Omar Achour
Magnolia Metrica Ltd
31/12/2024

On Informative Priors and Distribution Shapes in Networks:

Incentive models and shapes of temporal curves for migratory patterns of all pairs of countries

Abstract:

In this paper, we introduce a better approach to gravity model used to predict migrant count in fully connected networks of all countries. We then proceed to stratify the pairwise temporal trends, taking advantage of their low span, into the most probable curves and proceed to derive the subsequent part of the curves. This paper serves as feed for a following estimation using Monte-Carlo methods in the upcoming paper.

5000 words

Keywords: Network, Gravity Model, Incentives, Temporal curves, network strata

Software: RStudio, R (programming language)

Table of Contents

Table of Figures

I. Introduction

Ever since humans have started to settle, at the very core of all events in known history is the movement of people from one place to another. Starting from settlements along the main rivers of the known world, Mesopotamia, Egypt, China and the Indus Valley, dweller from all parts would come for an opportunity to attain a betterment of the self, a sense of belonging to an orderly society and attempt possibilities to contribute to it. Generations later, many of these civilisations, due to drought or war or other factors, end up exporting humans to other centres bringing with them what has been the product of that part of civilisation into another. Organisation, technology and the rule of law would enable settlements to grow into empires and empires to collapse into others, the only common theme here is the flow of people from one place to another in one time or another.

In recent times however, economic migration has also presented us with, perhaps, more sophisticated reasons why humans migrate, for it is not for lack of food or pottery but a sort of arbitrage of wages, disposable income and costs of living. Migration of unskilled workers to Europe in the 1950-1990s was an invitation to fill labour shortages in recovering European economies post world war. Mass migration to the Americas in the late 1800s was an opportunity to settle “vacant” land, fertile and water and resource rich lands, and so the population followed. In both cases, and in many other cases, in fact most, the trade-off is a higher reward for those willing to take a leap of faith into the unknown especially when home countries are ill-run impoverished or war-torn.

The door has to be open in the first place for the flow to occur. Otherwise, we are faced with illegal settlements, wage cuts, poor conditions and a push to the margins. Phenomena we see today.

In the case of lawful migration, incentives and disincentives play a major role in the decision process, factors like distance, languages and free movement agreements are highly insightful of the propensity to migrate from one country to another, whereas economic conditions in both the origin and destination countries seem to fall off statistically.

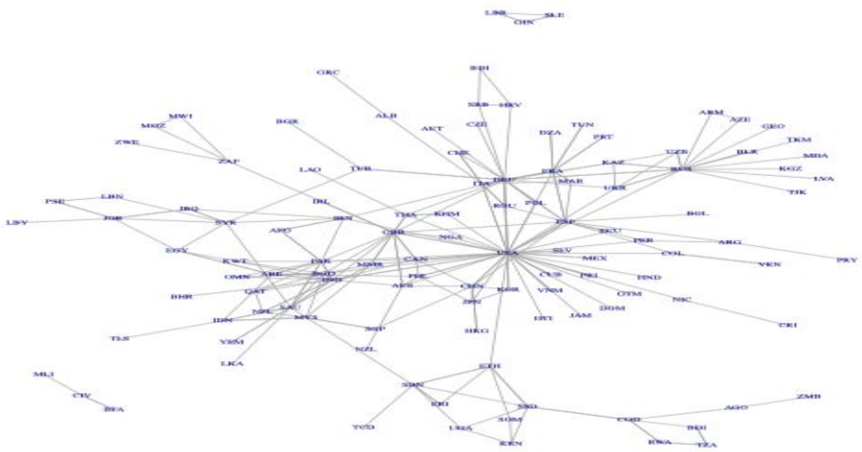
Perhaps, migration is not so purely driven by an economic engine looking to capitalise on wages, or its existence is largely shadowed by the other factors to such an extent that the findings only reflect the desire to belong in one society and to leave another and that people sometimes look for familiarity, in language for example, to geographic proximity, in distance, and to existing migratory routes canonised by free movement agreements.

The models however, in the frequentist sense do a poor job reflecting any temporal trends, since for the 40,000 pairs only a proportion can produce reliable estimates. All in all, a pairwise intercept is fitted to a large proportion of the pairs, reflecting a simple hovering around a mean. Therefore, the frequentist approach to this problem, from which we will only take these means would be used in a Bayesian setting in the next paper.

II. Data

The Data is collected from the United Nations World Population Projection, migrant count values are assembled into matrix form in the work of Azoze and Raftery (2017) and used as a reliable source to record our targeted variable.

To solve the problem, we collect data countries population totals, gross national income, nominal and PPP, we produce per capita estimates of income and then use a pairwise difference to create an economic incentive variable between average-earning citizen of every pair of countries. Throughout the process, compatibility between the set of countries from one dataset to another is corrected for. Anecdotally, the UN and the WB use different official names for some countries.

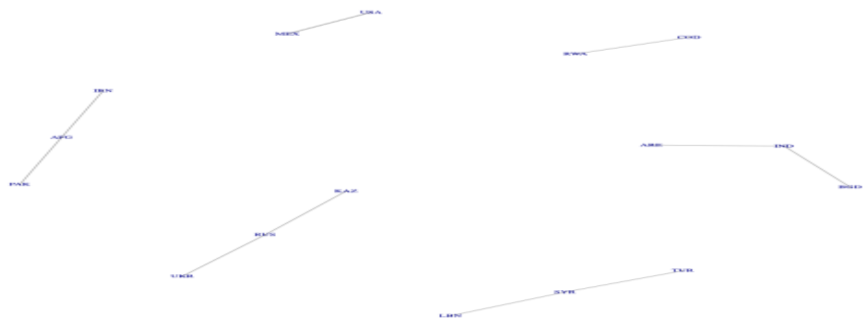


Source: Author's work, RStudio

Figure 1: Bidirectional Network of Migration Flows over 100,000 people.

From an early inspection of the data, well documented migration patterns appear on the graph with attractive centres in the United States, Russia, Germany, France, the UAE and so on. Relationships are easier to detect with flows exceeding a certain threshold and one can obviously see colonial relationships, the likes of Spain with Latin America, Russian with Central Asia and the Caucasus, France with North Africa, Britain with South Asia. Regional centres like Ivory Coast and Ethiopia attract neighbouring countries. Countries like Turkey and Britain, with extensive presence over multiple continents exhibit properties of importers as well as exporters of migrants.

While an interesting shift from one coloniser to another country that speaks the same language as the coloniser occurs often, Germany for example receives migrants from what was Habsburg Austrian territory, previously Ottoman Turkish, the United States receives migrants from the ex-British realm and so on. In these cases, the destination language has been introduced or at least cultural familiarity is present. The USSR is also a stark example, as a recipient of Turkic Central Asians citizen.



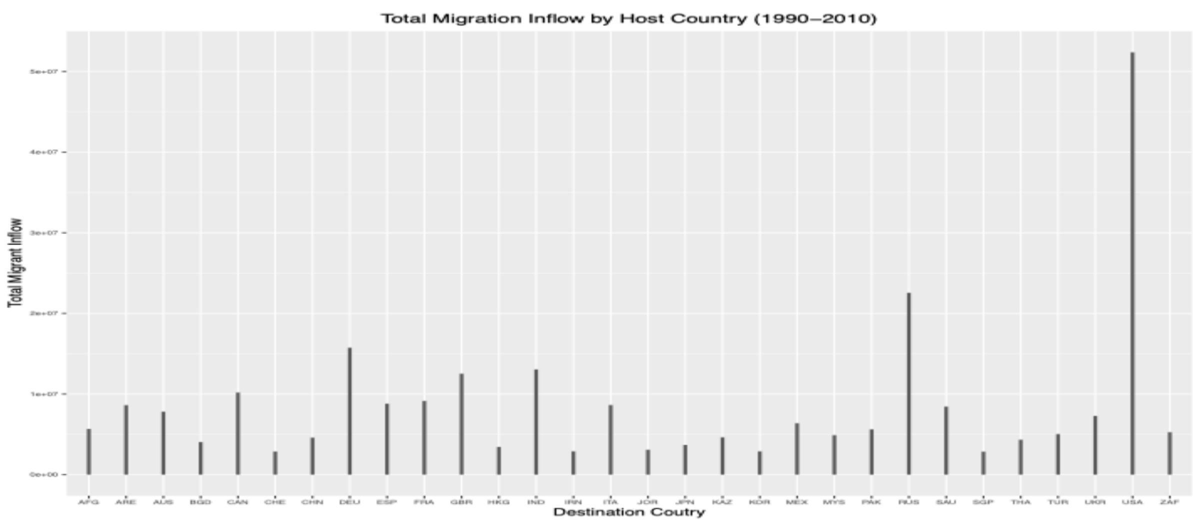
Source: Author's work, , RStudio

Figure 2: Bidirectional Network of Migration Flows over one million people.

Iran, Afghanistan and part Pakistan share multiple languages, in their respective regions, with Farsi being spoken in Afghanistan as Dari, Pashto in both Afghanistan and Pakistan and Balochi in all three countries. As well as the historical significance of the Persian language in modern Urdu among other factors cooccurs with large flows between these countries.

On the other hand, the continental proximity of Latin American countries produces large migrant stocks in the United States, in spite of the attainable language barrier. In also give rise to Spanish speaking communities forming in large urban centres and the southwestern US along the border. Bearing in mind that the Southwest was under Spanish and then Mexican hands one can see again the historical significance mixed with language that fuels destination choice.

The Gulf Region also emerges as a hub for migration from Arab countries as well as other Asian communities from India and the Philippines, the mix of Arabic and the widespread use of English seems to be the formula to attract labour from within the region and beyond to fill in shortages. With South Asia, being fairly close geographically the transition is highly optimised on both ends.

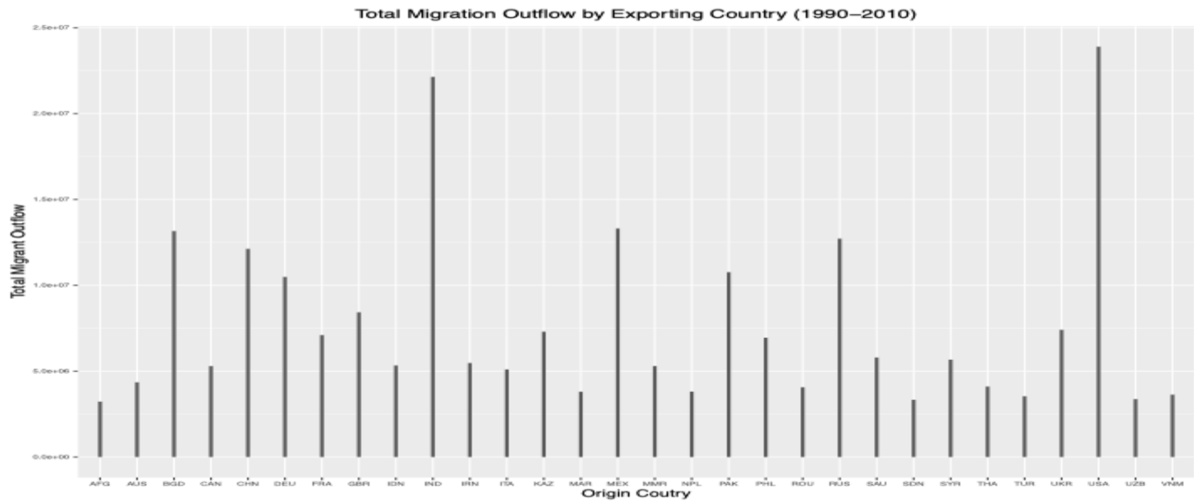


Source: Author's work, RStudio

Figure 3: Total Migration Inflow by destination during 1990-2010

Due to the nature of the migration count variable, being counted as the number of citizens of one country being recorded as residents in another for a minimum period during a specific time, many migrants end up returning in the next period to their home countries and are therefore considered exported back by their destination country to their country of origin.

Migration outflow is also intrinsically linked to population count and thus large countries in terms of population do export large numbers of migrants. As seen below, India, China, the USA, Pakistan, Bangladesh, Russia, Mexico each have exported over 10 million migrants during the 30-year period. In our analysis, we might investigate using a total population denominator to account for it.



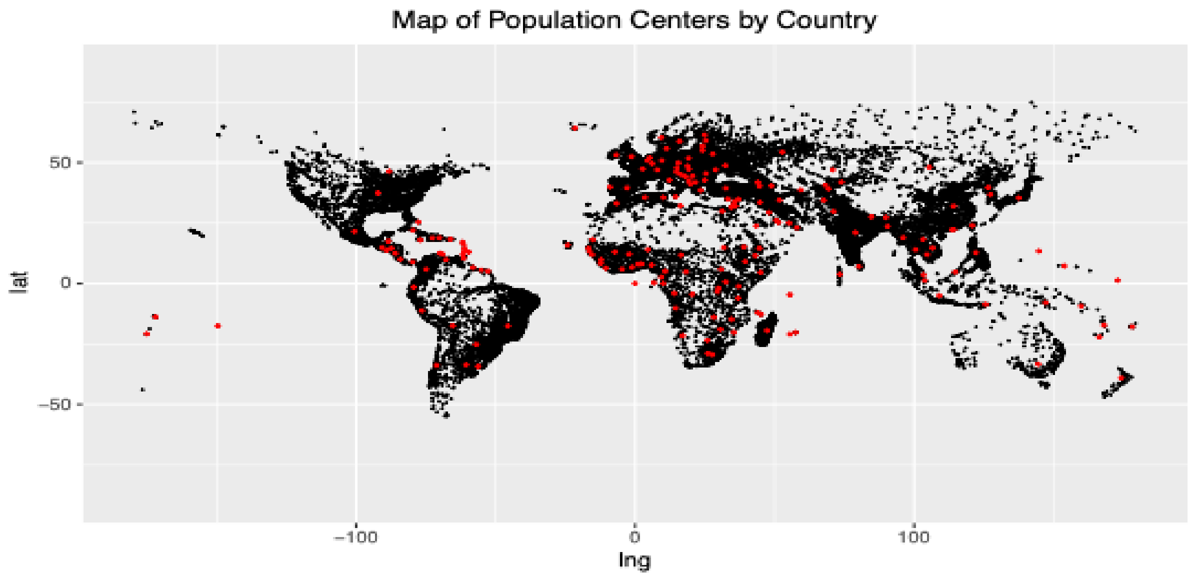
Source: Author's work, RStudio

Figure 4: Total Migration Outflow by destination during 1990-2010

We then look at disincentives from the perspective of migrants the first of which is distance, although airplane flights have partly solved the problem for the modern human costs are still involved in the number of hours, instead of days, the number of stops, visa-requirements and flight ticket costs. The notion distance is therefore, only subliminally embodied in physical distance, but it is nonetheless a factor to consider, especially when involving concepts such as familiarity, culture and migratory precedence.

We therefore produce a dataset of population centres for each country using a weighted mean of all cities coordinates by their population as weights. We then compute a pairwise Haversine distance between population centres of all pairs of countries.

Needless to say, that proximity between Western European countries as well as the Levant coincides with large bilateral flows between each country. Whereas countries that land share borders on a geographic map but are separated by large natural barriers such as deserts, in the case of North and Sub-Saharan African, mountain ranges in the case of India and China and rainforests in the case of Brazil with Western South America creates striking migratory barriers further separated by language and culture.



Source: Author's work, RStudio
Figure 5: Population Centres by country (in red)

Natural barriers such as the Sahara Desert, the Himalayas, the Amazon Rainforest, the Australian Desert the Siberian Tundra, the Central Asian Steppe the Rocky Mountains and the Canadian shield create very distinct cultural, linguistic and often economical concentrations of populations sometimes irrespective of political borders.

Fertile lands and plentiful water resources are to this day the magnet for population concentration and the creation of abundance of resources necessary for human sustenance and prosperity. It is almost as if one can see every major civilisation that has walked this earth in one picture.

We then produce datasets of dummy variables for the presence of a common language where a language is grouped into a set of countries where it is official, co-official, a minority language of widely used. This variable is used as an exponent with various non-linear transformations.

The choice of countries where a country is a minority language is fairly subjective as German settlements for example exist throughout eastern Europe scattered in small towns, as well as Greek settlements in the Middle East and the Black Sea. Italy would therefore have German as a minority language in the northern region along the Austrian border but also Greek in the southeast of the country, both languages have been present of centuries albeit at a smaller scale nowadays. The same goes for Kazakh, Kyrgyz and Tajik (Farsi) in Western China, where these communities speak these languages as their native tongue. Another example would be Caucasus region in Russia and Western Iran sharing many minorities in both countries. Kurdish is spoken natively in four highly settled countries but only recently is considered official in Iraqi Kurdistan.

If for example a Kurdish speaker from Western Iran moves to the Kurdish speaking Southeast of Turkey a common language would have been an incentive, obviously along with distance and others, whether he spoke Turkish or Farsi would not have reflected the choice incentive and therefore including minority languages does make sense as the official language is often not as absolutely widespread at least in daily life. This subjective choice of variable can bring to light deeper connections between countries as well as very improbable ones.

The dummy variable for free movement agreements between countries follows the same algorithm as that of common languages minus the stigma, where common trade blocks are grouped together then pairwise dummies are created. Major blocks include the EU, Mercosur, the Caribbean Community, the Gulf Cooperation Council.

Data is then collected in the end using an R array of 200x200x6 cubic matrices for each pair during the 6 time periods for each independent variable. We then proceed to add certain derived variables like total migrant inflow, outflow and define our target variable as migrant choice, being the proportion of migrants originating from one country choosing a certain destination country. Outflow variable would add up to one for origin countries and the same for inflow for destination counties.

We wish to also examine whether migrant stock from one origin could be an incentive for more migrants from the same country as it might constitute precedence. As well as generic gravity models computed as single variables to reduce dimensionality. These models will be discussed in more detail in the next section.

The data presents us with a challenge as in only spans 6 periods (1990, 1995, 2000, 2005, 2010, 2015) and is reduced to 5 periods when using pairwise constructs.

III. Literature Review

This paper, in its conception draws inspiration from the works of Welch and Raftery (2022) among others that sought to estimate migration flows between pairs of countries. Building up on previous works in the field, they derived Bayesian hierarchical model to make probabilistic projections for bilateral migration flows among 200 countries, excluding territories and dependencies. The model is fit to 5-year intervals between 1990 and 2010, which we preserve in this paper. The Bayesian approach outperforms the international migration literature and is considered as a new norm when approaching this problem. The Bayesian flow model outperforms, the Poisson hurdle, Gravity models and Autoregressive persistence models and mean flows discussed below.

Guy Abel (2011) proposed a model to estimate flows of migrants between European countries then extended into a global setting in 2013 using the place of birth of migrants as the origin. In parallel with Bijak and Wisniowski (2013), migration count was estimated assuming a classic Poisson distribution appropriate for count data. They also clarified terminology regarding what constitutes a migrant and which migrants need be included, long-term migrants that is. Migrants were legal resident of foreign birthplace living for a period exceeding 12-months in a given country, this persists in this work.

Azoze and Raftery (2015) propose an approach based on Gravity models, reappropriated from international trade literature, itself from the physical sciences, their work involves estimating migration rates rather than counts and involves demographic indicators like age structures and sex ratios. Their migrant count matrix derived from UNWPP was used in this paper. The hierarchical model with an autoregressive component with normal priors on the adjusting multiplicative factor and noise factor as well as a uniform prior on the mean, the following levels follow the classic inverse gamma for deviations and non-informative priors on its parameters.

Almquist et al. (2019) dissected closely-connected migration networks by using an information theory-based community detection algorithm resulting in 17 migration networks, of which North and Central America as well as Japan, Korean, the Philippines and Vietnam for example constitute one, the South Asian subcontinent with Arabia, the Mediterranean region including Southern Europe and North Africa but also South America, the Levant and Egypt, the Persosphere, Russosphere and Sinosphere among other. The algorithm was initially developed by Rosvall & Bergstrom in 2008.

The Bayesian flow model proposed by Welch and Raftery (2022) however has a different hierarchy for migrant count as it relies at the top on a multinomial distribution with 200 destinations (events), this distribution is specified using two parameter the number of migrants departing the country of origin during the period and a conditional probability vector of those migrants arriving in the indexed destinations, this parameter resembles our target variable of migrant choice albeit in logs. Although not used in this paper, we review this model at the end of the next section.

IV. Model

IV. I. Simple Linear and Logit models

As mentioned earlier, the pairwise data spans 5 periods so the model should be designed to only include at most one or two variables with an intercept. Having gone through the variables earlier it is therefore unpractical to fit the dependant variable to free variables and we would need a model that both encompasses all information available to us but is not allowed to have its own regression factor.

As for the dependant variable, since it adds up to one for each country and exhibit the properties of a probability, that of destination choice, it would be appropriate to use log probabilities as a proxy of choice. We would be looking to build a logit model of migrant destination choice.

The emphasis would shift to examine possible constructs that would encompass all information. We look into log-additive models, which would translate into a probability depending on a multiplicative construct of multiple variables. Where incentives and costs could be both grouped together and their effects separated by a fraction line. A common model in the literature is the gravity model.

IV. II. Gravity models

Gravity models have made their way into economic sciences via the physical sciences, the idea is that economies much like physical objects, would attract each other by a force of trade dependant on the scale of both economies and their distance from each other. In that literature, the volume of trade is thought to be proportional to the respective economies, in GDP, and inversely to the exponent of trade barriers (or the opposite of trade agreements).

In migration modelling, a reasonable extrapolation would be to assess bilateral migration flow using a combination of origin migrant population, destination migrant stock and physical distance as the basic components of a standard gravity model and then develop from there.

The initial attempts at this model tried to include migrant stock for the destination country j and migrant population of origin i as the multiplicative masses in the numerator, with the physical distance between the two countries population centres:

$$log(P_{i,j,t}) = \alpha_{i,j} + \beta.\frac{S_{j,t}.M_{i,t}}{D_{i,j}} + \epsilon_{i,j,t}$$

With $P_{i,j,t}$ representing the probability that a migrant from country i choses country j as a destination in time t.

$$P_{i,j,t} = \frac{M_{i,j,t}}{\sum_{j=1}^{200} M_{i,j,t}}$$

In the subsequent models we have opted to use log probabilities by entering logs on the whole equation. Adding total population to the fraction has not yielded desirable results and was omitted. A logit model would look something like this, however since we are not assigning to a class of choice probabilities but rather determining a fraction of choices out of a total the logit model approach although possible is discarder for the log probabilities. The accentuations resulting from log odds would not serve the purpose of our problem.

$$\log\left(\frac{P_{i,j,t}}{1 - P_{i,j,t}}\right) = \alpha_{i,j} + \beta.\frac{S_{j,t}.M_{i,t}}{D_{i,j}} + \epsilon_{i,j,t}$$

The first error of the model would be that $M_{i,t}$ is not independent, as a matter of facts it is a constituent of the fraction that is $P_{i,j,t}$. So is the case with migrant stock $S_{j,t}$. We would therefore have to replace population data with another. An approach would be to use economic mass rather than population with Gross National Income (GNI) or GNI per purchasing power parity (GNI PPP)

$$\log\left(P_{i,j,t}\right) = \alpha_{i,j} + \beta.\frac{E_{j,t}.{E_{i,t}}^{-1}}{D_{i,j}} + \epsilon_{i,j,t}$$

With $E_{k,t}$ representing the economic conditions of the country of origin and destination. A large destination economy, coupled with a smaller origin economy could trigger a migration decision but a more accurate approach is the use per capita GNI PPP. This way, a migrant from a low income, per capita, country would be economically incentivised to move across a certain distance to earn a higher income for the same work produced, this would be a form of arbitrage had barriers to entry been abolished. The arbitrage decision is better understood using a difference of average income between the two countries. The higher, the positive, difference is the more the agent stands to gain from his movement from i to j at time t.

$$\log\left(P_{i,j,t}\right) = \alpha_{i,j} + \beta.\frac{(E_{j,t} - E_{j,t})}{D_{i,j}} + \epsilon_{i,j,t}$$

It turns out however, that the data is fitted poorly using economic indicators, since income difference fits only 1 in 10 pairs of all countries.

As detailed in the discussion section, any economic and demographic indicators on their own seldom coincide with the temporal progression of migrants' choice. As gravity models are only multiplicative combinations of these, they also would not be able to completely trace out the progression in 40,000 pairs of countries.

As the movements from 1990 to 2015 are the result of multiple reasons and <u>not necessarily occur continuously throughout the period</u> whereas others increase or decrease monotonically and others fluctuate with multiple peaks and throughs. A brute force algorithm to determine the shape of each quintuple is used to determine the shape out of 16 possible combinations.

Subsequently, we seek to mix our previous indicators with new ones so as to improve the network's fit. We therefore look to enrich the denominator with exponential dummies.
This approach is out of scope of gravity models, we would call them network models of incentives, detailed in the next section.

IV. III. Synthetic models of incentives

Building our synthetic models of incentives of networks of all counties draws inspiration from the dynamic of gravity model. As we are deeply constrained with the number of independent variables, the model's idea to combine different indicators saves us regressors without compromising on information.
Another great attribute is that of its form, as incentives are both multiplicative and upward trending whereas disincentives in the denominator sink the values down as desired. Therefore, we will build on this structure without necessarily relying on a temporal element. What this would do is provide at least pairwise specific intercept, or mean, over which the 5 temporal occurrences hover, if correctly fitting, the temporal dynamics would be dealt with using a Bayesian approach in the next paper.

In other words, we build from this:

$$\log\left(P_{i,j,t}\right) = \Lambda_{i,j} + v(t) + \epsilon_{i,j,t}$$

Or,

$$\log\left(P_{i,j,t}\right) = \alpha_{i,j} * \frac{\exp\left(L_{i,j}\right)^{a} . \exp\left(F_{i,j}\right)^{b}}{D_{i,j}} + v(t) + \epsilon_{i,j,t}$$

With ${L_{i,j}}^{a}$ represents a dummy variable for the presence of a common language between the origin and the destination. ${F_{i,j}}^{b}$ represents a dummy for the presence of a free movement agreement and $D_{i,j}$ represents Haversine distance in Kilometres. While $v(t)$ potentially represent either or both components of an Autoregressive moving average, although neither of which fit over a tenth of the given pairs in practice.

$$v(t) = \theta_{AR} \sum_{m} \log\left(P_{i,j,t-m}\right) + \theta_{MA} \frac{1}{m} * \sum_{m} \log\left(P_{i,j,t-m}\right)$$

The dummies are used inside of and exponential, resulting in either e or 1, and then raised to a common power, 2 or 3, chosen based on the best fit. The square exponentiation would reflect a higher level of relationship between countries, be it a common language or common market, and therefore a higher scale of migration flow. This would also be reflected on the log scale by an increment of almost one for the decimal basis when one of the exponents is non-negative and close to two if both are raised to the power.

This approach has a certain intuition that is understandable, since countries with high synergies would be more inclined to admit and send immigrants to and from. This baseline would represent a certain level or scale of migration between pairs of countries that fits almost all the pairs as detailed in the next section.

As for the temporal parameter, no indicator, combination of indicators, or non-linear transformation of it using the classical economic and demographic trends fits more than 1 in 6 pairs of the model and therefore approaching a solution for this part of the problem using the frequentist paradigm seems erroneous. We focus on the pairwise intercepts for the rest of the paper as they would serve as informative priors when building a Bayesian model, the specifications of which, for the sake of exemplification resemble the following model mentioned in the literature review.

IV. IV. Intermezzo: Bayesian Flow Model

A hierarchy is derived for each of the two parameters representing outflow and inflow. The former factorised as the product of the origin's population and migration rate, the rate is log-normal with its mean being a linear combination, given a uniform prior, of its long-term migration rate, with normal prior and normal-beta sub-priors, and its previous period lag. The latter as the fraction of a pairwise exponentiated spatial interaction term over its sum for all destinations, in turn given a normal prior with normal-beta sub-priors. Both normal and beta subpriors are non-informative.

Expectation:

$$E(m_{i,j,t}|\, \pi_{i,j,t}, \delta_{i,t}) = \pi_{i,j,t}\delta_{i,t}N_{i,t}$$

Observations:

$$m_{i,.,t}|\, \pi_{i,.,t}, \delta_{i,t} \sim Multinomial(N_{i,t}, \pi_{i,.,t})$$

Outflows:

$$N_{i,t} = \left\lfloor \delta_{i,t}P_{i,t} + \frac{1}{2} \right\rfloor$$

$$\log(\delta_{i,t}) \sim Normal([(1-\phi)\mu_i + \phi\log(\delta_{i,t-1})], {\sigma_i}^2)$$

$$\phi \sim Uniform(0,1)$$

$$\mu_i \sim Normal(\nu, \tau_0^2)$$

$$\nu \sim Normal(\mu_0, 100^2)$$

$$\sigma_i \sim Beta(\alpha_{10}, \beta_{10})$$

Inflows:

$$\pi_{i,j,t} = \exp(\eta_{i,j,t}) / \sum_{j \neq i} \exp(\eta_{i,j,t})$$

$$\eta_{i,j,t} \sim Normal(\kappa_{i,j}, \psi_{i,j}^2)$$

$$\kappa_{i,j} \sim Normal(0, 10^2)$$

$$\psi_{i,j} \sim Beta(\alpha_{20}, \beta_{20})$$

V. Findings

V.I. Time-invariant Model of incentives

The pairwise specific model of mean deviation plus noise looks as follows:

$$\log{(P_{i,j,t})} = \alpha_{i,j} * \frac{\exp{(L_{i,j})^a}.\exp{(F_{i,j})^b}}{D_{i,j}} + \epsilon_{i,j,t}$$

With $P_{i,j,t}$ representing probabilities, $\alpha_{i,j}$ a scaling parameter, $\epsilon_{i,j,t}$ the noise parameter. The incentives are listed as $L_{i,j}$ for language, $F_{i,j}$for free movement and $D_{i,j}$ for spherical distance as the disincentive.
After iterating over integer combinations, a = b =2 as powers is chosen and justified using the property mentioned earlier.

We look at the fit for all pairs and compare with the other models discussed before:

	SMI I	**SMI II**	**Gravity I**	**Gravity II**	**AR (1)**	**MA (1)**	**Mean**	*
Intercept (Λ)	0.48	0.53	0.53	0.53	0.41	0.57		**SMI II**
Variable I	-	-	0.16	0.14	0.07	0.10	-	-
Variable II	-	-	-	-	(0.02)	(0.04)	-	-

Source: Author's work, RStudio

Figure 6: Models' fits as a fraction of all pairs of countries (40,000)

As mentioned earlier, the gravity model with economic indicator variables lacks any substance with regards to its temporal trend as well as the autoregressive processes, moving average and the combination of the two.

With temporal variables fitting less than 4000 pairs out of 40,000 and performing worse when combined as an ARIMA. While using a mean would still work in form, its correlation with the data corrupts any predictive power to the model and would therefore give rise to autocorrelation in the models.

Therefore, keeping a pairwise intercept, that we have made ourselves instead of a regression intercept adds comprehensibility to the model as well as incorporates certain information surely consequential in the decision process.

This would leave us with the time-invariant model of incentives alone fitting half of all pairs. Noting that using squared exponents in both the language and movement dummies improves the fit to the network.

The underlying model give rise to speculation about the true behaviour of network flows for each pair with time, whether it exists or not and whether the pairs are so heterogenous temporally that no one model is prevalent.

V.II. Pairwise temporal trends in the network

We therefore turn to real analysis to figure out the temporal trend of each pair. Making use of a simple brute force algorithm we sweep through the data and record any peaks, throughs and descents and ascensions in migrant count.

The low number of observations, once a disadvantage when adding regressors, is now facilitating the computation of our curve forms since using five data points, creates four ascending or descending lines, the possible combinations of the four lines results in 16 shapes of 40,000 pairs. Classifying the data into each allows us to resume our digging into the behaviour of the data.

Strictly ascending '/' '/' '/' '/' 1524* 3.81%, 10.98%	Double ascending peak '/' '/' '/' '\' 1100 2.75%, 7.92%	Ascending peak through '/' '/' '\' '/' 886 2.21%, 6.38%	Ascending peak descending '/' '/' '\' '\' 674 1.68%, 4.86%
Peak through ascending '/' '\' '/' '/' 1587* 3.97%, 11.44%	Peak through peak '/' '\' '/' '\' 1009 2.52%, 7.27%	Peak descending through '/' '\' '\' '/' 1019 2.54%, 7.34%	Peak double descending '/' '\' '\' '\' 320 0.8%, 2.3%
Through double ascending '\' '/' '/' '/' 770 1.92%, 5.55%	**Trough ascending peak** '\' '/' '/' '\' 1687* 4.21%, 12.16%	Trough peak through '\' '/' '\' '/' 847 2.12%, 6.10%	Trough peak descending '\' '/' '\' '\' 849 2.12%, 6.12%
Descending through ascending '\' '\' '/' '/' 542 1.35%, 3.90%	Descending through peak '\' '\' '/' '\' 448 1.12%, 3.23%	Double descending through '\' '\' '\' '/' 400 1%, 2.89%	Strictly descending '\' '\' '\' '\' 213 0.53%, 1.53%

Source: Author's work, RStudio

Figure 7: Temporal behaviour of migrant count for all responding pairs of countries

We are able to produce a sort of heat diagram with a certain ordering of the data, the first line is ascending in the first 8 then iterated, the second through the first four, the third through the first two and the fourth is iterated every form. Noting that only 34.69% of the data is stratified using this method, the rest either contain at least one missing variable or some sort of misfit of the condition or the form (i.e. characters).

If we are to consider the sampling independently distributed and that third of the pairs that fits the conditions is randomly picked, we can start to compare the occurrence of certain shapes and their frequency throughout the sampled pairs.

At a first glance, the possible shapes seem fairly well distributed with hardly any one being present more than four times the other. This even distribution still shows certain properties regarding the overall behaviour of the count data.

First, descending behaviour is less often present than ascending behaviour, this could be as a result of the overall population growth and the parallel need to invite labour and migrate to work.

Second, the presence of throughs in the second and third lines, third and fourth period minima, reflect a change in immigration policy fuelled by pushback from working class citizen in the early 2000s.

Third, a high presence of ascending curves marked by single peaks, double peaks and through or peak and through. Reflecting, the pairwise specific policies regarding immigration for certain countries, late responders and early movers into the imposition of restrictions throughout the late 1990s and early 2000s.

	Ascending	Descending	
First leg	58.51%*	41.49%	**1.41:1**
Second leg	60.08%*	39.92%	**1.5:1**
Third leg	62.46%*	37.54%	**1.66:1**
Fourth leg	54.59%*	45.41%	**1.20:1**

Source: Author's work, RStudio

Figure 8a: Temporal behaviour of migrant count for responding sample of pairs

The table above shows a penchant towards ascending behaviour on a 3:2 scale. Although not absolute, data definitely drifts more upwards than downwards, more precisely for every five changes in trend three are upward and two downward.

	Ascending	Descending	
First and second leg	30.14%*	11.55 %	**2.6:1**
Second and third leg	36.61%*	14.06%	**2.6:1**
Third and fourth leg	31.82%*	14.81%	**2.15:1**
Three legs {1-3; 2-4}	13.47%*	5.19%	**2.6:1**
All legs	10.98%*	1.53%	**7.17:1**

Source: Author's work, RStudio

Figure 8b: Temporal behaviour of consecutive legs for responding sample of pairs

The contrast is starker in the case of successive legs where strict ascendance occurs between one point and its two successors accounting for at least of third of the sampled data and outnumbering descending trends 5:2. The ratio persists considering three legs at a time. Strictly ascending trends are highly present outnumbering strictly descending ones by seven-fold.

The analysis can also be extended to the occurrence of peaks and throughs by combining two legs at a time, three points, the results produced are even more pleasant.

	Peak	Through	
First (1-2)	28.35%	29.93 %*	**1:1**
Second (2-3)	23.46%	24.84%*	**1:1**
Third (3-4)	30.58%*	22.71%	**1.34:1**

Source: Author's work, RStudio

Figure 8c: Temporal behaviour of consecutive legs for responding sample of pairs

Peaks and throughs occur just as frequently in the beginning and middle of the curve while peaks are more frequent towards the end. This reflects a higher entropy with regards to assigning a probability when restricted only to figure 8c, but combining it with the lines from figure 8b, we are able to produce a fully encompassing look at the behaviour of the data:

	Ascending	Descending	Peak	Through	
First (1-2)	30.14%*	11.55 %	28.35%	29.93 %	**3:1:3:3**
Second (2-3)	36.61%*	14.06%	23.46%	24.84%	**3:1:2:2**
Third (3-4)	31.82%*	14.81%	30.58%*	22.71%	**3:1:3:2**

Source: Author's work, RStudio

Figure 8d: Temporal behaviour of consecutive legs for responding sample of pairs

The data, in its first three instances, is just as likely to produce an ascending order as it is to produce a peak or a through, and three times less likely to produce a descending order. In the middle instances, it is more likely to produce and ascending curve that a peak or a through by a factor of 3:2, and over 2.5 times less likely to produce a descending trend. Towards the end of the quintuplet, data is just as likely to produce an ascending tail or a peak, with the other third of possible outcomes is shared amongst a descending tail and a through and the latter outnumbering the former 2:1.

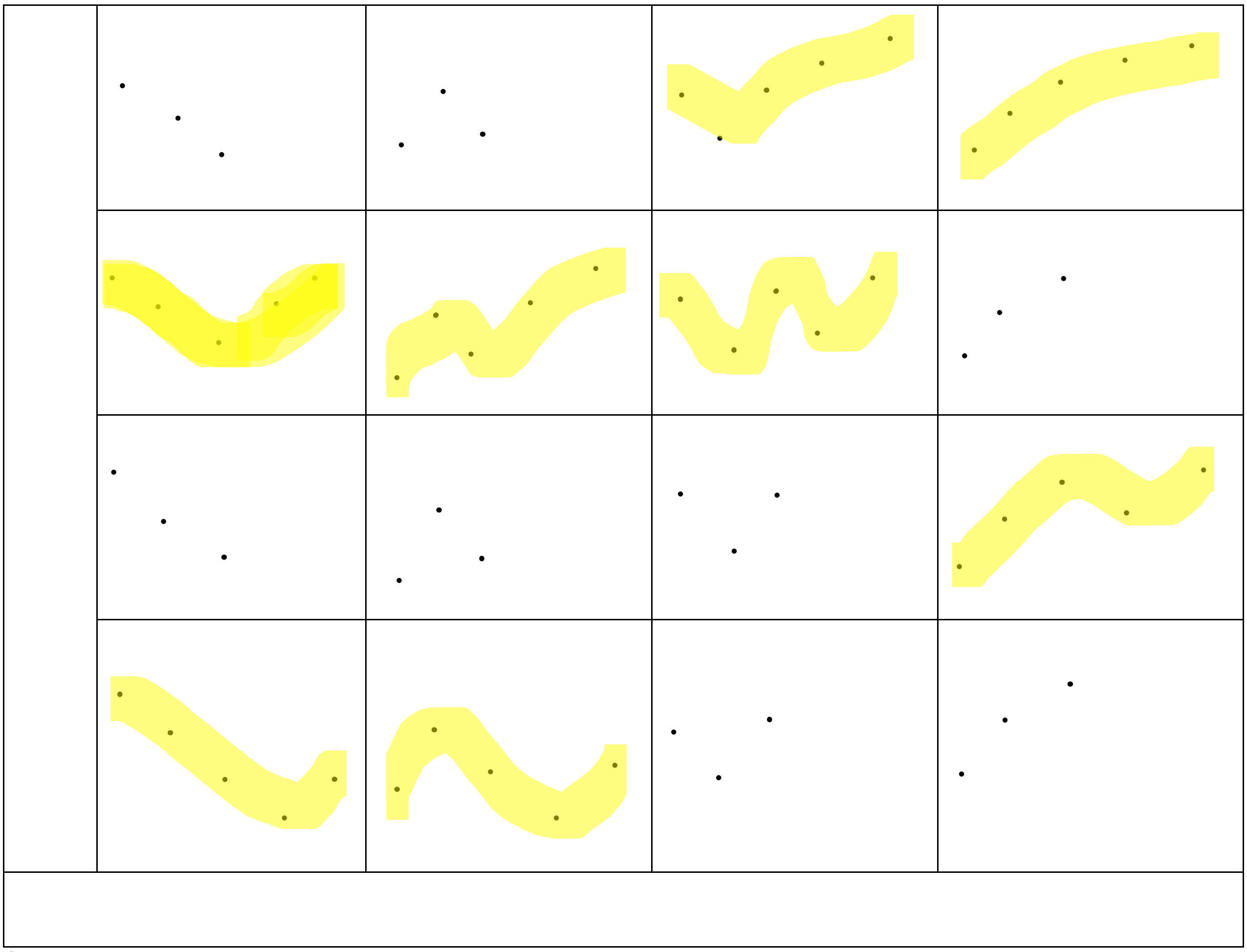
By drawing these figures, we can assign to their distribution's probabilities of occurrence as well as produce mixtures and hierarchies using Monte-Carlo methods.

Double peaks and double throughs for example could signify the presence of a mixture of Gaussians. Peaks in the middle would resemble a simple normal curve. Whereas, peaks or throughs, followed or preceded by a strict order could resemble a more tilted curve similar to a Gamma, a t-distribution, a beta distribution of a combination involving others.

We would look at identifying these properties as well making use of them in the next. Paper.

Using only strata from figure 8 we are able to build 8 legally most likely shapes, by the most probable first legs, middle and tail, noting however that the other 8 combinations are rendered illegal, in this fremework, by the shape of the first three points, i.e. combining a peak start with an ascending middle or other peak and vice-versa, or a through with a descending middle or other through and vice-versa.

We are left with the following curves:



Source: Author's work

Figure 9: Most likely temporal behaviour strictly considering figure 7 stratifications

The elimination of half the curves does not render them completely improbable, the point of this process is only to draw the most likely next part of the curve irrespective of transcendent correlations, but respective of the feasibility constraints.

We can compare the shapes produced here with the most likely shapes in figure 7, the strictly ascending curve and the peak-through-ascending do make it but the through-ascending-peak does not. We could obviously add them by hand, see appendix.

VI. Conclusion

Estimating migration count for a fully connected network of all countries, presented us first with a challenge limiting the number of regressors, which was solved by employing gravity models. The gravity models however, despite multiple combinations and transformations only poorly fitted the dependant trend. We therefore introduced a new set of synthetic intercepts sensitive to similarities between countries like language and free flow as well as the spherical distance between population centres of the two countries. The pairwise-specific intercept picks us well and forms a constant with noise model that we try to aliment with ARIMA(m) lags and moving averages, without substantial improvements we only keep the constant with noise model.

The second part of the paper looks at improving our knowledge about the various temporal trends occurring independently in each pair, this leads us to employ a simple algorithm to stratify the time curve into 16 different possible shapes and then rank the most likely succession of three points, forming either an ascending or descending line, a peak or a through then proceeding to derive and rank the most likely middle shape and tail based on the initial form and a feasibility constraint. This approach allowed us to derive most of the likely shapes we had statistically counted earlier.

This paper forms a great base for a Monte-Carlo simulation using mixtures or hierarchical models employing as priors our constant derived in the first part. While, the choice of distributions would largely rely on the derived curves from the second part.

VII. References

- Abel, Guy, et al. *The IMEM Model for Estimating International Migration between Countries in Europe*. 2011.
- Abel, Guy J. "Estimating Global Migration Flow Tables Using Place of Birth Data." *Demographic Research*, vol. 28, 15 Mar. 2013, pp. 505–546, https://doi.org/10.4054/demres.2013.28.18. Accessed 19 Sept. 2020.
- Abel, Guy J. "The Form and Evolution of International Migration Networks, 1990–2015." *Population, Space and Place*, vol. 27, no. 3, 16 Feb. 2021, https://doi.org/10.1002/psp.2432. Accessed 9 Dec. 2022.
- Abel, Guy J., and Joel E. Cohen. "Bilateral International Migration Flow Estimates for 200 Countries." *Scientific Data*, vol. 6, no. 1, 17 June 2019, https://doi.org/10.1038/s41597-019-0089-3.
- Allenby, Greg M., et al. "Hierarchical Bayes Models: A Practitioners Guide." *SSRN Electronic Journal*, 2005, https://doi.org/10.2139/ssrn.655541.
- Azose, Jonathan J. , and Adrian E. Raftery. *Estimation of Emigration, Return Migration, and Transit Migration between All Pairs of Countries*. Dec. 2018, doi.org/10.1073/pnas.1722334116.
- Azose, Jonathan J., and Adrian E. Raftery. "Bayesian Probabilistic Projection of International Migration." *Demography*, vol. 52, no. 5, 10 Sept. 2015, pp. 1627–1650, https://doi.org/10.1007/s13524-015-0415-0. Accessed 1 Jan. 2025.
- Azose, Jonathan J., and Adrian E. Raftery. "Estimation of Emigration, Return Migration, and Transit Migration between All Pairs of Countries." *Proceedings of the National Academy of Sciences*, vol. 116, no. 1, 2 Jan. 2019, pp. 116–122, www.pnas.org/content/116/1/116, https://doi.org/10.1073/pnas.1722334116.
- Bijak, Jakub, and Arkadiusz Wisniowski. *Statistical Modelling of International Migration Flows*. 22 Aug. 2011.
- Raymer, James. "Measuring Flows of International Migration." *IZA World of Labor*, 2017, https://doi.org/10.15185/izawol.354. Accessed 17 June 2020.
- Tunaru, Radu. "Hierarchical Bayesian Models for Multiple Count Data." *AUSTRIAN JOURNAL of STATISTICS*, vol. 31, 2002, pp. 221–229. Accessed 1 Jan. 2025.
- Vargas, Mauricio. *A Crash Course on Gravity Models Basic Model*. 2023.
- Welch, Nathan G., and Adrian E. Raftery. "Extended Abstract: Modeling and Forecasting Bilateral Migration Flows for All Countries." *University of Washington Department of Statistics*, May 2021.
- Welch, Nathan G., and Adrian E. Raftery. "Probabilistic Forecasts of International Bilateral Migration Flows." *Demography Statistics*, 15 July 2022, doi.org/10.1073/pnas.2203822119.
- Wisniowski, Arkadiusz. *Bayesian Modelling of International Migration with Labour Force Survey Data*. 2011.