# The Effects of Information on the Formation of Migration Routes and the Dynamics of Migration

Martin Hinsch*
University of Southampton
Department of Social Statistics and Demography
University of Glasgow
MRC/CSO Social and Public Health Sciences Unit
hinsch.martin@gmail.com

Jakub Bijak
University of Southampton
Department of Social Statistics and Demography



**Abstract** Most models of migration simply assume that migrants somehow make their way from their point of origin to their chosen destination. We know, however, that—especially in the case of asylum migration—the migrant journey often is a hazardous, difficult process where migrants make decisions based on limited information and under severe material constraints. Here we investigate the dynamics of the migration journey itself using a spatially explicit, agent-based model. In particular we are interested in the effects of limited information and information exchange. We find that under limited information, migration routes generally become suboptimal, their stochasticity increases, and migrants arrive much less frequently at their preferred destination. Under specific circumstances, self-organised consensus routes emerge that are largely unpredictable. Limited information also strongly reduces the migrants' ability to react to changes in circumstances. We conclude, first, that information and information exchange is likely to have considerable effects on all aspects of migration and should thus be included in future modelling efforts and, second, that there are many questions in theoretical migration research that are likely to profit from the use of agent-based modelling techniques.

**Keywords**
Migration, communication, beliefs, migration routes, agent-based modelling

## 1 Introduction

International migration has important economic, humanitarian, and cultural consequences not only in countries of origin and destination but also in countries that lie on common migration routes (Castles et al., 2014). Nevertheless migration is to date one of the least well understood demographic processes (Bijak et al., 2021). The majority of older theoretical efforts to understand migration follow the economic tradition where migrants' behaviour is typically described as an optimisation process that weighs the costs of migration against a combination of push and pull factors in the countries of origin and destination, respectively (Greenwood, 2005). While some of these models have become quite sophisticated and have in some cases even been empirically validated, the approach has repeatedly been criticised for oversimplifying many aspects of the system (Klabunde & Willekens, 2016).

In particular, it is usually assumed that migrants' decisions follow a simple and rational process. Furthermore variation between individuals as well as interactions between them are usually not taken

* Corresponding author.

 

into account. Why these assumptions might limit the applicability of these models is demonstrated by empirical results that show that in many cases prospective as well as actual migrants are substantially misinformed concerning the conditions in the country of destination (Gilbert & Koser, 2006). It has also been found that connections to and opinions of a country within an individuals' social network can play an important role in the migration decision, thus making interactions between individuals relevant for the process (Sačer et al., 2017).

Some of these concerns have been adressed in newer modelling efforts, in particular those using agent-based modelling (Frydenlund & De Kock, 2020). By explicitly simulating single individuals, agent-based models make it straightforward to model variation and interactions within a population. Furthermore, since these models are usually computational there is no inherent limit to the complexity of behaviour that can be modelled (for an overview see Hinsch & Bijak, 2021).

An aspect of migration that has not received much attention amongst modellers, even in newer studies, is the migration journey itself. The main reason for this is probably that in most models of migration the focus lies on the decision to migrate and then on the choice of destination. Some predictive models tailored to a specific time and place explicitly include the migrants' travels (e.g., Frydenlund et al., 2018; Hébert et al., 2018; Suleimenova & Groen, 2020) but apart from our own earlier work (Hinsch & Bijak, 2019), we are not aware of any theoretical models that directly investigate or take into account individuals' movement. Migrants are instead assumed to make their way from origin to destination without further complication.

We know, however, that migrants' journeys are anything but simple, direct movements from a country of origin to a destination (Crawley et al., 2016; Kingsley, 2016). More importantly, the specificities of the journey might have consequences in other areas as well. They can be relevant in a practical context, as, for example, political as well as humanitarian reactions to migration depend on the timely localizing of migrants. In a theoretical context on the other hand they might affect our understanding of migration itself, as decisions made during travelling might have profound carryover effects on other aspects of migration such as choice of destination (Brekke & Brochmann, 2015). Furthermore the difficulty of the journey a migrant expects will change the perceived attractiveness of destinations and might therefore itself affect their choice of destination or even the decision to migrate in the first place (Bertoli & Fernández-Huertas Moraga, 2013).

While the effect of limited information about migrants has been considered at least in the economic literature (Katz & Stark, 1987), migrants themselves are usually assumed to be perfectly informed. Information can, however, be an important yet often scarce resource for migrants during their journey. Surveys of migrants show that knowledge about the destination and the ways to reach it is often limited and might come from unreliable sources (Borkert et al., 2018; Dekker et al., 2018; Gilbert & Koser, 2006). In some cases this information precarity is exacerbated by a general distrust towards information sources other than personal contacts (Emmer et al., 2016). If, however, migrants base their travel decisions on incomplete or erroneous information, it can be expected that they will experience difficulties on their journeys leading to delays, detours, or failure.

As we showed in an earlier theoretical simulation study, this scarcity of information and the way knowledge is obtained and exchanged can indeed strongly affect the development of migration routes. We found that under limited information, migration routes can become an emergent effect of the migrants' communication, which makes them unpredictable and leads to suboptimal travel (Hinsch & Bijak, 2019). This suggests that the assumption of a straightforward, successful migration journey might often be misleading.

Here we expand on this effort using an improved version of the model. Our aims in this are twofold. First we want to test the robustness of our previous results in a more general context and with a better model. Mainly, however, we are interested in how misleading we expect the assumption—as made in most migration models—of a simple journey with perfect information to be. Our question therefore is: How different are migration journeys under perfect information from those in a scenario with limited information? What might the consequences of these differences look like?

It is important to note that as with our previous study this is a purely theoretical work. We are not modelling a specific real-world situation but performing *computational sociology* (Macy & Willer, 2002) by attempting to understand the effect of certain assumptions on the behaviour of an entire class of systems.

## 2 Model Description

The model described below is a strongly modified version of a model we have presented before (Hinsch & Bijak, 2019). Along with many smaller modifications, we transitioned from stepwise updates to a continuous-time, event-based paradigm (with commensurate changes from probabilities to rates and updates to processes) and simplified the model by removing capital, resources, and the two-tier link system.

An earlier version of the model that the present study is based on was also used as a didactic running example in our book (Bijak et al., 2021).

Since a full description of the model would exceed the available space, we provide in the following only a brief overview. The source code and detailed documentation for the model can be accessed on Comses (https://www.comses.net/codebases/b57ed5e2-47cb-4b61-9d95-8785398c6c8d/releases/1.1.0/). Please note that a few of the mechanisms described in the full documentation (risk, resources, and capital) were not used in the current study and were therefore switched off in the simulation runs by setting the appropriate parameter values. A full list of model parameters including default values can be found in the appendices.

### 2.1 Overview

In our model a population of migrants travels from a location of origin to a destination, crossing a landscape of cities and transport links. Agents attempt to navigate this world optimally based on subjective knowledge that is not necessarily complete or correct. They gain additional knowledge through experience and by exchanging information with other agents.

### 2.2 Entities

The simulated world consists of *locations* ("cities") that are connected by *links* (see Figure 1). Cities and links are static entities with properties that do not change over the course of the simulation. Cities have a 2-dimensional position and a *quality* that determines their *attractiveness* to agents. Quality represents, for example, the (lack of) presence of police, the availability of resources, or the level of safety. Links connect two cities and have *friction* as their only property. Friction affects the time it takes for an agent to transverse the link and is determined by the link's length as well as a stochastic component.

Nearly the entire behaviour of the model consists of the *actions* of agents or their interactions with each other or the world (see below). Agents are at all times positioned either in a city or on a link unless they have arrived at their destination. Agents have some amount of *information* about the world (see below) as well as a number of *contacts* among the population of travelling or arrived agents.

### 2.3 World

The simulated world is constructed as a random geometric graph (Gilbert, 1961) of 600 cities connected by transport links. Cities have a random quality $q \sim \mathcal{U}_{[0,1]}$. The positions of cities are distributed uniformly on a unit square. Any two cities that are closer than a threshold distance are connected by a transport link. In addition one departure location at $x = 0, y = 0.5$ and 10 exit

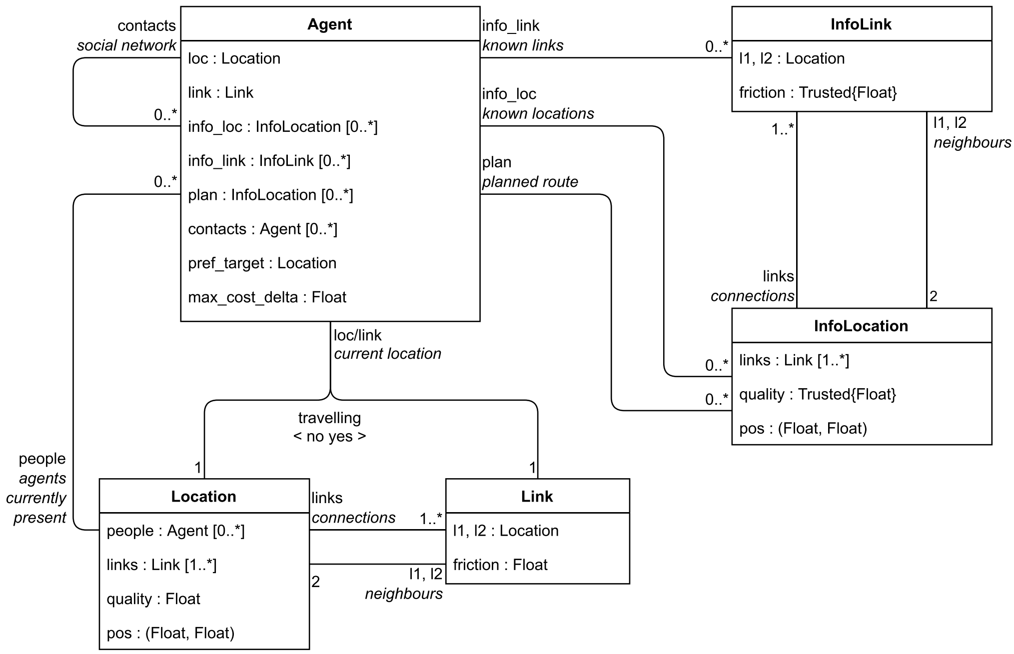


Figure 1. Diagram of the entities in the model and their relationships.

locations placed in regular intervals at $x = 1$ are added to the world. Departure and destination locations are connected by links to the five closest cities, respectively.

Links' only property is *friction*, which is calculated from distance $d$ as $f_i = d_i r$ with random $r \sim \mathcal{U}_{[0.75,1.25]}$.

## 2.4 Actions and Interactions

All events in the model are assumed to be Poisson processes in continuous time. With the exception of the creation and departure of new agents, all changes of model state are the result of the action of an agent. Which actions an agent can perform and their rates of occurrence depend on the agent's state, in particular on whether it is currently travelling on a link or staying in a city.

**create agents** Agents are created with a fixed time-dependent rate. They enter the world at the departure location. Unless noted otherwise, agents start out without contacts and without any knowledge.

**plan** During planning, an agent either plans a route to an exit or, if it does not have sufficient knowledge, decides which neighbouring city to go next.

**explore** An exploring agent gains new knowledge about closeby cities and links.

**add contact** An agent adds agents that are currently situated in the same city to its list of contacts.

**forget contact** An agent unilaterally forgets a randomly selected contact.

**exchange information** An agent communicates with one of its contacts and exchanges information about the world topology, i.e., the existence and connectedness of cities and links, as well as their properties.

**depart** An agent departs from its current location and starts travelling to the next location in its plan.

**arrive** A travelling agent finishes traversing a link.

## 2.5 Information

We are interested in how reliance on and exchange of possibly incomplete or wrong information affect the agents' decision-making. Therefore we decided to explicitly model the agents' knowledge of the world as well as the information exchange between agents. The submodel on information exchange presented in this section is largely identical to earlier versions published elsewhere (Bijak et al., 2021; Hinsch & Bijak, 2019).

An agent's knowledge is comprised of a number of information items each of which represents a city or a link. Topologically this information is accurate—all connections an agent knows about are correct—but not necessarily complete—an agent may know only a small number of cities and links. Information items have the same properties as the real-world entities they represent; however, their values may be inaccurate.

To model this, the real values of properties are in their subjective counterpart replaced by an estimate of the value together with a certainty that the value is correct. Agents can gain information either directly from the world by "exploration" (action 'explore') or by communicating with other agents (action 'exchange information'). As explained in the following, both processes can add new information items and update an estimate as well as the certainty of an information item's properties.

If agents encounter unknown (to them) cities or links (through exploration or communication), they add a new information item corresponding to that entity to their knowledge, setting property estimates to a default value and certainty to 0. When exploring a known entity, values are updated, with the new value being a weighted mean between the previous estimate or certainty and the real value (or 1 in case of certainty).

Information exchange between agents is more complicated as it needs to exhibit a number of specific properties: If two interacting agents have similar estimates for a property their corresponding certainty should increase. If, on the other hand, their estimates differ, both individuals should decrease their certainty. At the same time, an agent should always adjust its estimate in the direction of that of its interaction partner; however, it should do so in proportion to its relative certainty. That is, in an exchange between an agent with high certainty and one with low certainty, the agent with the low certainty should change its estimate more.

While there is a substantial theoretical literature on belief and opinion dynamics, previous models seem to focus largely either on adversarial exchange of opinions, i.e., situations where individuals attempt to convince each other, or on situations where individuals change their beliefs according to social norms or consensus (e.g., Duggins, 2017). An interesting approach by Martins (2009) and extended by (among others) Adams et al. (2021) uses Bayesian inference to derive updating rules for beliefs about the value of continuous real-world variables. The resulting model is, however, computationally quite expensive. We therefore designed our own model of information exchange.

We based our information model on the well-known mass action dynamics (Horn & Jackson, 1972). To understand the model it is best to imagine that an agent's belief consists of two "substances", certainty and doubt, in proportion $t$ and $d = 1 - t$. When two agents interact a "reaction" between their respective belief components takes place, potentially transforming them: Doubt reacting with doubt produces doubt. Certainty of one agent interacting with the other agent's doubt can "convince" the latter, changing parts of its doubt into certainty. Depending on the difference in estimate, certainty interacting with certainty can lead to confusion and increased doubt or just change the estimate.

More formally, for an interaction between agents $A$ and $B$ with an estimate $v$ we define the difference in estimate as

$$\delta_v := \frac{|v_A - v_B|}{v_A + v_B}. \tag{1}$$

Using parameters $c_i$ ("convince"), $c_u$ ("confuse"), and $c_e$ ("convert"), we then calculate the new doubt value $d'_A$ based on the previous values of certainty $t.$ and doubt $d.$ as

$$d'_A = d_A d_B + (1 - c_i) d_A t_B + c_u t_A t_B \delta_v. \tag{2}$$

The estimate $v_A$ changes accordingly:

$$v'_A = \frac{t_A d_B v_A + c_i d_A t_B v_B + t_A t_B (1 - c_u \delta_v)((1 - c_e) v_A + c_e v_B)}{1 - d'_A} \tag{3}$$

It is important to note that this is a purely phenomenological model. It was chosen for being based on a well-known, simple formalism and showing all required properties, but does not claim to be psychologically or empirically accurate. As we can see, for the special case where different opinions do not lead to doubt, i.e., $c_u = 0$, doubt will disappear, i.e., $d$ will approach 0 (as long as $c_i > 0$), and the model reverts to a simple weighted mean (as in, for example, Nordio et al., 2018):

$$v'_A = (1 - c_e) v_A + c_e v_B \tag{4}$$

## 2.6 Decisions

Agents attempt to find the least costly route from their current position to an exit, based on their current knowledge. The cost of a route is a function of the links' friction and the quality of cities visited on the way. If they are not able to find a complete path they instead select the best city in the vicinity based on distance (friction), quality, and proximity to the destination.

## 2.7 Setup

We are investigating the effects of (limited) information and information exchange on the formation of migration routes. In order to obtain a baseline with which to compare our results, we first ran all scenarios under the assumption of perfect information. That is, agents received full and perfect knowledge about every link and city in the simulated world. In order to avoid any additional effects through communication errors we also switched off communication in these scenarios entirely (see Appendix 1).

To test the effects of information exchange we then ran the model under various levels of communication frequency and intensity (see Table A1 in Appendix 1). We also varied the strength of communication error and the fidelity of the information agents receive through exploration.

We explored further potential real-world consequences of information in additional scenarios where agents had a preference for a specific destination (scenario 'preferred destinations') or where after a certain amount of time some links became difficult to navigate (scenario 'intervention').

We ran 10 random replicates for each parameter combination. As preliminary runs showed that the simulation approaches equilibrium after 300–500 time units, we ran all simulations up to $t = 750$.

# 3 Results

We wanted to know whether discrete migration routes form in the first place and, if so, how predictable and optimal they are. For this we used three key measurements:

**route concentration** We calculate the relative standard deviation of transit counts across all links as a proxy for the degree to which travel routes are similar between agents.

**optimality** We determine the correlation coefficient between realised transit counts for all links and transit counts in a hypothetical scenario where each individual travelled optimally.

**unpredictability** The unpredictability of transits for a given city is measured as the standard deviation across all replicates of the proportion of transits for that city. We calculate overall unpredictability as average unpredictability of arrivals over all exits.

## 3.1 Baseline Scenario

If individuals are perfectly informed, every agent is able to find and travel on the optimal route, resulting in maximum route concentration and predictability (Figure 2). With imperfect or incomplete information agents do not necessarily know enough to find the objectively best route and will instead travel suboptimally (Figure 3). This leaves scope for variation between individuals as well as between replication runs (see Figure 4), therefore route concentration as well as route predictability are substantially lower in scenarios without perfect knowledge (Figure 2).

As we can see in Figure 2, however, for anything but perfect exploration, the unpredictability of agent arrivals decreases when changing from low to medium communication but increases again for high communication. Together with the increase in route concentration with communication, this indicates that what we observe is a phase transition between three regimes:

For low communication, agents receive only little input from each other. On the other hand exploration is not sufficient to produce a reliable map. Routes therefore differ between agents and

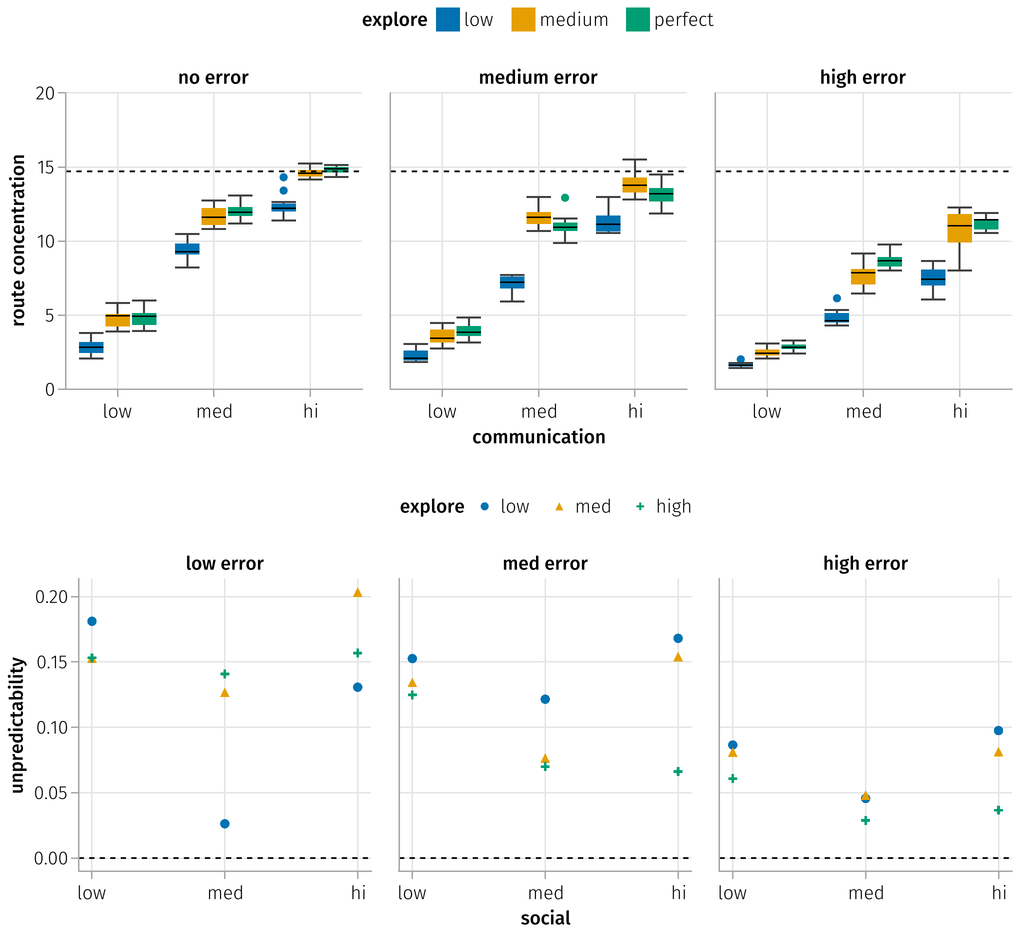


Figure 2. Route concentration (top, see text for definition) and unpredictability (bottom, see text for definition) for different values of exploration, communication, and communication error. The black line indicates values in a scenario where individuals have perfect information and do not communicate. We see that while higher levels of communication lead to an increase in route concentration, arrivals are most predictable at intermediate levels of communication.

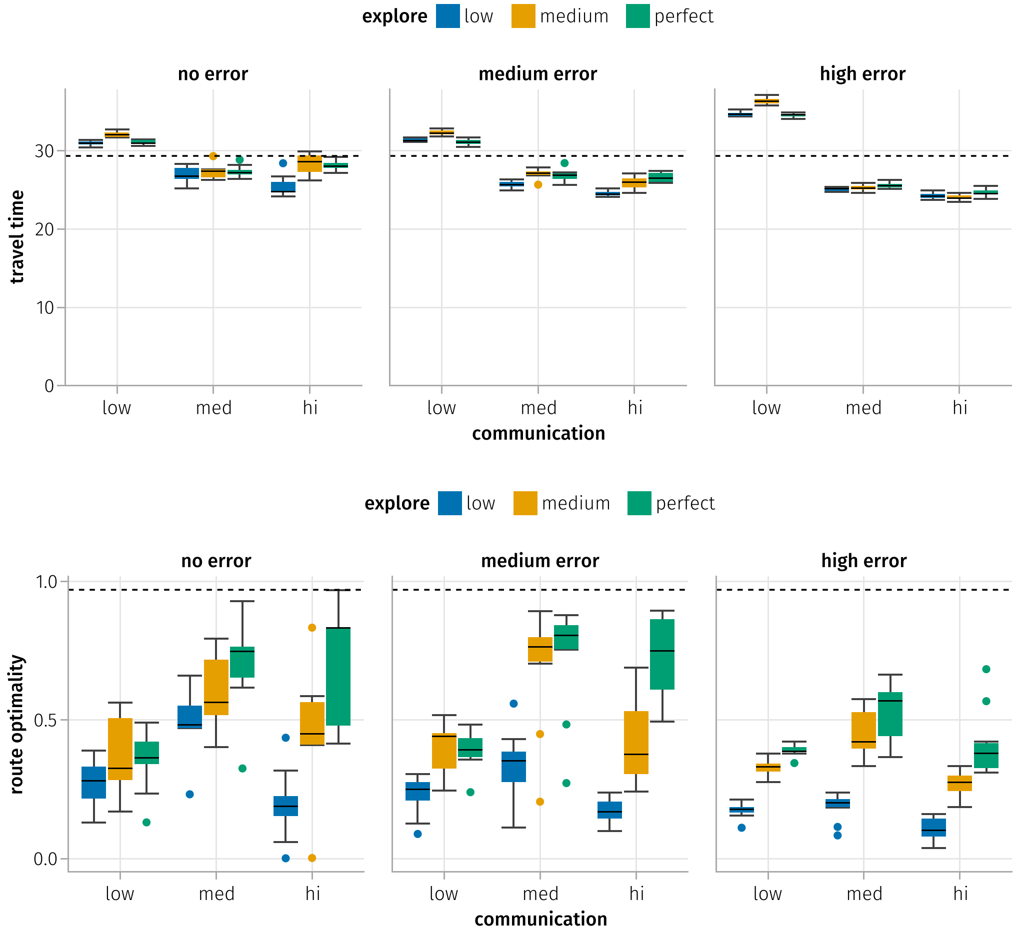


Figure 3. Average travel time (top) and route optimality (bottom, for definition see text) for different values of exploration, communication, and communication error. The black line indicates values as obtained in a scenario where individuals have perfect information and do not communicate. Travel time is high and increases with error for low communication while it is low and decreases with error for medium and high communication. Routes are generally closer to the optimum for intermediate levels of communication.

from the optimal route, leading to strong stochasticity across replicates (and thus high unpredictability).

For medium communication information, transfer between agents is high enough that a relatively accurate and complete consensus map emerges in the population. This leads to the emergence of similar, predictable, and relatively optimal routes in most replicate runs.

For high communication, the consensus between agents is even stronger. However, now the effects of information transfer override the effects of exploration so that unreliable consensus maps emerge. Therefore, while most agents take a similar route, that route is less optimal than for medium communication and can vary from case to case (implying lower predictability).

## 3.2 Preferred Destinations

With our second set of scenarios, we investigated how information and communication affect the chances of migrants to reach their preferred destination. For this we assumed that each agent at random picks one of the 10 destinations as its preferred target. The strength of preference

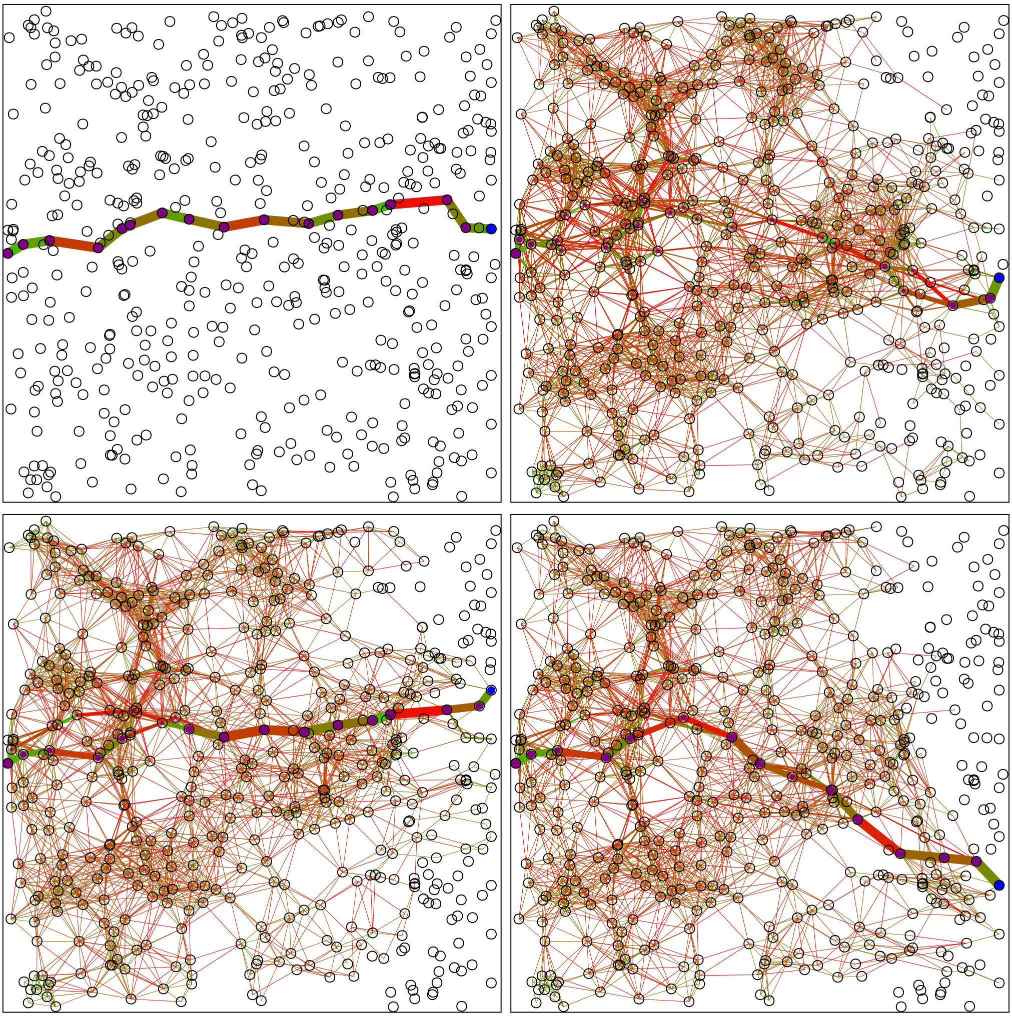

Figure 4. Migration trajectories for different scenarios. Thickness of the lines indicates traffic; colour represents friction (red: high). The top left panel shows the result for a full knowledge scenario (and thus the optimal path). The other panels are taken from communication scenarios. Top right: no error, low exploration, low communication; bottom left and right: low error, high exploration, high communication, different random seeds.

then indicates the increase in travel costs an agent is willing to incur in order to arrive at that destination.

Except for decreased route concentration (due to agents attempting to reach their target exits), adding preferences has little effect on the behaviour of the model as presented above (not shown). With respect to the ability of agents to follow their preferences, we find that if agents have perfect information a preference of 30% is sufficient to let the vast majority reach their preferred destination (Figure 5). Without prior information, however, in most scenarios less than half of the agents manage to arrive at their target. As before, agents travel most optimally for medium communication and high exploration, but even under these conditions arrival at target remains below 70%.

## 3.3 Interventions

A common response to a sudden increase in migration is the erection of physical or administrative barriers in the form of, for example, border closures or transport restrictions (Andersson, 2014). In our third set of scenarios we investigate how the reaction of migration routes to the sudden

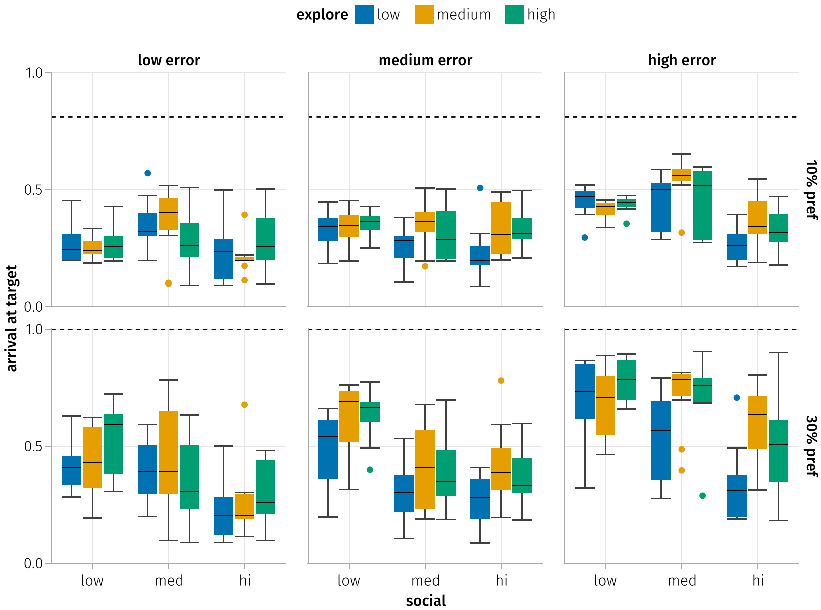


Figure 5. Proportion of agents arriving at their preferred destination for different values of exploration, communication, communication error, and strength of preference. The black line indicates values as obtained in a scenario where individuals have perfect information and do not communicate. Only for high error rates during communication and if agents are willing to incur an additional cost of 30% (bottom graph) do substantial proportions arrive at their preferred destination.

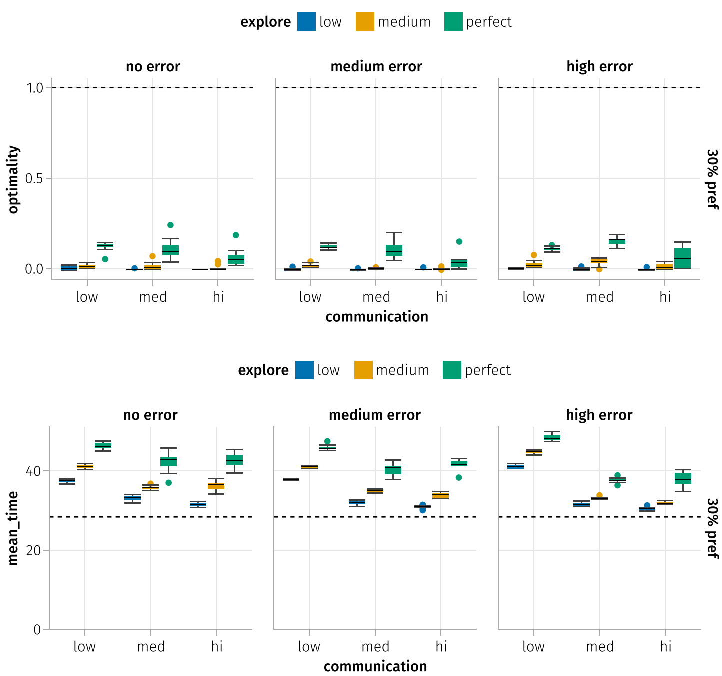


Figure 6. Properties of migration routes for an intervention scenario (see text for definitions). The black line indicates values from an equivalent scenario with full knowledge. After the intervention the quality of routes decreases dramatically while travel times increase substantially (cf. Figures 2 and 3).

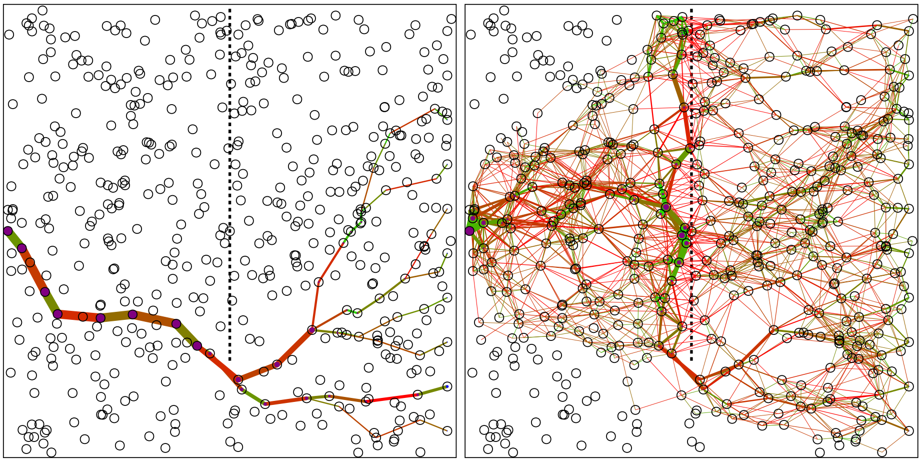

Figure 7. Migration trajectories for scenarios with interventions. Thickness of the lines indicates traffic; colour represents friction (red: high). The vertical dashed line represents the barrier. Shown are the results for full knowledge (left) and limited knowledge (right) with high error, perfect exploration, and low communication; both with a preference value of 30%. While agents easily manage to circumvent the obstacle when they are fully informed, only a small proportion of agents does so in the limited-information scenario.

appearance of barriers depends on the information regime. We implement barriers by, at timestep 500, increasing friction in all links that intersect with a vertical line across 80% (see Figure 7) of the world to 0.9 (which corresponds to an increase in travel time of about 8 time units). As we can see in Figures 6 and 7 migration routes in scenarios with full knowledge change to accomodate the barrier, so that neither quality nor travel time are substantially affected, although the number of agents reaching their preferred destination decreases as an effect of the detour.

In information-limited scenarios on the other hand, migration routes are largely unable to adapt. The quality of routes plummets and travel times increase substantially.

## 4 Discussion

We have shown that limitation and exchange of information can have a strong influence on the formation of migration routes. Migration routes can become less optimal, less predictable, and less centralised if migrants do not have perfect knowledge. Furthermore the proportion of migrants reaching their preferred destination is substantially lower in scenarios with more realistic informational logistics, and migrants find it much more difficult to adapt their routes to changing circumstances. The exchange of information in particular has a counterintuitive effect in that under certain conditions higher levels of communication can lead to less predictable routes (see also Hinsch & Bijak, 2019).

Even though this is a relatively simple, theoretical model, we can already at this stage draw a number of conclusions concerning migration modelling as well as the real-world dynamics of migration.

First and foremost we can conclude that information and information exchange are likely to be relevant for the formation of migration routes in the real world. In our model, how much information the agents have available and the frequency and accuracy of information exchange can lead to qualitatively different properties of the migration routes observed in the system. We know that in reality migrants do in fact often make travel decisions based on limited knowledge (Borkert et al., 2018; Crawley et al., 2016). It has also been found that (depending on country of origin) official sources of information are often met with very little trust and that in these situations most information is gathered from peers (Emmer et al., 2016; Prike et al., 2022). It seems therefore reasonable to expect that effects similar to those observed in our model can be found in reality.

Consequently any modelling attempting to predict migrants' movement in detail or on a small scale will need to incorporate these effects. This is particularly salient where models are meant to be used to support humanitarian measures in crisis situations. Previous modelling efforts in this area assume perfect knowledge (albeit sometimes with a limited range of perception) and thus optimal decision-making (e.g., Frydenlund et al., 2018; Hébert et al., 2018; Łatek et al., 2013; Suleimenova & Groen, 2020). We expect that including the effects of information in these models would change at least some of the observed results.

We also see that the migration journey itself not only shows considerable variations in dynamics depending on which scenario we assume but also can have important effects on other aspects of migration. Our results show that introducing a (more) realistic information regime can halve the number of migrants that arrive at their preferred destination. This contradicts the assumption of many models of migration that migrants *always* arrive at their chosen destination (e.g., Ahmed et al., 2016; Lin et al., 2016). We can conclude that while the situation might be different for voluntary migration, at least models of forced migration should assume that a considerable proportion of migrants will be diverted on their journey and that this depends on the information regime in the population. Similarly the effects of introducing a barrier to migration differ considerably depending on whether we assume perfect information or not. Models that, for example, attempt to extrapolate the effect of border closures on migration will risk vastly overestimating the effectiveness in steering migration streams unless the role of information is included.

The situation becomes even more complicated when we look in more detail at how the specific variables we modelled correspond to aspects of real-world situations. The frequency and accuracy with which migrants communicate might be a result of cultural factors but will also depend on simple practical aspects of their circumstances, such as availability of mobile phones, opportunities to charge them, and accessibility of service in the travel area (Gillespie et al., 2018). Similarly the access to local information (exploration in our model) can be strongly affected by something as straightforward as a language barrier. Empirical studies furthermore show that how well informed migrants are about their journey and their destination as well as their capacity to obtain information can vary dependent on factors such as country of origin (Dimitriadi, 2018; Emmer et al., 2016). Based on our results it can therefore be expected that migrant populations will differ, for example, with respect to how predictable their travel routes turn out to be or how likely it is that migrants end up at their planned or preferred destination. Modelling studies aiming at predicting migrant arrivals therefore have to take the specific properties of the modelled population as well as how they relate to the situation into account.

Even though the importance of networks for migration decisions has been recognised in previous studies (Gurak & Caces, 1992), many models that explicitly include networks simplify them in at least one of two ways—by assuming that networks do not change over time (e.g., Simon, 2019), or, if so, then deterministically and/or or by summarising the effects of networks as a single numerical value (e.g., "strength" or "number of connections"; e.g., Lin et al., 2016) that then is used during decision-making. Our results show that the situation can be considerably more complicated. We find that not only the existence and strength of the network matters, but also what individuals use it for. In our case that is information, but it does not seem implausible that other, known, network effects such as monetary support or logistic aid have similarly fine-grained dynamics that affect the other parts of the system and therefore need to be taken into account.

## 4.1 Limitations and Future Work

While our results clearly show that informational logistics affect the migration journey, it is difficult to judge how exactly the scenarios we investigated relate to specific real-world situations. At this point our modelling efforts therefore have to remain a proof of principle. However, given the wide range of parameter values we tested, we can assume that similar dynamics will take place in real systems. Nevertheless, additional effort will be required to calibrate the model to empirical data in order to test the relevance of our results.

We intentionally kept our model of information and information exchange simple and abstract, partially due to a lack of reliable empirical information and partially in order to investigate the simplest scenarios first. At this point the model is therefore clearly "unrealistic" in many aspects. The two biggest simplifying assumptions concerning information in our model have to be, first, that agents (in the "communication" scenarios) have no prior knowledge and, second, that information is retained and exchanged entirely indiscriminately. Strictly speaking both assumptions are clearly wrong. In the absence of empirical data on either aspect, however, any attempt at making the model more realistic would have lead to a massive increase in the number of potential realisations and in the size of the parameter space. As it is, this version of the model and the scenarios we tested serve to describe both extremes of what is possible in reality. Any real population will likely be somewhere between our "full knowledge" and "no knowledge" scenarios.

In this version of the model we assume for the sake of simplicity that the only choice agents have is which route to take. We know, however, that in reality migrants have more options available. For one they may decide that they would be better off returning to their country of origin when, for example, faced with an obstacle. More importantly, however, there are many situations where it can be prudent or even necessary to delay the continuation of the journey (Anam et al., 2008; DeVoretz & Ma, 2002). If included this would add timing of migration decisions as an important dimension to the model.

We also completely ignored the heterogeneity that every human population shows. We know that means and circumstances often differ between early and late migrants on the same route (Lindstrom & López Ramírez, 2010). If we assume that access to information differs in a similar way, we can easily imagine that well- or better-informed early migrants serve as "trailblazers," chosing good routes and transmitting their experiences to followers who a priori might not be as well-informed.

Another aspect worth exploring in the future that was out of scope for this study is the role of network structure and density in information transmission and—ultimately—route formation. To a certain degree we can assume that, for example, the effects of an increase in information exchange due to higher network density are analogous to the effects of increased information exchange we modelled in our scenarios. However, new dynamics might emerge if networks interact with other aspects of the system, for example, if people have a tendency to travel in groups (Collins & Frydenlund, 2016) or if pre-existing networks are stratified by social status and thus access to information and capital.

We also—again for the sake of simplicity—did not include many of the additional factors known to be important in real-world migration systems. There are, for example, good indications that at least in some situations, smugglers play an important role in maintaining or even shaping migration routes, in particular when there are pre-existing non–migration-related smuggling routes (Triandafyllidou, 2018). We also completely ignored the effects of material means on the availability of information and transportation (see the point on temporal heterogeneity above).

Furthermore the difficulty of the journey a migrant expects might itself affect their choice of destination or even the decision to migrate in the first place. However, that difficulty itself might decrease over time if a migration route emerges and leads to the establishment of supporting infrastructure. In this case the migration decision is therefore part of a feedback loop and cannot be understood without taking into account the journey.

### 4.2 Conclusions

We can conclude that information is an important, yet largely neglected aspect of migration that deserves more attention in the future. This is likely to apply to all stages of the migration journey, from the decision to leave to the journey itself to the decision to remain in the country of arrival or to move on, and finally in the decision to return if the opportunity arises. Our model is a simple first step in exploring this issue that—as discussed above—leaves ample scope for extension. We are looking forward to seeing the interesting future developments in this area.

Our work also confirms that—as is the case for many other social phenomena—small-scale interactions between individuals can have substantial effects in the context of migration. While it might for a given situation be possible to find macroscopic approximations for the effects of microscopic interactions this can be a difficult and time-consuming process. If we assume that information exchange is not the only relevant interaction between migrants (others include direct interactions such as transfer of capital and indirect interactions via environmental factors, such as economic effects of transit zones or the establishment of smuggling services)—we have to conclude that in many if not most situations some form of bottom-up modelling strategy will be required when dealing with the dynamics and effects of migration (Willekens, 2018). This further strengthens the case for the use of agent-based modelling in the social sciences (Chattoe-Brown, 2013).

## Acknowledgments

This work has received funding from the European Union's Horizon 2020 research and innovation programme; European Research Council grant No. 725232 BAPS: Bayesian Agent-Based Population Studies (MH and JB). This article reflects the authors' views, and the Research Executive Agency of the European Commission is not responsible for any use that may be made of the information it contains. We are grateful to the anonymous reviewers for their comments that helped improve an earlier draft of the manuscript.

## References

Adams, J., White, G., & Araujo, R. (2021). The role of mistrust in the modelling of opinion adoption. *Journal of Artificial Societies and Social Simulation*, *24*(4), Article 4. https://doi.org/10.18564/jasss.4624

Ahmed, M. N., Barlacchi, G., Braghin, S., Calabrese, F., Ferretti, M., Lonij, V., Nair, R., Novack, R., Paraszczak, J., & Toor, A. S. (2016). A multi-scale approach to data-driven mass migration analysis. In R. Gavaldá, I. Žliobaitė, & J. Gamma (Eds.), *SoGood 2016: Proceedings of the first workshop on data science for social good* (CEUR Workshop Proceedings series, Vol. 1831). Available at http://ceur-ws.org/Vol-1831/paper_4.pdf

Anam, M., Chiang, S.-H., & Hua, L. (2008). Uncertainty and international migration: An option cum portfolio model. *Journal of Labor Research*, *29*(3), 236–250. https://doi.org/10.1007/s12122-007-9033-y

Andersson, R. (2014). Time and the migrant other: European border controls and the temporal economics of illegality. *American Anthropologist*, *116*(4), 795–809. https://doi.org/10.1111/aman.12148

Bertoli, S., & Fernández-Huertas Moraga, J. (2013). Multilateral resistance to migration. *Journal of Development Economics*, *102*, 79–100. https://doi.org/10.1016/j.jdeveco.2012.12.001

Bijak, J., Higham, P. A., Hilton, J., Hinsch, M., Nurse, S., Prike, T., Smith, P. W. F., Uhrmacher, A. M., & Warnke, T. (2021). *Towards Bayesian model-based demography: Agency, complexity and uncertainty in migration studies*. Springer. https://doi.org/10.1007/978-3-030-83039-7

Borkert, M., Fisher, K. E., & Yafi, E. (2018). The best, the worst, and the hardest to find: How people, mobiles, and social media connect migrants in(to) Europe. *Social Media & Society*, *4*(1), Article 35. https://doi.org/10.1177/2056305118764428

Brekke, J.-P., & Brochmann, G. (2015). Stuck in transit: Secondary migration of asylum seekers in Europe, national differences, and the Dublin regulation. *Journal of Refugee Studies*, *28*(2), 145–162. https://doi.org/10.1093/jrs/feu028

Castles, S., de Haas, H. G., & Miller, M. J. (2014). *The age of migration: International population movements in the modern world* (5th ed.). Palgrave Macmillan.

Chattoe-Brown, E. (2013). Why sociology should use agent based modelling. *Sociological Research Online*, *18*(3), 31–41. https://doi.org/10.5153/sro.3055

Collins, A. J., & Frydenlund, E. (2016). Agent-based modeling and strategic group formation: A refugee case study. In *2016 winter simulation conference (WSC)* (pp. 1289–1300). IEEE. https://doi.org/10.1109/WSC.2016.7822184

Crawley, H., Jones, K., McMahon, S., Duvell, F., & Sigona, N. (2016). *Unpacking a rapidly changing scenario: Migration flows, routes and trajectories across the Mediterranean* (Report No. 1). Centre for Trust, Peace and Social

Relations, Coventry University. Available at https://pure.coventry.ac.uk/ws/portalfiles/portal/11553989/MEDMIG_research_brief_1_rapidy_changing_scenario.pdf

Dekker, R., Engbersen, G., Klaver, J., & Vonk, H. (2018). Smart refugees: How Syrian asylum migrants use social media information in migration decision-making. *Social Media and Society*, *4*(1), Article 80. https://doi.org/10.1177/2056305118764439

DeVoretz, D. J., & Ma, J. (2002). Triangular human capital flows between sending, entrepot and the rest of the world regions. *Canadian Studies in Population*, *29*(1), 53–69. https://doi.org/10.25336/P6F603

Dimitriadi, A. (2018). Deciding to be mobile. In *Irregular Afghan migration to Europe* (pp. 31–62). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-52959-2_2

Duggins, P. (2017). A psychologically-motivated model of opinion change with applications to American politics. *Journal of Artificial Societies and Social Simulation*, *20*(1), Article 13. https://doi.org/10.18564/jasss.3316

Emmer, M., Richter, C., & Kunst, M. (2016). *Flucht 2.0: Mediennutzung durch Flüchtlinge vor, während und nach der Flucht* [Escape 2.0: Media use by refugees before, during and after the flight] (Technical report). Institut für Publizistik- und Kommunikationswissenschaft, Freie Universität Berlin.

Frydenlund, E., & De Kock, C. (2020). Agent-based modeling within forced migration research: A review and critique. *Refugee Review*, *4*(1), 53–68. Available at https://espminetwork.com/wp-content/uploads/2020/06/Frydenlund-and-De-Kock.pdf

Frydenlund, E., Foytik, P., Padilla, J. J., & Ouattara, A. (2018). Where are they headed next? Modeling emergent displaced camps in the DRC using agent-based models. In *2018 winter simulation conference (WSC)* (pp. 22–32). IEEE. https://doi.org/10.1109/WSC.2018.8632555

Gilbert, A., & Koser, K. (2006). Coming to the UK: What do asylum-seekers know about the UK before arrival? *Journal of Ethnic and Migration Studies*, *32*(7), 1209–1225. https://doi.org/10.1080/13691830600821901

Gilbert, E. N. (1961). Random plane networks. *Journal of the Society for Industrial and Applied Mathematics*, *9*(4), 533–543. https://doi.org/10.1137/0109045

Gillespie, M., Osseiran, S., & Cheesman, M. (2018). Syrian refugees and the digital passage to Europe: Smartphone infrastructures and affordances. *Social Media and Society*, *4*(1). https://doi.org/10.1177/2056305118764440

Greenwood, M. J. (2005). Modeling migration. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (pp. 725–734). Elsevier. https://doi.org/10.1016/B0-12-369398-5/00352-2

Gurak, T., & Caces, F. (1992). Migration networks and the shaping of migration systems. In M. M. Kritz, L. L. Lim, & H. Zlotnik (Eds.), *International migration systems: A global approach* (pp. 150–176). Oxford University Press.

Hébert, G. A., Perez, L., & Harati, S. (2018). An agent-based model to identify migration pathways of refugees: The case of Syria. In L. Perez, E.-K. Kim, & R. Sengupta (Eds.), *Agent-based models and complexity science in the age of geospatial big data* (pp. 45–58). Springer. https://doi.org/10.1007/978-3-319-65993-0_4

Hinsch, M., & Bijak, J. (2019). *Rumours lead to self-organized migration routes* [Workshop presentation]. ABMHuB 2019 International Workshop on Agent-Based Modelling of Human Behaviour, Newcastle, UK. Available at http://abmhub.cs.ucl.ac.uk/2019/

Hinsch, M., & Bijak, J. (2021). Principles and state of the art of agent-based migration modelling. In *Towards Bayesian model-based demography: Agency, complexity and uncertainty in migration studies* (pp. 33–49). Springer. https://doi.org/10.1007/978-3-030-83039-7_3

Horn, F., & Jackson, R. (1972). General mass action kinetics. *Archive for Rational Mechanics and Analysis*, *47*(2), 81–116. https://doi.org/10.1007/BF00251225

Katz, E., & Stark, O. (1987). International migration under asymmetric information. *The Economic Journal*, *97*(387), 718–726. https://doi.org/10.2307/2232932

Kingsley, P. (2016). *The new odyssey: The story of Europe's refugee crisis*. Guardian Books.

Klabunde, A., & Willekens, F. (2016). Decision-making in agent-based models of migration: State of the art and challenges. *European Journal of Population*, *32*(1), 73–97. https://doi.org/10.1007/s10680-015-9362-0, PubMed: 27069292

Łatek, M. M., Rizi, S. M. M., & Geller, A. (2013). Verification through calibration: An approach and a case study of a model of conflict in Syria. In *2013 winter simulation conference (WSC)* (pp. 1649–1660). IEEE. https://doi.org/10.1109/WSC.2013.6721547

Lin, L., Carley, K. M., & Cheng, S.-F. (2016). An agent-based approach to human migration movement. In *2016 winter simulation conference (WSC)* (pp. 3510–3520). IEEE. https://doi.org/10.1109/WSC.2016.7822380

Lindstrom, D. P., & López Ramírez, A. (2010). Pioneers and followers: Migrant selectivity and the development of U.S. migration streams in Latin America. *The Annals of the American Academy of Political and Social Science*, *630*(1), 53–77. https://doi.org/10.1177/0002716210368103, PubMed: 24489382

Macy, M. W., & Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology*, *28*(1), 143–166. https://doi.org/10.1146/annurev.soc.28.110601.141117

Martins, A. C. R. (2009). Bayesian updating rules in continuous opinion dynamics models. *Journal of Statistical Mechanics: Theory and Experiment*, *2009*(2), Article P02017. https://doi.org/10.1088/1742-5468/2009/02/P02017

Nordio, A., Tarable, A., Chiasserini, C. F., & Leonardi, E. (2018). Belief dynamics in social networks: A fluid-based analysis. *IEEE Transactions on Network Science and Engineering*, *5*(4), 276–287. https://doi.org/10.1109/TNSE.2017.2760016

Prike, T., Bijak, J., Higham, P. A., & Hilton, J. (2022). How safe is this trip? Judging personal safety in a pandemic based on information from different sources. *Journal of Experimental Psychology: Applied*, *28*(3), 509–524. https://doi.org/10.1037/xap0000415, PubMed: 35201841

Sačer, S., Palić, M., Grünhagen, M., & Kundid, T. (2017). Determinants of choice of migration destination: Evidence from the Western Balkan transit route. *International Journal of Sales, Retailing and Marketing*, *6*(1), 48–60. Available at https://www.circleinternational.co.uk/wp-content/uploads/2021/01/IJSRM6-1.pdf

Simon, M. (2019). Path dependency and adaptation: The effects of policy on migration systems. *Journal of Artificial Societies and Social Simulation*, *22*(2), Article 2. https://doi.org/10.18564/jasss.3970

Suleimenova, D., & Groen, D. (2020). How policy decisions affect refugee journeys in South Sudan: A study using automated ensemble simulations. *Journal of Artificial Societies and Social Simulation*, *23*(1), Article 2. https://doi.org/10.18564/jasss.4193

Triandafyllidou, A. (2018). Migrant smuggling: Novel insights and implications for migration control policies. *The Annals of the American Academy of Political and Social Science*, *676*(1), 212–221. https://doi.org/10.1177/0002716217752330

Willekens, F. (2018). Towards causal forecasting of international migration. *Vienna Yearbook of Population Research*, *16*, 1–20. https://doi.org/10.1553/populationyearbook2018s199

# Appendices

## Appendix 1: Scenario Parameters

Table A1 shows the parameters that vary between scenarios. Scenarios 'preferences' and 'obstacle' were run in combination with all configurations for 'full info' and 'communication', respectively.

Table A1. Parameters that vary between scenarios.

| Parameter | Explanation | Full info | Communication | Preferences | Obstacle |
|---|---|---|---|---|---|
| *n_ini_contacts* | Initial number of contacts | 0 | 5 | var | var |
| *p_know_target* | Probability to know an exit at the start of the simulation | 1.0 | 0.0 | var | var |
| *p_know_link* | "Link" | 1.0 | 0.0 | var | var |
| *p_know_city* | "City" | 1.0 | 0.0 | var | var |
| *speed_expl_ini* | Exploration on departure | 1.0 | 0.0 | var | var |
| *n_contacts_max* | Maximum number of contacts | 0 | 20 | var | var |
| *p_drop_contact* | Probability to lose a contacts | 0 | 0.05 | | |
| *pref_target* | Preference for specific destination | 1.0 | 1.0 | 1.1, 1.3 | 1.0, 1.3 |
| *convince* | See section 2 | 0.0 | 0.5 | var | var |
| *convert* | See section 2 | 0.0 | 0.1 | var | var |
| *confuse* | See section 2 | 0.0 | 0.3 | var | var |
| *error*, *error_frict* | Communication error | n.a. | {0.0, 0.0}, {0.12, 0.015}, {0.36, 0.045} | var | var |
| *rate_explore_stay*, *p_find_links*, *p_find_dests*, *speed_expl_stay*, *speed_expl_move* | Rate of exploration and quality of information gained when exploring | 0 | {1.0, 0.1, 0.1, 0.5, 0.5}, {4.0, 0.8, 0.5, 1.0, 1.0}, {10.0, 1.0, 1.0, 1.0, 1.0} | var | var |
| *p_keep_contact*, *p_info_contacts*, *p_transfer_info* | Probability to gain contacts, rate of information exchange | 0 | {0.1, 0.1, 0.1}, {0.3, 0.3, 0.3}, {0.6, 0.6, 0.6} | var | var |

Some parameters were changed as a set (indicated as '{}'). Error level for example changes both, *error* and *error_frict*. A value of '{0.12, 0.015}' then corresponds to a value of 0.12 for *error* and 0.015 for *error_frict*.

## Appendix 2: Default Parameter Values

Table A2 shows the values of all parameters that do not change across the scenarios. The submodels on risk and resources, respectively, were not used and the corresponding parameters have been omitted.

Table A2. Values of all parameters that do not change across scenarios.

| Parameter | Default | Parameter | Default |
|---|---|---|---|
| *n_cities* | 600 | *n_nearest_exit* | 5 |
| *link_thresh* | 0.12 | *qual_entry* | 0.0 |
| *n_exits* | 10 | *res_entry* | 0.0 |
| *regular_exits* | true | *qual_exit* | 1.0 |
| *n_entries* | 1 | *res_exit* | 1.0 |
| *regular_entries* | true | *dist_scale* | 1.0 |
| *exit_dist* | 1.0 | *frict_range* | 0.5 |
| *entry_dist* | 0.0 | *p_unkown_city* | 0.0 |
| *n_nearest_entry* | 5 | *p_unknown_link* | 0.0 |
| *rate_dep* | 20.0 | *move_rate* | 0.0 |
| *rate_plan* | 100.0 | *move_speed* | 0.1 |
| *res_exp* | 0.5 | *p_notice_death_c* | 0.0 |
| *qual_exp* | 0.5 | *p_notice_death_o* | 0.0 |
| *frict_exp* | 1.25 | *qual_bias* | 1.0 |
| *qual_weight_x* | 0.25 | *path_penalty_loc* | 1.0 |
| *qual_weight_res* | 0.0 | *path_penalty_risk* | 0.0 |
| *qual_tol_frict* | 2.0 | | |