facts about Return Migration

The following are some simplified facts on return migration.

5.4.1. Links

Links are a migration-specific element related to the migrant's network, meaning that with each move, the migrant builds new connections that aid in finding a job, information, and more. The determinants behind the formation of links are pre-established in the model. This concept is also known as relationship capital. See Klabunde (2014) for further details.

5.4.2. Distribution of Migrants

Links influence the distribution of migrants across our model's neighborhoods in a heavy-tailed manner, assuming that it follows a power law. See Klabunde (2014) for more details.

Additionally, people from one neighborhood migrate to the same places due to positive network externalities. For additional details, check Klabunde (2014).

5.4.3. Behavioral Assumptions and Migration Hypotheses

Our statistical model's logit regressions express that the implemented variables are valid as an input into the model. More specifically:

1) The hypothesis that higher expected salaries in the destiny province attract migrants is not refutable by the statistical model.

2) The hypothesis that previous migrants in the host province, that is, Stockij, increase the chances of migration is confirmed by our statistical model.

3) The model assumes that the higher is a migrant's home preference, that is, their utility is higher when at home, the less likely they are to emigrate. See Klabunde (2014) for further confirmation.

4) The model assumes that the more abundant the ties to the previous place of a migrant, e.g., family, the more likely they are to return. This is the relationship capital. For empirical relevance, consult Haas & Fokkema (2011). Since we lack official interprovincial data on return migrations, we take the assumptions as a given. Note that this applies to both home and host provinces and that the importance of ties diminishes over time.

5) Initially, the model hypothesized that the higher a migrant's savings are, the more likely they return home after accumulating enough wealth so that their purchasing power is higher at home. This is plausible if we assume that migrants have a high enough home preference. When empirically contrasted, one can find that, for example, in Klabunde's original US and Mexico case, migrants tended to buy more property in the host country (US), instead of home (Mexico). Though the savings parameter in this case study is statistically significant, it is small. Hence, a saving parameter is not in the NetLogo model. Instead, a set of idiosyncratic skew-normal savings are randomly assigned to each individual during each period.

6) The hypotheses related to education have a wide range, and the empirical evidence is mixed (Kladunde, 2014). Ultimately, it is chosen not to implement education into the NetLogo model. Hence, a uniform level of education is assumed since we cannot account for the education and skills of interprovincial Catalan migrants since we have no figures regarding the matter.

5.5. The Non-behavioral Parameters

5.5.1. Average Neighborhood Size

Through this parameter, we set the radius size of our average neighborhood for the simulation in the case of Barcelona province. To determine the radius value, it is necessary to calculate the average area of half a municipality by dividing the total area of the province of Barcelona by double the number of municipalities. This yields us the average area of our neighborhoods. By analogy, to calculate our average area of a neighborhood in the simulations, the total area of the left part of the grid (122x60), which corresponds to Barcelona, is divided by the number of municipalities multiplied by two. This is the average size of our neighborhoods, which is the area of a small rectangle. We use the following formula to transform this area into a circle of the same size, where we solve for the radius r with A as area and π is pi: $A = \pi * r^2$.

5.5.2. Number of People

The number of people set in the model is "Home Province Active Population + Stockij," but scaled down to 1:225 as the original numbers are too high.

5.5.3. The Initial Percentage at Home

The initial percentage at home between 0 and 1 is:

$$1 - \frac{Stockij}{Home\ Province\ Active\ Population + Stockij}$$

5.5.4. Number of Firms

The formula for getting the number of firms in a province is $\frac{Number\ of\ People\ Initially\ in\ The\ Home\ Province}{Average\ People\ per\ Firm\ in\ Given\ Province}$.

This figures are obtainable for each province through eInforma's website.

In the case of the host province, the corresponding number of firms in host is calculated given the Stockij at host, i.e., only migrants are counted, but the amount of average employed people per firm is the host's one.

5.5.5. Moving and Return Costs

Our moving and return costs are the sum of the costs of the monetary costs of transportation and one month's loss in salaries of the province we move out from. The reported values are at 2017 constant prices, that is, the 2017 nominal salaries + costs of transportation.

5.5.6. Salary and the Salary's Standard Deviation

The input salaries are the same nominal salaries for each province from 2008 to 2020. To further see the formulas the NetLogo model uses regarding the salary and CPI input data and their adjustments for each year, check Klabunde (2014).

Firms pay a uniform and idiosyncratic salary adjusted exogenously for inflation. For calculating the standard deviation of the salary, INE's data on the Catalan average monthly salaries per decile for the year 2020 is used as a reference.

For the standard deviation of salaries in terms of each province reference 2020 salary, the following formula is used:

$$\frac{Standard\ Deviation\ of\ the\ Second\ to\ Ninth\ Decile\ of\ Catalan\ Salaries}{Average\ Monthly\ Nominal\ Salary\ of\ 2020} = Relative\ Standard\ Deviation\ of\ Province$$

5.5.7. Minimum Consumption

Although the original case study calculates the minimum consumption (Cmin) by taking a household's consumption from the first quintile since the study is about the migration history of household heads, this is not the approach we take. This is because, firstly, we lack data about people's migration history. Secondly, since first is the case, we are consequently aiming at studying migration fluxes of individuals regardless of their marital and family status.

Due to this, for our minimum consumption, we take the per-person consumption expenditure statistics for Catalonia for the year 2017 (12833 €). This data is accessible from IDESCAT. Note that this value is coincidently slightly below the minimum consumption of a Catalan household from the first quintile.

We assume that the relation between average salary and the minimum consumption is constant throughout the years, so the formula used for calculating the minimum consumption of each province for the model's code is: $\frac{12833}{Average\ Anual\ 2017\ Salary\ of\ a\ Province} = Cmin.$

5.5.8. Savings Rate

The savings rate is obtainable in the following way:

a) For each year, we calculate the annual tax collected per person in Catalonia by dividing the total personal income tax by the total Catalan-occupied population of that year. The tax figures are accessible via IDESCAT.
b) Next, we subtract the average collected tax per occupied person from the corresponding salaries of each province. This yields us the net income.
c) We proceed to subtract the annual average spending in Catalonia from each province's corresponding average salary. This gives us the annual savings for each province. Note that debt and sources of earnings other than salaries are not considered for calculating savings.
d) Finally, we divide each saving by its corresponding net income by province and year to calculate the savings rate.

The savings rates are converted into the skew-normal distribution. The parameters of the skew-normal distribution are estimated by maximum likelihood using our figures on the savings rate for each province.

Note that the range of our savings rate is restricted between 0 and 1. A savings rate smaller than 0 would mean that someone has to use their wealth in that period, while savings higher than 1 would mean that someone received some earnings other than salary during a given period, such as an inheritance. You can find attached in Appendix the code to generate the skew-normal distribution.

5.5.9. Distance

The original model accounts for border control, but there is no border control in our interprovincial case. Instead, we use the distance to factor in the proclivity toward home, assuming that no migrant is completely indifferent towards home, the home's surroundings, and the non-monetary costs of establishing at the host province. For this, kilometers are normalized from 1 to 356 between [0,1], where 1 kilometer represents the least amount of moving effort, and 356 kilometers is amongst the largest road route found between the furthest point in Catalonia. This is so that the max number of KM between the province, which reaches 226,2, does not become 1. Hence, the normalized value for each migration flux case is different. The probability of migrating is inversely proportional to the distance.

Accounting for distance in this manner helps us differentiate each type of interprovincial migration flux better by assigning its border.

5.5.10. Production Parameter

For each case, the production parameter of Barcelona is 0.5. For the other provinces, this parameter is relative to the 0.5 GDPWA value for Barcelona.

5.5.11. Home Preference

For home preference, figures from a Fotocasa report on housing done via surveys are utilized. The report states that, according to the analyzed survey sample, 14% of Spaniards do not own a home, 20% are not owners but they feel like they have ownership, 51% have one property, 12% have two properties, and 3% have three or more. Since a similar source about Catalunya was not found, to simplify, we assume that, in most cases, at least one owned property per owner is in their home province. Hence, we use this statistically accurate information on the types of property owners in Spain as representative of the Catalan case given that, according to the same report, the total amount of ownership in Catalunya is similar to the Spanish one, just slightly lower (Radiography of the Housing Market 2017 - 18, 2018).

In our model, the home preference is proportional to the number of properties a worker owns. Hence, those with no property have a home preference of zero, while those with the most properties have a home preference of 4.

5.5.12. Decrease in Relationship Capital

The ties to our home province are our model's relationship capital. The depreciation in relationship capital is set to 2% arbitrarily for each simulation. This depreciation is introduced because the longer the migrant is away from home, the weaker the link gets due to the distance. The phenomenon is empirically relevant (Haas & Fokkema, 2011).

5.5.13. Parameter Comparative Earnings and Parameter Homebias

These parameters are left at a value of 1 but are alterable for sensitivity analysis.

5.6. Calibrated Parameters and Robustness

The parameters that need calibration through simulation results are the baseline probability of moving (p_0), baseline return probability (q_0), the behavioral parameter for the number of network neighbors in the host country (p_2), and the behavioral parameter for ties to the home province (q_1). For this, NetLogo's "BehaviorSpace" tool is used. Via this tool, the following experiments are run for each case study:

1) A set of the most suited combinations of all parameters for both our three cases are simulated with centesimal increments. The intervals values are:
    a. [0,01 , 0,09] for "parameter-other-migrants" (p_2)
    b. [0,01 . 0,09] for "baseline_migration-probability" (p_0)
    c. [0,01 . 0,1] for "baseline_return-probability" (q_0)
    d. [0,01 . 0,09] for "parameter-ties-to-home" (q_1)

    Exceptionally, to calibrate further Lleida's case, the interval for baseline return probability is changed to [0,05 . 0,2].

2) The purpose of these 7290 combinations for each flux is to generate their corresponding migrant and return flows. Once generated, they are filtered out in the following manner:
    a. First, the generated values that are around the maximum and minimum of each type of empirically observed flux are filtered.
    b. Next, the frequency of each run is plotted, that is, the frequency of each set of simulations with their steps, where each step represents a year.
    c. Thirdly, the runs amongst the highest frequencies are picked. This is because these are the ones that remain the most under the empirically observed interval.
    d. Finally, for each case study, the simulated runs that best fit the trends of the empirical values both for migrant fluxes and the total amount of migrants at the host are picked.

This leaves four degrees of freedom. Additionally, the model passed the robustness check, that is, after changing input data and parameters, the results vary accordingly.

Alternatively, it is also possible to obtain the best combinations for the flows of migrants by minimizing the following distance function:

$$f_1 = \left(\sum_{t=1}^{t=4} \left(m_{t,emp} - m_{t,sim}\right)^2 + \left(r_{t,emp} - r_{t,sim}\right)^2\right) : 4$$

Check Klabunde (2018) for further details on the model's distance functions.

6. Simulated Time Series of Interprovincial Migrations

The following are the empirical data graphs:

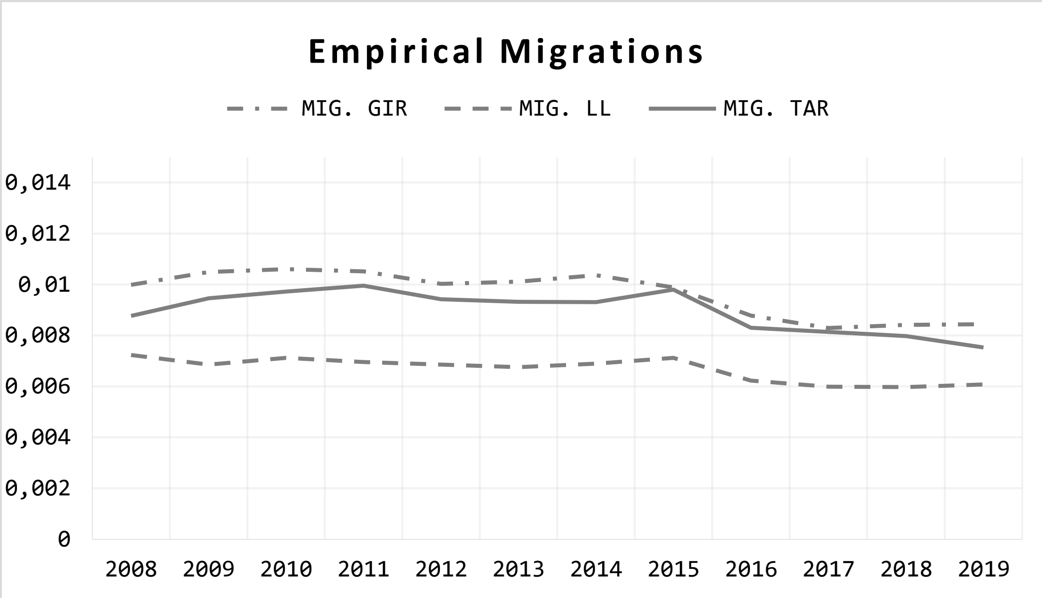

Figure 2. Empirical Migrations

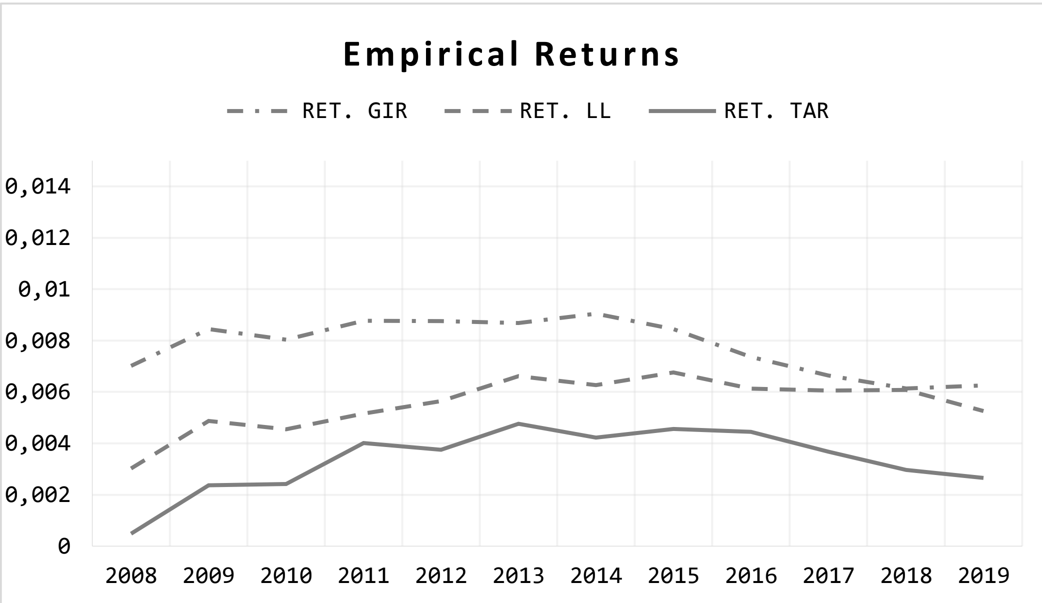

Figure 3. Empirical Returns

The following six graphs represent the results of 3330 Monte Carlo simulations each for both migration and return fluxes of our case studies:

Girona's Case

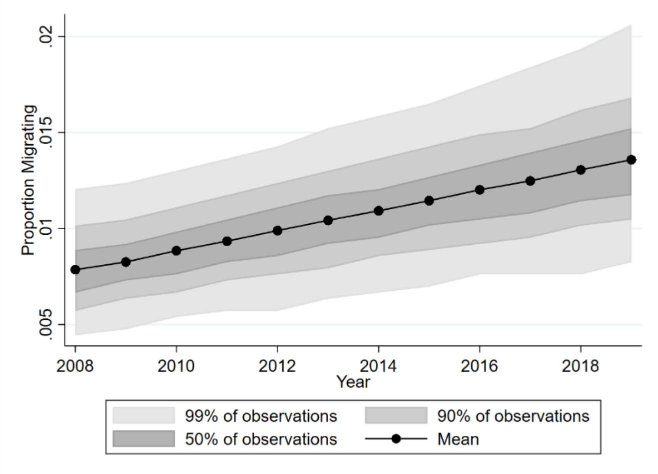

Figure 4. Proportion Migrating - Girona

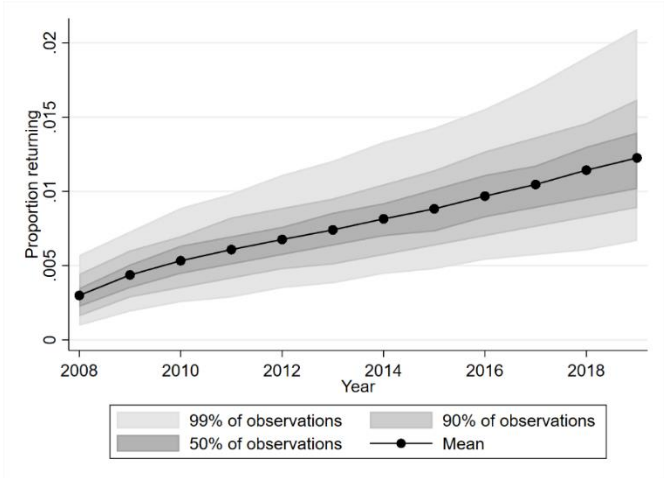

Figure 5. Proportion Returning - Girona

Lleida's Case

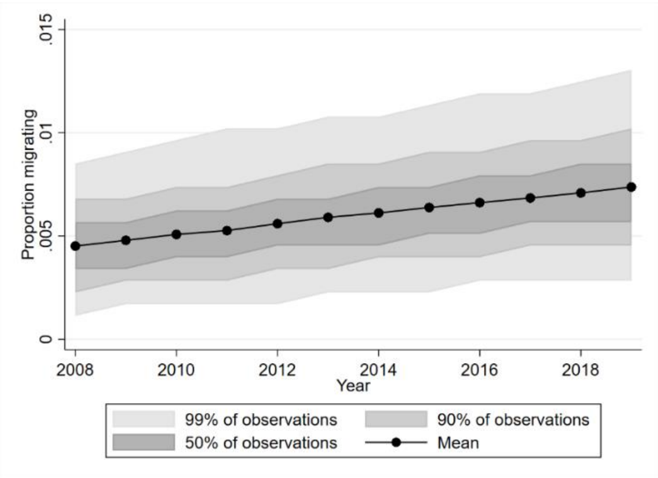

Figure 6. Proportion Migrating - Lleida

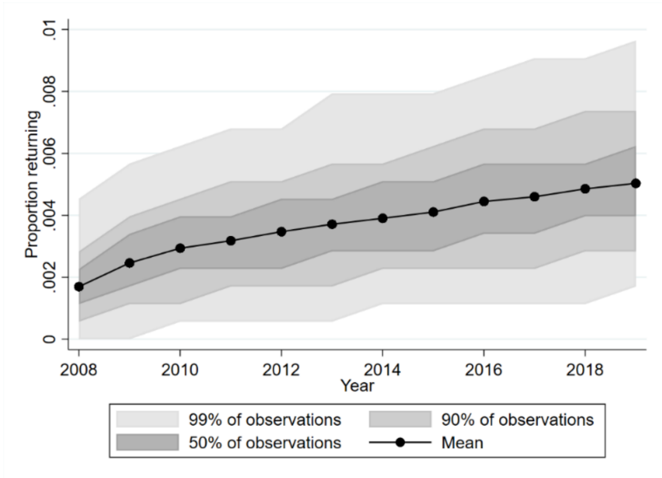

Figure 7. Proportion Returning - Lleida

Tarragona's Case

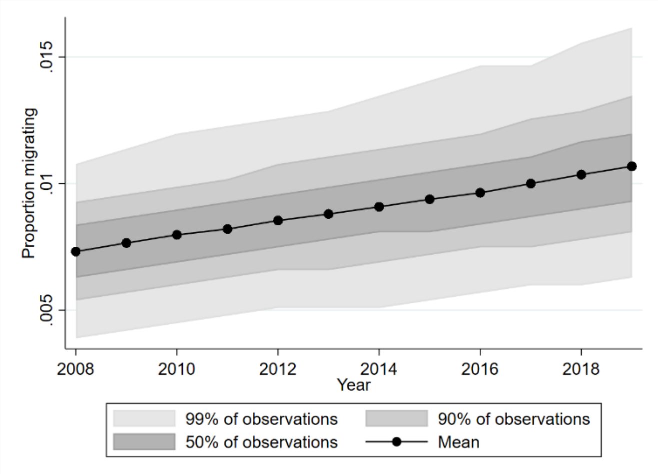

Figure 8. Proportion Migrating - Tarragona

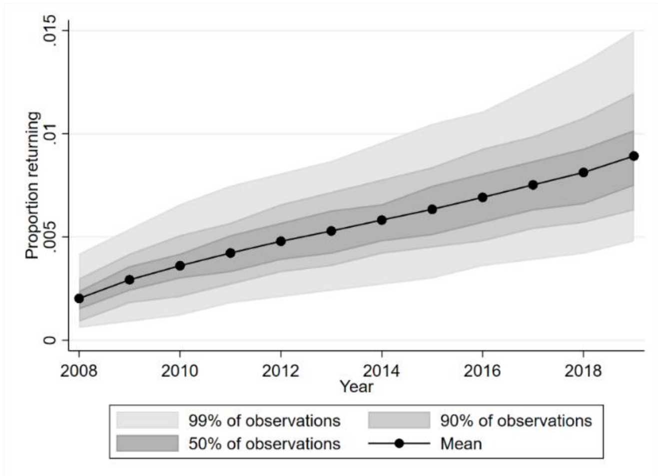

Figure 9. Proportion Returning - Tarragona

Figures four to nine are the simulation graphs. In the simulation graphs, the black line represents the mean of each type of simulation. The upper bounds of each graph represent the maximum reached among 99% of observations during each year of the simulation. The lower bounds represent the minimum of each year. The different colors classify the range under which are a certain percentage of observations. For example, the darkest color contains 50% of the observations.

In contrast with the empirical information in Figures 2 and 3, our simulated figures adjust well into the range of the empirical illustrated data and follows the upward trend well up until the year 2015. Future study is necessary to fine-tune the models so that they adjust to the downward trends after the year 2016. Overall, our models do a good job and adjust well into the empirical data range.

The series of graphs in the next pages, composed of Figures 10 to 15, represent 100 simulations for each type of graph legend, where the first series of graphs show the average stock after an increase in average home salary against the calibrated model. The second series shows the average stock after increments in the savings against the calibrated model. GIR refers to migrants from Girona, LL from Lleida, and TAR from Tarragona.

The following graphs show us that higher home salary prospects and higher savings indeed incentivize individuals in the model to move back and stay at home. The same is true for the opposite, which the non-altered figures reflect. The only case in which this is not as clear is the second graph of series 2, which can be due to a relatively low population, but even in that case, we can appreciate that the number of migrants at host starts to decrease over time after an initial increase. This effect occurs because the higher the savings are, the more are workers able to migrate, but later are incentivized to go back due to the much higher salaries at home.

Series 1

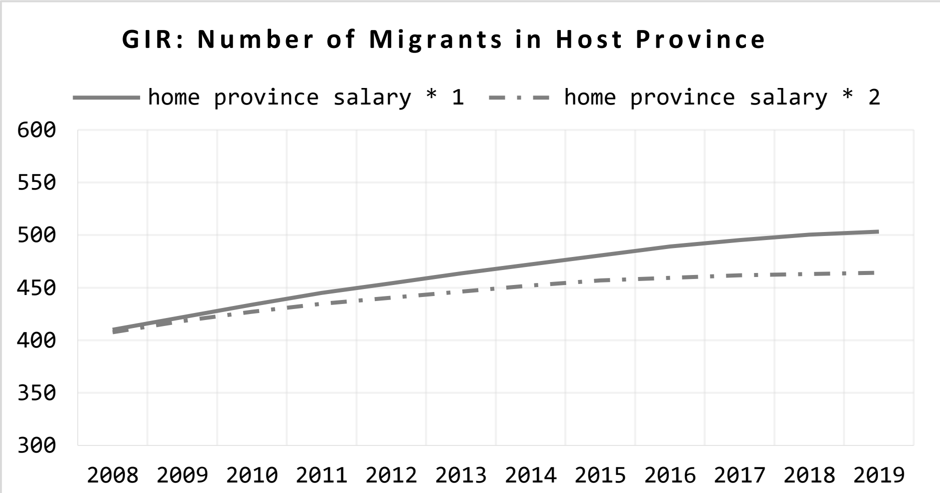

Figure 10. Average migrant stock per year of 100 model runs at different values of home salary - Girona

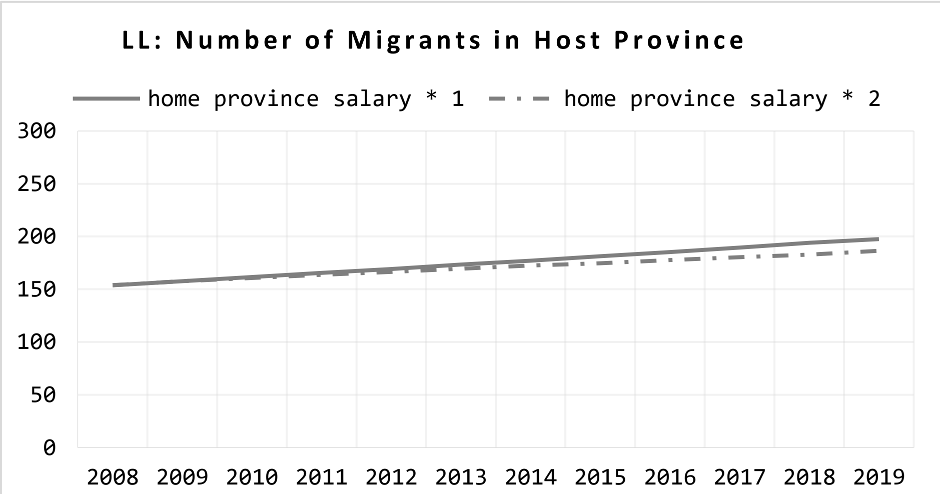

Figure 11. Average migrant stock per year of 100 model runs at different values of home salary - Lleida

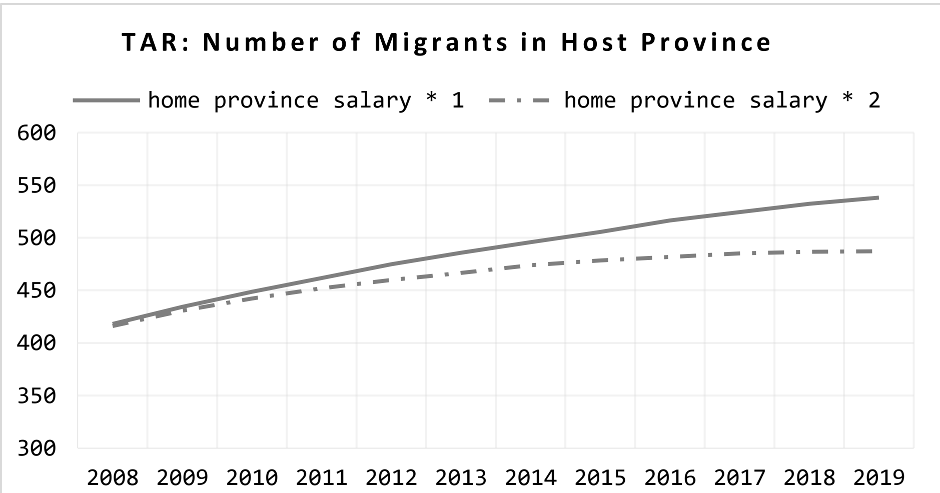

Figure 12. Average migrant stock per year of 100 model runs at different values of home salary - Tarragona

Series 2

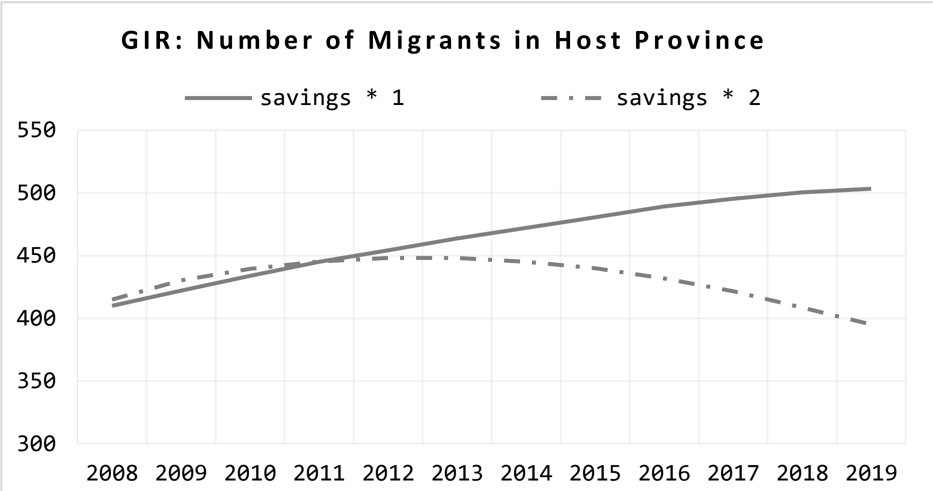

Figure 13. Average migrant stock per year of 100 model runs at different values of savings - Girona

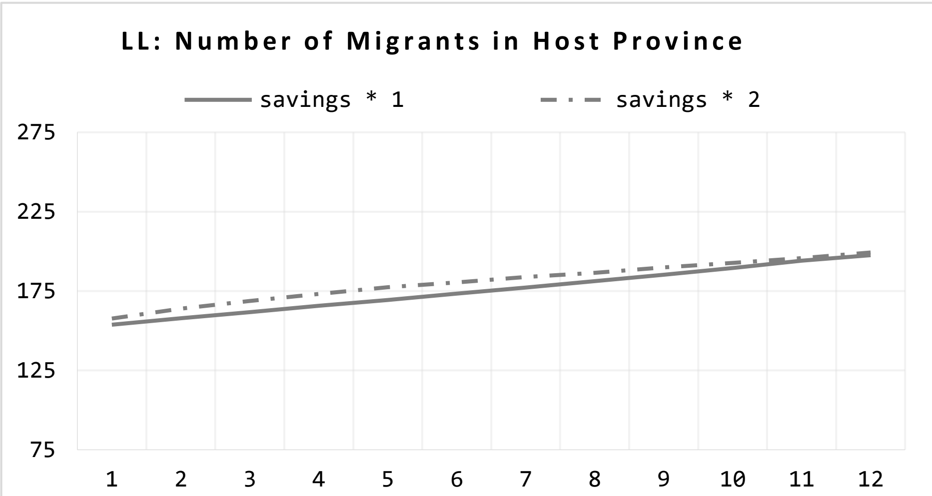

Figure 14. Average migrant stock per year of 100 model runs at different values of savings - Lleida

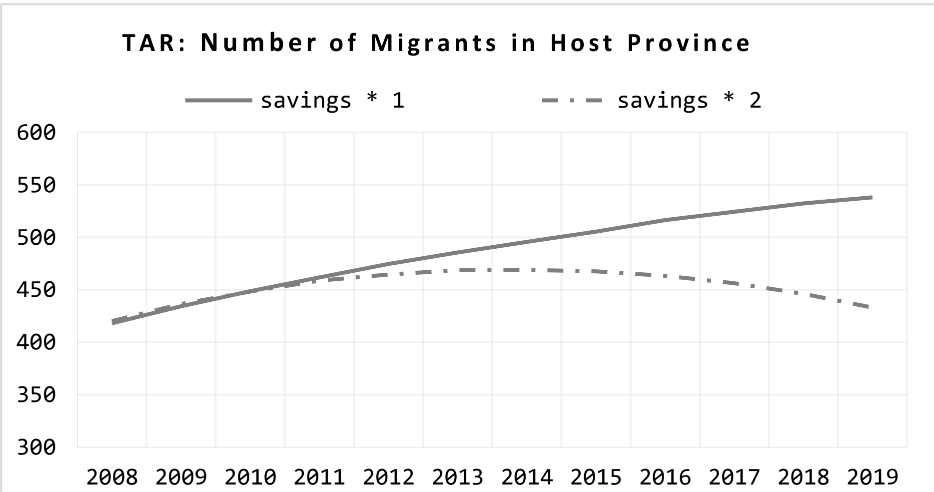

Figure 15. Average migrant stock per year of 100 model runs at different values of savings - Tarragona

7. Policy Suggestions and Conclusions

This work highlights that it is possible to simulate circular migrations and that using ABM can generate predictions for returning migrants.

It is fair to infer that since Barcelona is the host and that, according to our calculated empirical return data, not all migrants end up staying, this phenomenon could be partially due to seasonal labor migrations.

A weakness of our NetLogo model is the lack of data before the year 2008. Additionally, it could be possible to fine-tune the model if more data on the migrant's education, age, and migration history are accessible. This process could be easier for government departments to handle discreetly.

For future studies, models should be at the municipality or county levels. One interesting thing about the model is its capacity to make predictions for the future. For this, it is necessary to calibrate the model with empirical data. Next, predictions must be made of the data for the corresponding steps, be these in years, trimesters, or months. Then, the empirical and predicted data must be put together as input for the model after calibrating it with empirical data. This process makes it possible to make predictions within the model. A robust model like this could help mold policy in real life to treat the main reasons for interprovincial migrations in the future.

Advanced Agent-Based Modeling has the potential to become a cogwheel in taking the required steps to tackle the phenomenon of the empty Spain of the future. The resulting policies can be many, such as subsidizing the living costs for a few years. The process can let more people gather additional savings to establish well in their origin homes. Another policy could be to strategically incentivize new investment projects in lesser population-dense areas, which allows the people in less populated areas to remain there thanks to new jobs created via investment in said areas. This process is feasible assuming that the projects pay well by adjusting to the living costs, education, and skill profiles in said areas, making good use of local talent.

Appendix

Values of data-derived parameters and code

Slider Values	Host (BCN)	Girona's Case	Lleida's Case	Tarragona's Case
Neighborhood size (applicable to host)	1,8	-	-	-
Number of red firms (at home)	-	276	202	367
Number of people	-	3152	1763	3341
Number of blue firms (at host)	-	33	15	50
Moving costs	-	1843	1797	1916
Percentage at home	-	87,5	91,5	88
Return costs	-	2106	2106	2106
Production parameter blue	-	0,5	0,5	0,5
Production parameter red	-	0,45	0,46	0,48
Par. Comparative earnings	-	1	1	1
Parameter homebias	-	1	1	1
Parameter other migrants	-	0,01	0,03	0,01
Parameter ties to home	-	0,08	0,1	0,02
Baseline return probability	-	0,1	0,18	0,1
Baseline migration probability	-	0,03	0,04	0,02
Code Values				
Home preference	-	0,34; 0,51; 0,12; 0,027	0,34; 0,51; 0,12; 0,027	0,34; 0,51; 0,12; 0,027
Decrease in relationship capital	-	0,02	0,02	0,02
Minimum consumption	0,52	0,6	0,61	0,57
Standard deviation salary	0,32	0,37	0,38	0,35

Table 2. The table displays the corresponding slider and code values for each type of case study. Own source.

Input Figures with Barcelona as host and Girona as home:

Barcelona Nominal Salaries: 23391.80937 23505.3109 23644.16566 23701.74582 23834.78647 23850.86889 24163.93034 24828.92667 24163.47662 24670.32431 25621.30707 26550.26461 26720.75648

Barcelona CPI: 0.878426297 0.882435262 0.90040721 0.928996075 0.956269663 0.974746678 0.977314099 0.977293054 0.978997654 1 1.017950903 1.027462988 1.024222145

Girona

Distance: 0.282817 0.282817 0.282817 0.282817 0.282817 0.282817 0.282817 0.282817 0.282817 0.282817 0.282817 0.282817 0.282817

Nominal Salaries: 20442.58485 20508.99779 20637.46191 20676.45414 20733.34577 20764.54189 21074.81668 21640.06196 21057.32018 21513.78096 22331.56174 23095.42577 23282.7633

CPI: 0.89726186 0.892631645 0.913351854 0.946026434 0.970187734 0.982405186 0.983320705 0.977543461 0.976491139 1 1.019867828 1.02681315 1.018783937

Empirical Migration Rate (for reference): 0.009981757 0.010484588 0.010598868 0.010506934 0.01001724 0.010108313 0.010357032 0.009877069 0.008782636 0.008290705 0.008407174 0.008434538 0.006769115

Empirical Return Rate: 0.007018275 0.008444236 0.008036085 0.008764864 0.008760925 0.008682389 0.009046568 0.008446709 0.007368572 0.006634111 0.006135946 0.006248701 0.00239757

Stocks at host: 0.12425627 0.124440992 0.125100009 0.125057034 0.125415309 0.126539659 0.126528871 0.126577682 0.126197109 0.126015605 0.125779215 0.125046958 0.127689916

Savings Rate skew-normal distributions RStudio code

Using 2008-2020 annual data on savings rate we estimate the *location*, *scale*, and *skewness* parameters, per province, of a fitted skew-normal distribution.
We then use these estimations to generate the proper number of simulated observations.

We first load the 2008-2020 data

```
setwd("/home")
Savings_Provinces <- read.csv("savings_provinces.csv")
attach(savings_provinces)
```

Estimation

The estimation is done by maximum likelihood using the `fGarcg` package in *R*

```
library(fGarch)
parameters_GIR <- snormFit(GIR)
parameters_LL <- snormFit(LL)
parameters_TAR <- snormFit(TAR)
```

Simulation

Again, the `fGarcg` package is used to generate the simulated data

```
Savings_Rate_GIR <- rsnorm(3152, parameters_GIR$par[1],parameters_GIR$par[2],parameters_GIR$par[3])
Savings_Rate_LL <- rsnorm(1763, parameters_LL$par[1],parameters_LL$par[2],parameters_LL$par[3])
Savings_Rate_TAR <- rsnorm(3341, parameters_TAR$par[1],parameters_TAR$par[2],parameters_TAR$par[3])

write.csv(Savings_Rate_GIR, file ="skewnormal_GIR.csv")
write.csv(Savings_Rate_LL, file ="skewnormal_LL.csv")
write.csv(Savings_Rate_TAR, file ="skewnormal_TAR.csv")
```

NetLogo Model’s Example Code for the BCN (host) & GIR (home) Case

```
breed [workers worker]
breed [firms firm]

globals
[wall-open
profit-history
gdp-red
gdp-blue;
gdp-growth-blue
gdp-growth-red
skew-normal
datamigration
datareturn
realmigration
realreturn
BCNpriceindex
BCNpricenow
GIRpriceindex
GIRpricenow
distanceh
distancec
dataBCNsalary
dataGIRsalary
realBCNsalary
realGIRsalary
minimum-consumption-red
minimum-consumption-blue
stocksbluedata
realstocks
stockserror
mean-stockserror
flowserror
mean-flowserror]

patches-own
[residents
county]

firms-own
[employed-workers
counter
people-in-neighborhood
salary-here
random-prod]

workers-own
[employed
wealth
relatives
salary-earnings ;employed people receive salary-earnings
network-destination
home-preference
migrant
return-migrant
beforefirstmig ;indicates whether worker has migrated at least once
moving-probability
return-probability
hometown1 ;x-coordinate of hometown
hometown2 ;y-coordinate of hometown
move-list ;registers whether someone moves in a period
failed-migrant
potential-migrant
neighborlinks
clustering-coefficient
destination-lasttrip
already-counted
movelist-at-tick28]

undirected-link-breed [informations information ]

links-own
[life
 weight1
 weight2
 weighted-earnings
 ]

to setup
  __clear-all-and-reset-ticks
setup-wall
setup-people
setup-firms
setup-links
load-data
end

to setup-wall

ask patches with [pxcor = max-pxcor] [set pcolor yellow]
ask patches with [pxcor = min-pxcor] [set pcolor yellow]
ask patches with [pxcor = 0] [set pcolor yellow]
ask n-of (max-pycor * 2 * (1 - distancec)) patches with [pxcor = 0]
      [set pcolor pink]

end

to setup-people
ask n-of ((number-people * percentage-at-home) / 100) patches with [pxcor > 1 and pxcor < max-pxcor]
    [ sprout-workers 1 [ set color red
                         set shape "circle"
                         ]
    ]

ask n-of ((number-people * (100 - percentage-at-home)) / 100) patches with [pxcor < (- 1) and pxcor > min-pxcor + 1 ]
    [ sprout-workers 1 [ set color blue
                         set shape "circle"

                        ]
     ]

ask workers [set move-list []
             set beforefirstmig 1

             set already-counted 0]
ask workers [set home-preference 4]

ask n-of (number-people *  0.34) workers [set home-preference 0]
ask n-of (number-people *  0.51) workers [set home-preference 1]
ask n-of (number-people *  0.12) workers [set home-preference 2]
ask n-of (number-people *  0.027) workers [set home-preference 3]

ask workers with [color = red]

     [set relatives workers-on neighbors
      create-informations-with relatives

      set hometown1 xcor
      set hometown2 ycor]

ask workers with [color = blue]

     [set relatives one-of workers with [color = red ]

      set network-destination other workers-on patches in-radius neighborhood-size with [pxcor < 0]
      create-information-with relatives
      create-informations-with network-destination
      set hometown1 [xcor] of relatives
      set hometown2 [ycor] of relatives
     ]

end

to setup-firms
ask n-of number-firms-blue patches with [pxcor < 0 and pxcor > min-pxcor] [sprout-firms 1  ]
ask firms with [pxcor < 0] [set color green

                           set shape "square"
                           set size 1.5
                           set employed-workers 0
                           set random-prod random-normal 0 0.32
                           if random-prod < -0.99 [set random-prod -0.99]
                           ]

ask n-of number-firms-red patches with [pxcor > 0 and pxcor < max-pxcor] [sprout-firms 1  ]
ask firms with [pxcor > 0]  [set color yellow

                            set shape "square"
                            set size 1.5
                            set employed-workers 0
                            set random-prod random-normal 0 0.37
                            if random-prod < -0.99 [set random-prod -0.99]
                            ]

end

to setup-links
ask links [set life 1]

end

to load-data

ifelse (file-exists? "skewnormal.txt")
   [
     set skew-normal[]
     file-open "skewnormal.txt"
     while [not file-at-end?]
        [set skew-normal lput file-read skew-normal
        ]
    file-close
```

```
   ]
   [user-message "There is no skewnormal.txt file in current directory!"
   ]

ifelse (file-exists? "migration.txt")
   [
     set datamigration[]
     file-open "migration.txt"
     while [not file-at-end?]
        [set datamigration lput file-read datamigration
        ]
    file-close
   ]
   [user-message "There is no migration.txt file in current directory!"
   ]

ifelse (file-exists? "return.txt")
   [
     set datareturn []
     file-open "return.txt"
     while [not file-at-end?]
        [set datareturn lput file-read datareturn
        ]
    file-close
   ]
   [user-message "There is no return.txt file in current directory!"
   ]

ifelse (file-exists? "GIRsalary.txt")
   [
     set dataGIRsalary []
     file-open "GIRsalary.txt"
     while [not file-at-end?]
        [set dataGIRsalary lput file-read dataGIRsalary
        ]
    file-close
   ]
   [user-message "There is no GIRsalary.txt file in current directory!"
   ]

ifelse (file-exists? "BCNsalary.txt")
   [
     set dataBCNsalary []
     file-open "BCNsalary.txt"
     while [not file-at-end?]
        [set dataBCNsalary lput file-read dataBCNsalary
        ]
    file-close
   ]
   [user-message "There is no BCNsalary.txt file in current directory!"
   ]

ifelse (file-exists? "distance.txt")
   [
     set distanceh []
     file-open "distance.txt"
     while [not file-at-end?]
        [set distanceh lput file-read distanceh
        ]
    file-close
   ]
   [user-message "There is no distance.txt file in current directory!"
   ]

ifelse (file-exists? "BCNpriceindex.txt")
   [
     set BCNpriceindex []
     file-open "BCNpriceindex.txt"
     while [not file-at-end?]
        [set BCNpriceindex lput file-read BCNpriceindex
        ]
    file-close
   ]
   [user-message "There is no BCNpriceindex.txt file in current directory!"
   ]

ifelse (file-exists? "GIRpriceindex.txt")
   [
     set GIRpriceindex []
     file-open "GIRpriceindex.txt"
     while [not file-at-end?]
        [set GIRpriceindex lput file-read GIRpriceindex
        ]
    file-close
   ]
   [user-message "There is no GIRpriceindex.txt file in current directory!"
   ]

ifelse (file-exists? "stocksblue.txt")
   [
     set stocksbluedata []
     file-open "stocksblue.txt"
     while [not file-at-end?]
        [set stocksbluedata lput file-read stocksbluedata
        ]
    file-close
   ]
   [user-message "There is no stocksblue.txt file in current directory!"
   ]

set stockserror []
set flowserror []
end

to go

reset-values
consume
if ticks > 14 [migrate]
move

stop-at-walls

employ-people

update-salaries
if ticks > 14[return]
if ticks > 14 [update-data]

if ticks > 14 [do-plots]
tick
if ticks = 27 [ask workers with [color = blue]
                  [set destination-lasttrip round xcor * 100 +  round ycor]
             ask firms with [color = green and any? workers-here] [calculate-population]
              ask patches with [pxcor < 0 and any? workers-here ][people-here]
               ask workers [set movelist-at-tick28 length move-list]
               set mean-stockserror mean  stockserror
               set mean-flowserror mean flowserror ]
;
;                ]
if ticks = 28 [

               stop
               ]
end

to reset-values
ask workers
     [if color = blue
                [ create-informations-with other workers-on patches in-radius neighborhood-size with [pxcor < 0] [set life 1 ]
                  set network-destination information-neighbors with [pxcor < 0]

                 ]
     ]
ask workers [set moving-probability 0
             set return-probability 0
             if migrant = 2
                [set migrant 0]
             if migrant = 1
                   [set migrant 2]
            if return-migrant = 2
                [set return-migrant 0]
             if return-migrant = 1
                   [set return-migrant 2]
            if failed-migrant = 2
                 [set failed-migrant 0]
            if failed-migrant = 1
                 [set failed-migrant 2]
            if employed = 0 [set salary-earnings 0]
            ;set label length move-list
            set potential-migrant 0

           ]



ask links [set life life * (1.02) ;links get 2% older each year

           set weight1 0
           set weight2 0
           set weighted-earnings 0
           ]

ask workers [ifelse any? workers-on neighbors ;except links between close neighbors
                     [let closefriends workers-on neighbors
                       ask my-links with [other-end = one-of closefriends] [set life 1]
                     ]
                     [
                     ]
            ]

ifelse ticks > 14 [

set realBCNsalary item (ticks - 15) dataBCNsalary
set realGIRsalary item (ticks - 15) dataGIRsalary
set distancec item (ticks - 15) distanceh
set BCNpricenow item (ticks - 15) BCNpriceindex
set GIRpricenow item (ticks - 15) GIRpriceindex
```

```
set minimum-consumption-red realGIRsalary * 0.6
set minimum-consumption-blue realBCNsalary * 0.52
]
[set realBCNsalary item 0 dataBCNsalary
set realGIRsalary item 0 dataGIRsalary
set distancec item 0 distanceh
set BCNpricenow item 0 BCNpriceindex
set GIRpricenow item 0 GIRpriceindex
set minimum-consumption-red realGIRsalary * 0.6
set minimum-consumption-blue realBCNsalary * 0.52
]

ask firms [ifelse color = green
               [set salary-here realBCNsalary + random-prod * realBCNsalary
               ]
               [set salary-here realGIRsalary + random-prod * realGIRsalary
               ]
          ]

end

to consume

ask workers [let saving-rate item random 3152 skew-normal
             set wealth wealth + (saving-rate * salary-earnings)  ;no credit restrictions
             if salary-earnings = 0 ;if people have no salary-earnings they have to use up their wealth
                      [ifelse color = red
                        [set wealth wealth - minimum-consumption-red]
                        [set wealth wealth - minimum-consumption-blue]
                       ]
             ]
end

to migrate
compute-probability
make-decision

end
;
to compute-probability

ask workers with [color = red][if wealth >= moving-costs
                                     [set moving-probability baseline_migrationprobability

                                              if home-preference = 0 []
                                              if home-preference = 1 [set moving-probability moving-probability - parameter-homebias * 0.003]
                                              if home-preference = 2 [set moving-probability moving-probability - parameter-homebias * 0.005]
                                              if home-preference = 3 [set moving-probability moving-probability - parameter-homebias * 0.007]
                                              if home-preference = 4 [set moving-probability moving-probability - parameter-homebias * 0.003]

                           ifelse any? information-neighbors with [color = blue]
                                   [
                                     let comparative-earnings (sum [salary-earnings] of information-neighbors with [color = blue] ) / count information-neighbors with [color = blue]
                                      ifelse salary-earnings < comparative-earnings
                                             [
                                              ifelse beforefirstmig = 1
                                                          [set moving-probability moving-probability + 0.0000011 * (comparative-earnings - salary-earnings) * parameter-comparative-earnings
                                                          ]
                                                          [set moving-probability moving-probability + 0.0000035 * (comparative-earnings - salary-earnings) * parameter-comparative-earnings
                                                          ]
                                              ask my-links
                                                   [if [color] of other-end = blue
                                                            [
                                                             ask myself [set moving-probability moving-probability + parameter-other-migrants ]
                                                            ]
                                                   ]
                                             ]
                                             [set moving-probability 0
                                             ]
                                   ]
                                    [
                                   ]
                            if moving-probability >= 1 [set moving-probability 1]
                            ]
                        ]
end
;
to make-decision
ask workers with [color = red]
    [if (random-float 1 < moving-probability)
            [set potential-migrant 1
             set wealth wealth - (moving-costs * GIRpricenow)
             go-to-distance
             ]
    ]
end
;
to go-to-distance
ifelse random-float 1 < distancec

                [set failed-migrant 1
                ]
                [set migrant 1
                  set beforefirstmig 0
                  set color blue
                  set employed 0
                ;now: transform links to relatives into links that can decay
                  set salary-earnings 0

                decide-destination
                ]

end

to decide-destination ; people go either to where they know someone, or to a random patch in the left country.

ifelse any? information-neighbors with [color = blue]
                     [let informants information-neighbors with [color = blue]
                      move-to one-of informants with-max [salary-earnings]
                      if any? firms-here [set destination-lasttrip round xcor * 100 +  round ycor ]
                     ]
                     [move-to one-of patches with [pxcor < 0]
                     ]
set move-list lput 1 move-list

end
;
to move
ask workers
[if color = blue and employed = 0
    [ifelse any? information-neighbors with [employed = 1 and xcor < 0]  ;network helps migrants to find a job
         [move-to one-of information-neighbors with [employed = 1 and xcor < 0] with-max [salary-earnings]
         set destination-lasttrip round xcor * 100 +  round ycor
         ]
         [
         rt random 40
         lt random 40
         fd 1
         ]
    ]
 if color = red and employed = 0
   [ifelse any? information-neighbors with [employed = 1 and xcor > 0]
     [move-to one-of information-neighbors with [employed = 1 and xcor > 0]

     ]
     [rt random 40
       lt random 40
       fd 1
     ]
   ]
]
end
;
to stop-at-walls
ask workers [ifelse patch-ahead 1 = nobody
                     [rt 180]

                     [ if pcolor != black
                                [set heading (- heading)
                                         ifelse patch-ahead 1 = nobody
                                         [rt 180 fd 1
                                         ]
                                         [fd 1
                                         ]
                                ]
                       ifelse patch-ahead 1 = nobody
                       [rt 180 fd 1
                       ]
                       [if [pcolor] of patch-ahead 1 != black

                                [set heading (- heading)
                                        ifelse patch-ahead 1 = nobody
                                         [rt 180 fd 1
                                         ]
                                         [fd 1
                                         ]
                                   ]
                            ]
                       ]
             ]
end

to update-firms

end

to employ-people
ask firms [ifelse color = green
               [ifelse any? workers-here with [employed = 0]
                  [while [any? workers-here with [employed = 0]]
                    [set employed-workers employed-workers + 1
                      ask one-of workers-here with [employed = 0]
                      [set employed 1
                       set salary-earnings [salary-here] of myself
                       set destination-lasttrip round xcor * 100 +  round ycor
                       ]
                    ]
                  ]
                  [                  ]
```

```
              ]
              [ifelse any? workers-here with [employed = 0]
                [
                  while [any? workers-here with [employed = 0]]

                   [set employed-workers employed-workers + 1
                     ask one-of workers-here with [employed = 0]
                     [set employed 1
                      set salary-earnings [salary-here] of myself
                      ]

                   ]
                ]
                [
                ]

              ]
          ]
end
;
to update-salaries
ask firms  [if any? workers-here
                [ask workers-here with [employed = 1] [set salary-earnings [salary-here] of myself]]
            ]
end

to return
compute-return-probability
return-home

end

to compute-return-probability

ask workers [if color = blue and wealth > return-costs
                  [set return-probability baseline_returnprobability

                   if any? information-neighbors with [color = red]
                           [ask my-links with [[color] of other-end = red ] [let indicator 1 / life
                                                          ask myself [set return-probability return-probability + parameter-ties-to-home * indicator]
                                                           ]
                           ]
                  if return-probability > 1 [set return-probability 1]
                  ]
             ]
end

to return-home
ask workers with [color = blue]
    [if (random-float 1 < return-probability)
            [set return-migrant 1
             set wealth wealth - (return-costs * BCNpricenow) ;pay for return
            set color red
            set employed 0
            set salary-earnings 0
            set xcor hometown1 set ycor hometown2
            create-informations-with workers-on neighbors [set life 1];create new stable links to remaining relatives
            set move-list lput 1 move-list
            ]
    ]
end

to update-data

set realmigration item (ticks - 15)  datamigration
set realreturn item (ticks - 15 ) datareturn

set realstocks item ( ticks - 15) stocksbluedata

if ticks = 16 or ticks = 20 or ticks = 24
[set flowserror lput ((realreturn - ( count workers with [return-migrant = 1] / number-people)) ^ 2 + (realmigration - ( count workers with [migrant = 1] / number-people)) ^ 2) flowserror
  set stockserror lput (((count workers with [color = blue] / count workers ) - realstocks ) ^ 2) stockserror
]
end

to do-plots

set-current-plot "stocks-of-people" ;counts people in both countries
set-current-plot-pen "number-migrants"
plot count workers with [color = blue]
set-current-plot-pen "number-home"
plot count workers with [color = red]
set-current-plot-pen "realstocks"
plot realstocks * count workers
;;
set-current-plot "migrant-flows"
set-current-plot-pen "migrants"
plot count workers with [migrant = 1] / number-people
set-current-plot-pen "data-migrants"
plot realmigration
;;
;;
set-current-plot "return-flows"
set-current-plot-pen "return-migrants"
plot count workers with [return-migrant = 1] / number-people
set-current-plot-pen "data-return-migrants"
plot realreturn
;;
;;
end

to calculate-population ; firm procedure
;
set residents workers with [destination-lasttrip = round [pxcor] of one-of patches in-radius neighborhood-size * 100 +  round [pycor] of one-of patches in-radius neighborhood-size]
set people-in-neighborhood count residents with [already-counted = 0]
ask residents with [already-counted = 0][set already-counted 1]
end

to people-here;patch-procedure
;
let mayberesidents workers with [destination-lasttrip = round [pxcor] of one-of other patches in-radius neighborhood-size * 100 +  round [pycor] of one-of other patches in-radius neighborhood-size or
                                 destination-lasttrip = round [pxcor] of myself * 100 +  round [pycor] of myself]
set residents count mayberesidents with [already-counted = 0]
ask mayberesidents with [already-counted = 0][set already-counted 1]
;
end
```

Bibliography

Agent-Based Model. (2022). Retrieved from Wikipedia.

Bandrés, E., & Azón, V. (2021). *La despoblación de la España interior*. Funcas.

Demography. (2022). Retrieved from Wikipedia.

Haas, H., & Fokkema, T. (2011). The effects of integration and transnational ties on international return migration intentions. *Demographic Research 25*, 755-782.

Human Migration. (2022). Retrieved from National Geographic.

Klabunde, A. (2016). A model of circular migration (version 1.2.0). *A model of circular migration" (Version 1.2.0). CoMSES Computational Model Library. Retrieved from: https://www.comses.net/codebases/3893/releases/1.2.0/.* Retrieved from A model of circular migration (version 1.2.0).

Klabunde, A. (2018, February). *Computational Economic Modeling of Migration*. Retrieved from Oxford Handbooks Online.

Kladunde, A. (2014, 02 23). *Computational Economic Modeling of Migration*. Retrieved from SSRN: https://ssrn.com/abstract=2470525

Migration. (2022). Retrieved from United Nations.

(2018). *Radiography of the Housing Market 2017 - 18*. Fotocasa.

Santillana, I. (1981). The Economic Determinants of Internal Migration in Spain, 1960-1973. *Economics Notebooks: Spanish Journal of Economics and Finance*, 381-407.

Yaukey, D., Anderton, D. L., & Lundquist, J. H. (2015). *Demography: The Study of Human Population, 4th Edition*. USA: Waveland Press, Inc.