

CORVINUS UNIVERSITY OF BUDAPEST

UNDERSTANDING INTERNAL MIGRATION
THROUGH A NETWORK PERSPECTIVE

DOCTORAL DISSERTATION

Supervisor: László Lőrincz PhD

Gergely Mónus

Budapest, 2026

Gergely Mónus

Understanding Internal Migration
Through a Network Perspective

Doctoral School of Sociology and Communication
Science

Supervisor: László Lőrincz, PhD

© 2026 Gergely Mónus

Corvinus University of Budapest
Doctoral School of Sociology and Communication
Science

Understanding Internal Migration
Through a Network Perspective

Doctoral Dissertation

Mónus Gergely

Budapest, 2026

Content

1. INTRODUCTION	5
1.1. Traces of flow network thinking through (internal) migration theory	8
1.1.1. The main concept and importance of internal migration	8
1.1.2. The origins of migration theory – Ravenstein revisited.....	9
1.1.3. The migration transition hypotheses.....	11
1.1.4. Push-pull theories and gravity models – independent units and independent dyads.....	13
1.1.5. System theories of (internal) migration	17
1.2. Inter-urban flow networks.....	24
1.2.1. Current trends in urbanization: Polycentric development.....	25
1.2.2. The new macro level network literature of migration	29
1.3. The Hungarian internal migration system after the regime change	33
1.4. The network perspective: a possible synthesis	37
1 st ARTICLE: RURAL-URBAN FLOWS DETERMINE INTERNAL MIGRATION STRUCTURE ACROSS SCALES.....	41
1. Introduction.....	43
2. Data and methods	48
2.1. Data.....	48
2.2. Methods.....	49
3. Results.....	52
3.1. Comparison of the two networks.....	52
3.2. Multiscale community structure	58
4. Discussion	65
5. Summary and conclusion.....	68
2 nd ARTICLE:.....	71
INTERNAL MIGRATION FLOW NETWORK REORGANIZATION DURING A HOUSE PRICE SHOCK IN A CAPITAL CITY.....	71
1. Introduction.....	72
2. Data and Methods.....	79
2.1. Data.....	79

2.2.	Modeling strategy	80
3.	Results.....	83
3.1.	Change in push-pull factors across the study periods.....	83
3.2.	The inter-district migration network across the study period	85
3.3.	Analysis of triangles and weight transitivity in the migration network	86
3.4.	ERGM results	89
4.	Discussion	94
5.	Summary	96
	Appendix	99
A.1.	Model diagnostics.....	99
A.2.	Structure of the house price boom in Hungary	103
3 rd	ARTICLE:	105
	SOCIAL CONNECTIONS AND LINKS TO COMMUNITIES DRIVE INTRA- METROPOLITAN MIGRATION WHEN HOUSING IS SCARCE	105
1.	Introduction.....	106
2.	Data and methods	114
2.2.	Data sources	114
2.3.	Methods.....	116
3.	Results.....	119
3.1.	Migration and housing development in the Budapest agglomeration area	119
3.2.	Links to communities	121
3.3.	Regression modeling of migration flows.....	123
4.	Discussion	129
5.	Summary	131
	Appendix	133
A.1.	Measurement of clique cohesion - robustness of the density measure.....	133
4.	SUMMARY AND REFLECTIONS	136
	REFERENCES	143
	AUTHOR'S OWN PUBLICATIONS RELATED TO THE TOPIC	165

LIST OF TABLES AND FIGURES

List of Figures

Introduction

Figure 1: The currents of migration in the UK by Ravenstein 1885.

Source: E.G. Ravenstein (1885): The Laws of Migration.
<https://doi.org/10.2307/2979181>

Figure 2: Internal migration rates in municipalities of Hungary, for three periods.

Source: The National Atlas of Hungary, Volume 3 – Society.
nemzetiatlasz.hu/MNA/3_en.html

1st Article

Figure 1: Degree distribution in the two networks. Author's own.

Figure. 2: Node degree and node strength by settlement type. Author's own.

Figure 3: Migration flows with the largest volume for every settlement for the observed and the radiation network. Author's own.

Figure 4: Results for multi-scale community detection with the map equation. Author's own.

Figure 5: Multi-scale community structure in the observed network. Author's own.

Figure 6: Center role inside modules. Author's own.

Figure 7: Multi-scale community structure in the radiation network. Author's own.

2nd Article

Figure A.1.1. Distributions of model statistics across networks simulated from the fitted model

Figure A.2.1. Year on year mean house price change in Hungarian districts by population size categories through the study period. Author's own.

3rd Article

Figure 1: Schematic example for identifying connections to clique-like groups known by users of the online social networking site. Author's own.

Figure 2: Population growth vs housing growth by period across the 2010s. Author's own.

Figure 3: Number of online social connections between settlements vs number of cliques larger than 5 known in destination by origin users. Number of online social connections between settlements vs mean density of cliques larger than 5 in destination known by origin users. Author's own.

Figure 4: Johnson-Neyman plots for interaction effects in different model specifications. Author's own.

List of Tables

1st Article

Table 1: Descriptive statistics for the two networks. Own calculations.

2nd Article

Table 1: Evolution of Budapest-to-country ratios for key migration determinants. Own calculations.

Table 2: Main statistics of the inter-district migration network of Hungary. Own calculations.

Table 3: Analysis of middleman-triangles with Budapest as middleman. Own calculations.

Table 4: ERGM results for the three periods. Own calculations.

Table 5: ERGM results with past flows as controls. Own calculations.

Table A.1: count-ERGM model fit statistics

3rd Article

Table 1: Regression results for social connection measures. Own calculations.

Table 2: Regression results for social connection, housing, and urbanisation measures. Own calculations.

Table A.1.1. Correlations among clique statistics and connection counts at the settlement-pair level

Summary and reflections

Table 1: Short answers to the overarching research question

1. INTRODUCTION

The main argument of this dissertation is simple: internal migration happens through a path-dependent network of flows, which have internal mechanisms, and every factor affecting migration is embedded in this flow structure. As I will describe, this notion is as old as migration research. However, wide-spread contemporary models of internal migration do not account for it. They rely on modeling independent units, and independent dyads (bi-lateral flows), without a systematic framework. I argue that the application of new network science methods to empirical analysis can account for this simple and old notion, and achieve systematic quantitative models of internal migration flow systems. And if we apply these network methods, and become able to account for systematic inter-dependency, we are able to achieve empirical results that make incremental steps in migration theory that otherwise would be impossible.

As I will summarize in this Introduction chapter, the traces of network thinking were present since the very origin of migration research, in the groundbreaking works of Ernst Ravenstein, who built his model of internal migration in the United Kingdom on a system of currents between locations. And the network idea has been present ever since. Macro theories of migration, which connect development to migration transitions predict that in developed economies, migration becomes more complex, and happen on a lattice of locations. Systematic migration theories showed as early as the 1970's that through the idea of information feedback, the structure of migration flows capture the decision making behind individual migration decisions, leading to a path-dependency in flow structure. Migration theorists of the late 20th century deal with the idea how emergent path dependency of flows originates from micro-level social network dynamics. Researchers who connect urban hierarchy and internal migration showed for the last 30 years how hierarchies change from being determined by the power of central places to network based hierarchies.

Massey writes in his seminal synthetising paper from 1990 that migration studies is fragmented into a diverse set of semi-autonomous literatures with little interaction between them. This is especially true for internal migration, which has a special place in social science – because it does not have its own academic tradition. There is for example no specialized journal on internal migration processes, and monographs

with a unique focus on the topic are also rare. Usually, internal migration is covered by the topics of urbanization, sub-urbanization, residential mobility and residential segregation, housing, or as a special aspect of social and economic geography. It can be also a sub-category in demography. This makes it inevitable that the theoretical background of this thesis covers a lot of different topics like a patchwork. Besides summarizing the emergent inter-settlement network literature, and the traces of network thinking through the development of migration theory, I will briefly summarize status quo push-pull models with independent unit and dyad approaches. After, I will show how polycentric urban development theory, and the new, emerging literature of inter-urban networks can enrich internal migration research. I will also cover the special Hungarian context, and the post-socialist migration transition the country experienced after the regime change, and argue that the new network approach can help us understand special migration contexts better.

The three essays making up this dissertation do not only aim to achieve methodological novelty in the area. They all aim to critique simple individual unit thinking, where criticism is due, and use classical theoretical sources to interpret results for the novel methods. While they can only offer small steps in a hopefully right direction, I hope they can serve as experimental cases of how re-reading foundational classics can make us interpret results of new network methods, offering new evidence for internal migration research.

The argument outlined above can be condensed into the overarching research question of this dissertation:

How do interdependent flow structures shape the spatial patterning of internal migration in Hungary – and what can network methods reveal about this structure that unit-based and dyadic models cannot?

Strictly speaking, this question has two components: an empirical one, about the structure of the Hungarian internal migration system; and a methodological one, about what relational modeling recovers that individual unit and individual dyad approaches structurally miss. The two components cannot be fully separated: it is exactly the argument of this dissertation that the structure only becomes visible through relational methods. The question is addressed at three nested spatial scales, through three sub-questions, each corresponding to one article of the thesis. First, on

the national, inter-settlement scale: what elements of the flow structure connect the different spatial scales of the internal migration network, and can population distribution and geographical layout alone account for them? Second, on the inter-district scale, and on a short timescale: does a housing price shock in the dominant capital city restructure the migration flow network, or does path dependency constrain spatial reorganization? Third, on the intra-metropolitan scale: do social ties – and especially ties to densely knit communities – direct residential migration flows where housing supply cannot act as a pre-emptive pull factor? The concluding chapter returns to the overarching question and its two components, and summarizes what has (and what has not) been answered.

1.1. Traces of flow network thinking through (internal) migration theory

1.1.1. The main concept and importance of internal migration

Internal migration is the residential relocation of people from one defined locality to another within a country. Localities can be regions, counties, settlements (or rarely districts or other parts of settlements). Internal migration does not have clear definitions. Most studies differentiate internal migration and residential mobility based on whether a move displaces people's daily activity spaces. Short distance moves that let people stay in one daily activity space are usually categorized as residential mobility, while medium and long distance moves are considered internal migration (Roseman 1971, Niedomysl 2011). Contemporary research shows that the distinction between the two categories are not clear-cut, and studies using administrative border crossings can only achieve fuzzy separation of the two behaviours (Coulter et al 2015). Short distance residential mobility is usually attributed to life-cycle changes of individuals, and the medium and long distance moves to employment-related reasons (Jia et al. 2023). Internal migration is a population process with growing importance in Europe and Hungary. This importance mainly comes from the fact that the very low rates of natural increase mean that migration compose a larger fraction of population change (White and Lindstrom 2005). In fact, internal migration has replaced fertility and mortality as the primary demographic process shaping the spatial distribution of populations within countries in Europe (Rowe and Patias 2020).

Internal migration, while a relatively large phenomenon, occupies a small fraction of migration research (Pitoski et al. 2021a). This imbalance is striking given that based on reliable estimates, internal migration volumes are approximately three times larger than international migration (Bell et al. 2002), and is a widely recognized pre-eminent process underpinning shifts in the pattern of human settlement (Bell et al. 2015). In 2019 in Hungary, according to Hungarian Central Statistical Office (HCSO), there were more than 580,000 internal migration events, compared to the more than 138,000 international migration events (and the latter number includes both Hungarian and non-Hungarian citizens, and both in- and out migration). The large attention to international migration is understandable since its current level is

brought about by globalization, one of the most important themes in the last few decades of social science. But internal migration is also deeply affected by the large scale societal changes of recent history, only these changes are mostly examined from the point of large levels of urbanization in developing countries, and sometimes counter-urbanization trends in the west. Also, the first social scientific work, the Laws of Migration by Ernst Ravenstein (1885, 1889) focused on the internal migration system of the UK, so the origins of migration theory did not focus on international movement of people.

1.1.2. The origins of migration theory – Ravenstein revisited

The state of Ravenstein's work in migration theory is fascinating for several reasons. It is undoubtedly the earliest paper in migration research that had a constant influence through the development of the field, and even today (Rees and Lomax 2019). Obviously, for the contemporary reader, the usage of the term "laws" can be seen as pure XIX.th century positivism. But even today, 110-130 years later, migration researchers use them as some kind of an analytic null model (for example Macedo et al 2020). The striking aspect is not how many researchers find Ravenstein fruitful, but the fact the large majority of them miss one main aspects of his laws, that these describe an interdependent system. Not a system of individual localities, but one that is defined by the relations of those.

Ravenstein first worked with census data for the analysis for internal migration in 1876, but his seminal paper is the one from 1885, in which he revisited the topic after a new census. In that paper, he describes 7 laws of migration (summarized):

- 1. Most migrants only proceed to move short distances, and consequently there takes place an universal shifting of population which produces "currents of migration", directed towards great centers of commerce and industry.*
- 2. Population in the immediate surrounding of a town with rapid growth migrate into it, and the gaps they leave are filled by migrants from more distant areas. This way, step by step, the absorbing force of growing centers is spreading through the whole country. Migrants enumerated in a large absorbing center will consequently grow less with the distance to their homes.*

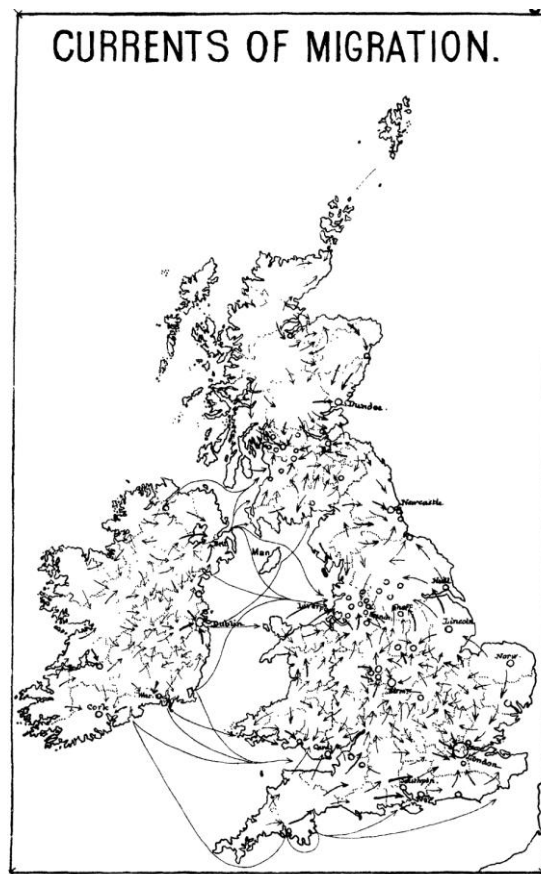
- 3. The process of dispersion is the inverse that of absorption, and it has similar features.*
- 4. Main currents of migration produce their compensating counter current.*
- 5. Migrants proceeding long distances tend to move to large centers of commerce and industry.*
- 6. Population of towns is less migratory than rural population.*
- 7. Females are more migratory than males.*

By applying “current” as the basic theoretical building block, Ravenstein’s work aims to explain relations. I do not want to claim Ravenstein practiced network science *avant la lettre*, but his laws can be translated easily to a network model: localities are nodes, where agents reside, and the movement of agents produces the emergent macro-level structure of migration streams. Ravenstein’s laws describe an interdependent system, where each locality is influenced by other localities, and this influence is the core of explanation. Surprisingly, this aspect of his work mostly disappeared during the 20th century – while there were few important theorists connecting to this aspect of his work, empirical research treated other aspects as more important. While some kind of pattern identification is the main and most useful heuristic of macro level internal migration research (Greenwood and Hunt 2003), the patterns researchers look for are patterns of individual units (localities).

However, one-by-one, many laws of Ravenstein laid foundations for migration research, without the systematic element. For example, patterns described in the 1st, 2nd, and 5th law are universal for all migration (and generally spatial social and economic processes, like trade), and during the 20th century, they were operationalized as the gravity laws, and later gravity type statistical models of migration. Gravity laws, which will be described thoroughly in this introduction later, state that large population attracts migration, while long distance between locations makes it less frequent. Also, the “step-by-step” process in the 2nd law got operationalized into the step migration model, according to which migration tends to happen one step upwards the urban hierarchy – an analytic null model for sociologists and demographers. Ideas covered in the 6th and 7th law ignited thinking about migratory potential based on societal context, which are widely covered in

works of economists focusing on migration. The 4th law of counter-currents resonates with contemporary migration theorists who focus on systematic migrant-network frameworks, like de Haas and Massey, described in the Systematic Theories of Migration section of this chapter. But before the contemporary advancements, it has to be described how migration research took a non-relational turn after Ravenstein.

Figure 1: The currents of migration in the UK, from Ravenstein 1885.



1.1.3. The migration transition hypotheses

The unfading popularity of Ravenstein can be understood from the large scale and holistic nature of his work. But this holistic nature still was only able to cover the very age and space he was focusing on: the UK during the last third of the 19th century. That age had some significant social processes affecting internal migration currents. In his second law 1885, centers of absorption are described not just as larger settlements with great weight in commerce and industry, but also: growing. And this growth process of these centers is what gives them absorbing power, nothing else. But what happens in other ages and societies, where such growth is either absent, or weaker?

During the last century, only one widely accepted large scale migration theory emerged as both an extension and a criticism for Ravenstein, the migration transition hypothesis. Here, large scale refers to an analytical strategy which does not aim to show abstract mechanisms independent of context, but one that exactly deals with social and economic history, and several stages and contexts of those simultaneously. The model was developed in 1971 by Wilbur Zelinsky, a geographer and demographer. According to the model, human migration patterns are directly influenced by the demographic and economic changes that take place in a society as it progresses through various stages of economic development (Zelinsky 1971, Faist 2013).

The migration transition model is composed of five stages, each of which corresponds to a different level of economic development. The first stage, known as the pre-industrial stage, is characterized by low levels of both migration and population growth. In this stage, people are typically tied to their home regions and do not migrate much due to the limited opportunities available. As societies progress to the second stage, which is the transitional stage, migration rates begin to rise due to economic growth and urbanization. People migrate from rural areas to cities in search of better job opportunities. Ravenstein described migration streams in the UK during (or slightly after) the second industrial revolution. The step-migration null model, where every migrant moves one step higher the urban hierarchy represent this process of industrialization, when agglomeration economies are created in order to maintain the pace and scale of economic change.

The third stage of the migration transition model is the industrial stage. In this stage, industrialization and urbanization are at their peak, and migration rates continue to rise due to the expansion of urban labour markets. At this stage, people move from one urban center to another, seeking better jobs and higher wages. The fourth stage is the post-industrial stage, where migration rates start to decline as the economy shifts towards service-based industries. Vigorous movement of migrants from city to city and within individual urban agglomerations, especially within a highly elaborated lattice of major and minor metropolises is observed. A large increase in urban to suburban migration can also occur. Stage five is a prediction for the future, when according to the model, almost all migration happens in cities or between them.

We are in stage four now, and Zelinsky's model turns out to be accurate by and large. But as with every large scale social theories, this accuracy is abstract, and cannot be tracked to concrete structures. Nevertheless, the migration transition hypothesis laid the direction of migration and urbanization research of the last decades in developed economies, such as in Hungary. New centralities emerged from migration streams, that are unexplainable by gravity, push-pull, or step-migration models. Urbanization theory tries to deal with these processes through polycentric urban development theory. But generally, it is interesting to point out, how the migration transition hypothesis also describes a network model for developmental stages when economic growth and centralization cannot drive significant absorption. For example: “a highly elaborated lattice of major and minor metropolises” is almost explicitly network-theory terminology. A lattice is a special type of network, and while it is certain that Zelinsky is not referring to that exact mathematical term (a grid-like network structure), we can see how he predicted that in post-industrial societies, complex internal migration mechanisms can only be understood through relational analytical frameworks, and not through the dominating influence of central places.

1.1.4. Push-pull theories and gravity models – independent units and independent dyads

While the origins of migration theory from the 19th century, and macro-theories from the 20th century all treat migration, and specifically internal migration as a systematic social process with interconnected agents, localities, and effects; theories of migration that successfully turned into the mainstream framework for applied quantitative research generally lack systematic elements. This is especially true for research that aims to develop models that explain aggregated migration behaviour between localities, and do not model individual behaviour. While it is hard to measure which theory is the most widely cited, push-pull and gravity theories of migration are generally considered as one of the most widespread analytical framework to understand net migration in localities and flow between them. It is important to point out, that while push-pull modeling is widespread, it is not a unified framework, but a fuzzy set of applied practices in a sense, it is only a most-accepted best practice, owing perhaps to its versatility and simplicity. I do not want to argue that every research practice in this set lacks a systematic element - rather,

that some applied and pedagogical uses reduce migration drivers to additive origin and destination attributes, thereby neglecting relational and systemic structures. And crucially, being the most accepted or most easily operationalised framework is not the same as being the best fitting one: the empirical chapters of this dissertation ask precisely whether models that omit interdependence fit observed migration systems.

It is hard to point to one paper where the core idea of push-pull thinking first appeared, because it is so intuitive: where people have attractive reasons to move, they move there, and where there are repulsive reasons to leave, they leave. This simple intuition resonates with the work of Ravenstein: big centers of commerce and industry have a pulling force for people in other localities (Ravenstein 1885). Everett Lee's seminal 1966 paper systematically reformulated Ravenstein's insights into a comprehensive theoretical framework, focusing on pulling and pushing forces. Lee identified four factor categories affecting migration: factors associated with the area of origin, factors associated with the area of destination, intervening obstacles, and personal factors. He proposed a plus-minus calculus for migration decision-making and derived hypotheses regarding migration volume, streams and counter-streams, and migrant characteristics. Lee emphasized that personal factors such as education, knowledge of destinations, and family ties mediate how individuals respond to push-pull forces (Lee 1966). This paper is still cited as the canonical foundation of push-pull theory.

Nowadays, in applied aggregate-level research push-pull thinking is typically operationalized as a plus-minus calculus of economic and social factors, resulting in a net migration value for a location. But the 1966 paper of Lee had strong elements of systems thinking – as the paper is kind of a direct expansion of Ravenstein. For example ideas like “*Migration tends to take place largely within well-defined streams*”, or “*For every major migration stream, a counter-stream develops*”. Even the plus-minus calculus is embedded in migration streams in his model: “*The efficiency of the stream (ratio of stream to counterstream or the net redistribution of population effected by the opposite flows) is high if the major factors in the development of a migration stream were minus factors at origin*”, etc. (Lee 1966). We can see how path-dependency, or the structure of migration flows, ideas well-described in later parts of this Introduction chapter appear.

While Lee is well cited as the foundational author of push-pull theories, push-pull models mostly miss the systematic element of his model, the fact that migration inducing or reducing factors are embedded in a flow structure with external determinations (i.e. economic structure of a nation, or the urban hierarchy), and work according to inherent mechanisms (i.e. reciprocity). It is hard to cite any paper when summarizing a research tradition as a whole, but if we read the most widely cited comprehensive reviews of the area (Greenwood 1975, 1985, 1997, Lucas 1997, Cushing and Poot 2003, Molloy et al. 2011, Etzo 2008), we can see that the push-pull calculus mostly refers to the migration enhancing or reducing qualities of individual units. They cite housing market opportunities, employment opportunities, amenity quality (like schools or hospitals) working on the level of the given unit of aggregation. In this introduction section, and in the paper of this thesis about the housing boom period restructuring of the internal migration system of Hungary, I call this approach the “individual unit” approach.

The comprehensive reviews cited above also cover one widespread approach to model migration streams between two localities, system decay, or when its formalized mathematically, gravity models. The general pattern that as distance grows between two locations, migration becomes less likely can also be found amongst the laws of Ravenstein, but he did not try to formalize this mechanism. The first formalization is coming from Zipf’s 1946 paper, “The $P_1 P_2/D$ Hypothesis: On the Intercity Movement of Persons”. The title summarizes the mathematical formulation that is still mostly used in research, although with constant minor adjustments. Zipf’s paper was shortly followed by Stewart’s 1948 paper on “demographic gravitation”, which provided deeper theoretical foundations. In following decades, the term “gravity model” was established as a general framework for spatial interactions between locations, and not just for migration studies, but for economics, transportation, etc. In his 1975 review, Greenwood treats the gravity model as a mathematical model that achieves proper accuracy for flow modeling for migration. According to the basic gravity law, populations in origin and destination locations raise the probability of migration (raise flow volume), but their product has to be divided by the straight geographical distance between them. As we see, gravity models are about estimating dyads of migration, but they treat every dyad as an individual, thus in this introduction and the housing boom restructuring article, I call

this approach the “individual dyad” approach. While gravity models are rarely used purely nowadays, gravity type variables are widespread in statistical modeling of flows as important controls, like in two of the three articles of this thesis.

To understand why gravity models are not just statistical controls, but the core mechanism is inherently meaningful on a sociological, theoretical basis also, it is good to look at the work of system theorists of migration. In his foundational 1970 paper, Mabogunje makes an interesting argument for treating certain theoretically derived migrant-distributions across cities as most probable random states. His argument lies in systems thinking, and especially information theory. Mabogunje argues that if we imagine a process of rural-urban migration in which every rural migrant who leaves the village simply and totally forgets it, information feedback loops cannot build up between the eventual destination of the migrant and the origin. This way the feedback loop cannot make certain migration decisions and directions easier for later migrants, who, in effect, will choose any destination of their liking (or end up somewhere by chance). Over time, this process would lead to a state in which every city has proportional amount of migrants from every village to the size of the given city. According to Mabogunje, this is "*the most probable state in which no order or organization is evident in the system*", in other words, a state of maximum entropy (Mabogunje 1970). One obviously missing element of this ideal-state is distance - chance of migrants getting to cities would also depend on the distance between origin and destination. If we expand the ideal-state with distance, we exactly get a pure gravity system. The argument of Mabogunje shows that gravity-networks can be seen as ideal-typical migration networks in the Weberian sense, which result from a process of migration without information, feedback loops, socio-historical determinants of established pathways. Thus comparison between such synthetic migration networks and observed migration networks exactly serves as a way to differentiate social order in internal migration systems. This is the core idea of the first article in this thesis.

Individual unit and individual dyad explanations however do not answer two seminal questions of migration research: which specific destinations will experience migration inflow, and what mechanisms will direct migration exactly to those locations. To give a concrete example, related to the third essay of this thesis, it is a necessity that the agglomeration of large cities experience population outflow from

the city - differentials in density, house price, rent, etc simply predict the inevitable. But as it is demonstrated in the article, not every location in the agglomeration experiences the same inflow - some locations experience population booms, some are almost unaffected. As Hein de Haas puts it: "*although the truism holds that economic and other opportunity differentials generally play a major role in migration, this alone cannot explain the actual, patterned and geographically clustered morphology of migration, typically linking particular places and regions*" (de Haas 2010). It is important to note that push-pull theories, or gravity models do not promise descriptions for exact patterns and relations, they average behaviour across the entirety of migration systems. They work on the aggregate, the population level, thus omit important detail. This is why after decades of migration research based on the aggregate a logical step forward is to understand the patterns and determinants of directions - which can easily be operationalized as migration flow networks. But before migration flow networks can be used as a heuristic, the possible micro- meso- and macro-level mechanisms that build up and sustain flows have to be addressed - which is what system theories of migration aims to do.

1.1.5. System theories of (internal) migration

System theories of migration are explanations of migration behaviour that focus on mechanisms that work as feedback mechanisms, through which we can understand how migration flows emerge and stabilize over time, and what origin-destination pairs develop resilient connections. The origin of this theoretical tradition is migration system theory (MST) developed by Mabogunje 1970. The theoretical work I describe in this section as system theories of migration and MST are not interchangeable, the former is the extension and generalization of the latter. Interestingly, while system theories almost always work with empirics of international migration, MST was originally developed to understand rural-urban migration in Africa.

Mabogunje builds MST on existing general systems theory. He defines a system as a complex of interacting elements, together with their attributes and relationships. Important to note, that an element in Mabogunje's formulation is a sub-system, an adjustment mechanism, or a migration channel – in this sense, original concept of

MST is quite different from papers that treat flow-networks as migration systems. Such work defines elements as locations, mostly settlements, and not social and economic institutions and mechanisms. A sub-system according to Mabogunje can be the family (by determining which member moves, and in what age can that happen), or the rural community itself. Such elements make the entirety of the rural control subsystem. The urban control subsystem includes city administration, employment agencies, generally mechanisms that facilitate attracting power towards potential rural migrants.

While MST is a macro-level theoretical framework, building explanation with societal and economic level feedback and control mechanisms, they are all linked to the individual. This way, the system can be understood through not just the propensity to migrate, but “*migration elasticity*”, the time impulses and stimuli have to work to trigger an individual migration event. Another important heuristic is energy – the stimuli acting on potential migrants. Towards this constant stimuli different individual reactions can happen, this way the question of who migrates can also be answered, in contrast with push-pull models. But the most important heuristic in MST is information – the idea that after an individual migrates, its effect on the whole system is amplified by the information that is sent back to the origin community. Such information makes future moves easier, and defines location choices of future migrants. Without information feedback, migration systems would converge to a most probable state, where migration flows are purely controlled by settlement size in destinations.

The core idea of MST is that through information feedbacks, migration has a self-perpetuating nature. It is important to note that MST was not developed as a closed theoretical framework – Mabogunje assembled a patchwork of different institutions and mechanisms, suggested that they interact and enhance each other’s effect, but did not provide an exact mechanism how that happens. This openness is not necessarily a weakness: migration systems are complex, multi-scalar, and historically variable, and a closed framework could impose false rigidity. MST is theoretically generative but operationally underspecified. It therefore invites selective formalisation – empirical tools such as network analysis can operationalise some of its mechanisms in a limited, testable way, which is precisely what this dissertation attempts.

Cumulative causation theory (CCT), introduced by Myrdal 1957, and reintroduced and developed as a standalone research tradition by Massey places this effect in the core of the explanation (Massey 1990, Massey et al. 1993). The core idea of cumulative causation is that migrant networks affect future migration processes between two locations. Migrant networks are sets of interpersonal ties that link migrants and former migrants, and people in the origin and destination communities. Migrant networks include ties of kinship, friendship, and shared origin. Each act of migration creates social capital for others to whom the migrant is related, reducing the costs and risks of future movement. As the number of network connections in an origin area reaches a critical threshold, migration becomes self-perpetuating because each act of migration creates the social structure to sustain additional movement. It is important to note, that Massey does not treat cumulative causation of migration solely as a social network phenomena – important meso and macro level social and economic effects also help it emerge. For example, as described by Stark and Taylor 1989, the relative deprivation of households can be solved through the household sending migrants to a new location for new resources and financial gains. But if this strategy succeeds for one household, it will raise relative deprivation in the community, leading to more relative deprivation thus more migration. Or as Greenwood describes, international labour migration many times filter for migrants with high human capital. That can create new workplaces through economic growth in the destination, which attracts more migration. Massey treats the idea of migration networks, and network-mediated community feedback as an integrating theoretical mechanism, that can connect fragmented explanations of migration (Massey 1990, Massey et al. 1993).

Empirically tested by Massey and colleagues for example for Mexican migration to the USA, having a parent with U.S. experience increases the hazard of first migration by approximately 40 percent, while sibling migration experience increase migration hazards by 60-80 percent depending on specification (Palloni et al. 2001). However, testing cumulative causation across different-sized communities show that network effects attenuate with urbanization. In rural communities and small towns, the migration prevalence ratio strongly predicts individual migration, with each percentage point increase in community prevalence raising migration odds by 3-5 percent. In cities above 100,000 population, this effect is much weaker or becomes

negative. Family network effects persist across all community types but are weaker in urban areas (Fussel and Massey 2004).

Hein de Haas synthesizes the internal dynamics of migration processes as an extension and integration of both MST and CCT. For de Haas, MST operates purely on the premise that previous migration defines future migration through information feedback loops, connecting origins and destinations while stabilizing flows over time, but lacks crucial contextual feedback mechanisms. These contextual mechanisms encompass not only migration-inducing effects, but processes whereby migration itself transforms the sending population, increasing the likelihood of future migration: social stratification, relative deprivation, and material and cultural remittances. While CCT attempts to incorporate these contextual feedback loops, it focuses exclusively on positive effects and neglects negative ones that cause flows and migration systems to decline. De Haas argues that migration systems can collapse without exogenous shocks like restrictive migration policies.

Conceptualizing migration as network diffusion reveals an inherent saturation dynamic: early moves generate large and growing positive returns, but once a migration network fully emerges, these returns diminish. Though this doesn't guarantee collapse – later moves remain easier with lower thresholds – it can produce declining migration frequencies (de Haas 2010, de Haas 2021).

System theories of migration, besides the original work of Mabogunje, almost always refer to international migration, like the above cited works of Massey and de Haas, and the synthesis of these theories with internal migration is still lacking. This can be attributed to the simple fact that international migration was in the forefront of migration research for the last decades, but also to the fact that the system-defining contextual feedback mechanisms the literature describes can be observed more clearly in international migration, or as Massey 1990 puts it, the self-reinforcing dynamism of migration is particularly pronounced for international migration. For example, the information and material thresholds that get reduced with migration chains through cumulative causation are more pronounced in an international migration event than in an internal one. If there is a steady flow of migrants between two distant countries, maybe even on different continents, the effects of feedback loops and cumulative causes can be seen crystallized. On the other hand, an intra-metropolitan move does not mean that the individual or the household changes its

whole social context; many times the workplaces, the schools, the social milieu remains the same. But even for an inter-regional move, the language barrier is non-existent, and information about the destination is automatically gained for example from media coverage.

However, there are feedback loops well documented in other parts of academia that show the systematic, feedback-driven nature of internal migration. For example in urban sociology, gentrification refers to the process when people with more capital appear in an otherwise deprived area, the mix of amenities starts to shift according to the new money in the neighbourhood, attracting similar residents from that point. Slum formation on the other hand means that if more affluent people leave an area on a downward trajectory, amenities, institutions, capital will follow them, making more people leave. Such feedbacks work on the regional level also, meaning that if a rural, poor region starts to lose population, it makes workplace creation and development less likely. Or for housing markets, one of the main drivers of shorter distance internal migration, falling house prices can lead to a lock-in effect, while rising prices enable the movement of homeowners through growing housing capital. But such feedback mechanisms of internal migration are usually connected to exogenous shocks, like the economic cycle, market busts and crises, or the collapse of some type of industry that defines the local job market of a region. The kind of theoretical work Hein de Haas developed for international migration, which focuses on endogenous system formation and erosion, is still missing for internal migration – maybe because it is less pronounced on this scale of spatial mobility. This shows that migration theories are maybe ought to be scale-dependent first, which sheds light on why settlement level internal migration networks are important to model and understand in themselves.

De Haas defines a migration system as a set of places that are linked by flows and counter-flows of people, goods, services, and information, which tend to facilitate further exchange. Operationalizing such a system can be easily done as an inter-location flow network – that works on the aggregate. Such formalizations already exist for international migration systems, for example the works of DeWaard et al. on differentiating migration systems in Europe through the analysis of harmonized flow data (DeWaard et al. 2012). It is important to note, that such networks are not enough in themselves to make analysis systemic – the original formulation of MST from

Mabogunje showed that systems are made up from emergent mechanisms, not patterns on the aggregate level. However, structural analysis of the patterns in these network can serve us as theory selection both on the micro and macro levels.

Before turning to contemporary urbanization research, a short historical bridge has to be built, because the two literatures this introduction connects - migration theory and urban research - are in some sense the same history, read from two different sides. Cities are accumulated migration. Ravenstein's centers of absorption describe urban growth as an outcome of flows; urbanization research describes the same growth as an attribute of places. Urban and migration historians worked exactly on this seam, and while this dissertation cannot cover their tradition, the most foundational works have to be pointed out - because the traces of flow network thinking are at least as strong in them as in the theoretical classics discussed above.

Traces of network thinking appear in the founding texts of migration sociology itself. Thomas and Znaniecki in *The Polish Peasant in Europe and America*, analysed peasant migration through letters, kinship obligations and community ties: migration organized by pre-existing social relations, decades before a formal network vocabulary existed to name them (Thomas and Znaniecki 1918). In another field-defining work Redford 1926 reconstructed labour migration in industrializing England, and found the same short-distance, stepwise currents rippling towards the absorbing centers that Ravenstein had described statistically forty years earlier. We can see that the systemic reading of Ravenstein argued for in this Introduction is not a retrospective projection: his contemporaries and immediate successors read internal migration similarly.

The historians also formulated the micro-mechanism behind path-dependent flows. MacDonald and MacDonald 1964 described chain migration for the arrangement in which prospective migrants learn of opportunities, and obtain transportation, housing and employment, through primary social relations with earlier migrants. Their opening already rejects the push-pull calculus as a "*mechanical reshuffling of heads*" - the same criticism this dissertation formulates against individual unit approaches. Chain migration can be interpreted as the path dependency of flows, observed on the scale of families. Tilly 1990 condensed this whole tradition into a single radical claim: the effective units of migration are neither individuals nor households, but

networks. The central argument of this dissertation was, in this sense, formulated from the archives long before it could be formulated from adjacency matrices. This once again shows how even non-sociological fields can be added to the patchwork-like theoretical background of internal migration studies, and that the network perspective can serve as a unifier even for these more distant scholarly traditions.

And the historians did not stop at the micro-mechanisms; they also worked on the system level. De Vries 1984 analysed European urbanization between 1500 and 1800 quantitatively, as one interdependent urban system, in which cities behaved as demographic sinks: their growth was fed by continuous rural inflow, and the fortunes of individual cities depended on their position in the system rather than on local conditions alone. Moch 1992 synthesized the history of Western European migration since 1650, and organized the whole synthesis around migration systems and networks, showing that mobility was a structural, network-channelled feature of European societies well before industrialization. What this tradition lacked was, again, not the systemic theory but the instrument: settlement level flow data, and formal models of flow interdependence. The parallel with the urban systems research of the 1960's and 70's, discussed in the next section, is not accidental: historians and geographers arrived at the same research program from two different directions.

The clearest bridge between the historical literature and the urbanization theories of the next section was built by Hohenberg and Lees 1985. They organized a millennium of European urban history around the duality of two ideal-typical systems: the Central Place System, in which settlements are ordered into a hierarchy of nested market areas, and the Network System, in which cities are nodes of long-distance flows of goods, people and information, and their importance derives from their network position, and not from their local hinterland. Importantly, they show that the two systems did not replace each other, but coexisted and interacted for a thousand years. This can be seen as the conceptual pair (hierarchy and network) that contemporary urbanization research rediscovered in the polycentric development debate, and that the first article of this thesis measures in the multiscale structure of Hungarian migration flows. The next section shows how this duality became the central question of urbanization research in the present.

It is important to state precisely what this short summary can and cannot claim. The problem is not that internal migration systems have not been studied – urban and migration historians studied precisely the mechanisms this dissertation is interested in. The problem is that their findings remain fragmented across disciplines, and have rarely been integrated into formal migration theory or network modeling. This dissertation cannot achieve that integration either, and can only present important examples for the historical literature. But the instruments it develops - null models, interdependence terms, tie-based measures - are exactly the kind of formal tools such an integration would require, and I hope the three articles can serve as small steps towards a research direction where the archive and the adjacency matrix can finally be connected.

1.2. Inter-urban flow networks

Up until this point, this introduction focused on the history and current developments of theories focusing on migration itself, as an endogenous system that can be explained by its own special rules and emergent structure. However, internal migration is also analysed by other fields, which on the surface focus on different phenomena, but in their core, population redistribution is the main mechanism that drives effects and mechanisms they explain. Most importantly, some aspects of studies focusing on urbanization have to be summarized, because as it is clear from previous chapters, cities and towns are the main drivers of internal population redistribution across the world. Also, during the last decades, urbanization studies shifted theoretical focus on polycentric structures, describing cities and flows between cities as a network-like system. The empirical shift following this theoretical trend also lead to the emergence of a new network science of inter-urban flow networks. This literature, mainly driven by the spillover of natural scientist to the study area of societies applies state of the art mathematical network science tools to model interactions between cities, and a fair share of this literature is focusing on migration and generally population redistribution flows between settlements. Often, this new empirical tradition does not connect to the system theories of (internal) migration, or to the foundational texts like Ravenstein, leading to an interesting

parallel world of different literatures focusing on similar phenomena, and using theoretically similar explanations. The main difference dissecting these parallel worlds is that while the system theories of migrations focus on the individual, and explains the flow structure between locations as emergent individual behaviour, the new network literature of inter-urban flows only focuses on the localities, the aggregate.

1.2.1. Current trends in urbanization: Polycentric development

Research on human settlement and urban structure largely originates from spatial thinking. This does not only mean the recognition that physical proximity enhances social, cultural, economic connectivity, but the very idea that research ought to use some spatial structure as an analytical ideal type in the Weberian sense. The most influential early example was central place theory, formulated by Christaller 1933 for the settlement system of Southern Germany, and formalized in economic terms by Lösch 1940. Strictly speaking, central place theory is a formal spatial-economic model rather than an ideal type in the exact Weberian sense, but it functioned as one in the literature: an abstract benchmark of spatial organization, against which observed structures can be compared. It is also one model within a broader tradition of spatial analysis - on the intra-urban scale, a parallel sociological tradition existed from the Chicago School onwards (Park et al. 1925). But the object of this dissertation is different: not the city, but the system of settlements and the flows between them. And for this inter-urban scale, central place theory and the urban systems research that grew out of it form the relevant classical lineage. The intra-urban tradition therefore falls outside the theoretical framework, although it can help interpret some of the intra-metropolitan results of the third article.

During the decades, different formalizations gained attention for economic and social research, urbanism and development - generally they gave an interdisciplinary foundation for space-related social sciences and policy. Thus many different formulations and definitions are available, but a general picture can be summarized: localities where human activity concentrates heavily (often called central business districts, indicating the attention on economic aspect) influence and define mechanisms in every other locality; and this influence decays with distance. This is a system way of thinking since the empirical observations are analysed as an outcome of the influence of the central place on all other places.

Central place theories led to a boom in urban system research during the 60's, 70's and 80's. It is important to note that this tradition already had an explicitly relational program. Berry 1964 defined cities as "*systems within systems of cities*": in his formulation, a settlement is not an independent observation, but an element whose growth and functions depend on its position among all other settlements. Pred 1977 went even further. In his model of city-systems in advanced economies, growth is transmitted between cities through flows of specialized information, and these flows form feedback loops: the growth of one city changes the growth prospects of the cities it is connected to, which changes the flows again. We can see that two core ideas of this dissertation, interdependence and feedback operating through inter-urban flows, were formulated decades before formal network science methods emerged.

This boom ended with the last decade of the century, firstly because while central place theories offer a meaningful tool for analysis they are insufficient to explain the detailed structures and differences of spatial patterns of human activity (Coffey 1998). There is a second reason for the fading of urban system research of the 80's besides the arguable analytical insufficiency of central place models: that urban development itself started to behave differently. The power of central places started to decay, a new structure started to emerge. This new structure is analysed from multiple perspectives. In the United States, the empirical starting point was the identification of employment subcenters: metropolitan regions like Los Angeles turned out to contain multiple concentrations of employment instead of one dominant core (Giuliano and Small 1991). In Europe, the change was first described in the language of city networks: Camagni and Salone 1993 show how Northern Italian cities form cooperation-based network structures that do not follow the central place logic of size and distance. In sociology, several parallel accounts of the new structure emerged: the time-space compression of Harvey 1989, or the global city of Sassen 1991. For this dissertation, the most important of these accounts is the space of flows concept of Manuel Castells: he described that globalization and the availability of functional telecommunication techniques led to a new geography of centers and peripheries. Castells argues that in the age of information society (vaguely the last few decades), the space of places (geographical reality) and space of flows (telecommunications) interact with each other, and this interaction defines what

becomes a new center, or which old center collapses into the new periphery (Castells 2020).

These recent trends of urban development gained attention in different forms in interdisciplinary urban and regional research, and a new paradigm started to emerge: polycentric urban development (Kloosterman and Musterd 2001). Polycentric development theories state that because of functional differentiation urban functions (business, residential, industrial, etc) evolve through different center-subcenter-hierarchy structures. One settlement can be a center for many functions, or just one, and every settlement can be placed somewhere in the center-periphery gradient of every function. In other words, a network of settlements replaces a strong hierarchy of settlements (Vasanen 2012). However, this replacement is not the disappearance of hierarchy: polycentric development produces hierarchical networks, in which centrality is distributed unevenly across nodes and functions. The first article of this thesis exactly demonstrates this coexistence – internal migration flow structure in Hungary is defined by polycentricity, but also by the classical pull force of the growing capital city.

Meijers 2007 summarizes two descriptions for this paradigm change from two influential theorists of polycentric development. Batten describes the change from central place systems to network systems through the changes from centrality to nodality, from size dependence to size neutrality, from vertical accessibility to horizontal accessibility. Van der Knaap describes the change from hierarchies to networks through the change from a fixed number of spatial scales to variable number of spatial scales, from economic function rising with scale to variable sets of economic functions on the same scale.

However, before connecting polycentricity to migration research, it has to be clarified that the concept is used in at least two fundamentally different senses. First, polycentricity is not only an analytical concept, but a normative one: in European spatial planning, polycentric development became an explicit policy goal, a desirable spatial structure that planning has to achieve (Davoudi 2003). This dissertation only deals with the analytical sense. Second, and more importantly, analytical polycentricity itself splits into a morphological and a functional interpretation (Burger and Meijers 2012). Morphological polycentricity describes the distribution

of nodes: how evenly size and importance are distributed across the centers of a territory. Functional polycentricity describes the structure of flows: how evenly relations between settlements are distributed, independently of the size distribution. Burger and Meijers 2012 show that the two do not necessarily coincide: a region can be morphologically polycentric, while its flows remain oriented towards one dominant node. This distinction is central for this dissertation, because functional polycentricity is defined and measured on flow networks. The largest empirical program in this direction, the POLYNET study of eight European mega-city regions, operationalized functional polycentricity exactly through flows of people and information between settlements (Hall and Pain 2006). This way, when this dissertation models internal migration as an inter-settlement flow network, it works inside the functional polycentricity tradition.

As it can be easily seen, the idea of polycentric development is vague, and developed to summarize a multitude of processes in different geographies around the globe. Parr 2004 provides a critical inspection of the concept: the term is applied to structures on very different spatial scales, with criteria that are rarely stated explicitly, and this way it risks becoming a label rather than a model. The use of the concept of networks in polycentric development research is also heterogeneous, most of the authors did not aim for formal definitions (van Meeteren et al. 2016). This is surprising since social network theory, and network science in general, experienced a parallel boom with the new paradigm of polycentricity. Pairing of these separate literatures is a natural fit, since network science offers the exact definitions for urban and regional social scientists. There are examples available in the literature, like for example Green 2007 who aims to set up formal definitions with tools of social network analysis, or Zhong et al. 2014 who aim to detect urban dynamics with spatial network analysis. However, none of these experiments for formal definitions became an universally accepted framework in the literature.

The more substantial reason probably is that these experiments relied heavily on concrete aspects of polycentric development, concrete urban functions. As Burger et al. 2014 point out, urban networks are not simple networks but multiplex ones. This means that the network of one urban function do not necessarily overlap with the network of another. This way the models of spatial organization of urban regions heavily depend on the data that is available for researchers. This way a formal

definition can be precise and insightful, but imprecise for another. Based on this recognition two possible research strategies emerge: one that aims for a holistic view handling multiplex functional networks; and one that aims for formal models in only one function, like this dissertation aims with the residential migration function.

Internal migration research is strongly connected to the research of the residential urban function, thus functional polycentricity. The classical example for this connection are agglomeration towns around large cities, settlements with high number of single family homes, and low number of workplaces. These towns attract high internal migration, but do not attract institutions or working capital. In the vocabulary of the network literature, such settlements have borrowed size: they gain population functions through their network position, and not through their own local weight (Meijers and Burger 2017). The article in this thesis about social networks and intra-agglomeration migration around Budapest exactly shows how the internal migration can work in the absence of amenities and local economic power.

1.2.2. The new macro level network literature of migration

White and Lindstrom in their seminal handbook chapter for internal migration define five main themes in contemporary internal migration research of “*more developed countries*”: population diffusion, urbanization and counter urbanization, intra-metropolitan population distribution, structure of migration streams, and policy concerns (White and Lindstrom 2005). These themes show that internal migration research pays growing attention to processes that are more fine-grained than classical region-based analyses, either in spatial granularity (intra-urban migration), or theoretical detail: not only focusing on individual streams between localities, but defining and modeling some kind of structure of streams. This led to a new, emerging literature, which aims to explain (internal) migration better by the application of network science to the field. Recently, the idea started to gain momentum, with the number of studies dedicated to migration networks tripling through the last decade (Pitoski et al. 2021b).

For example, Aksoy and Yıldırım 2024 criticizes the gravity models from the aspect of interdependence. They show that gravity-thinking only focuses on individual streams between two locations, but does not consider that if a person migrates from

one location to another, that person mechanistically cannot migrate to a third location. Likewise, the characteristics of the third location can affect the migration stream between the first two, by being a competitor for immigrants of one of them. By the application of network analysis, such effects are naturally incorporated in the modeling process. The explicit recognition of the importance of interdependence in internal migration systems appeared in early pioneering modeling works, for example in the work of William Alonso. Alonso states that unless the interdependence of a migration system is explicitly considered, important variables will be omitted or assumptions will be made implicitly that would not be readily granted if made explicit (de Vries et al. 2001). Despite the inherently relational nature of Ravenstein's laws, and early recognitions like Alonso's, and other outstanding isolated previous research examples, there was little research in internal migration from a network perspective before the 2010s (Desmarais and Cranmer 2012).

Gravity models can be integrated into the network perspective in a general sense, since practically these are tools to predict the weight of edges in migration networks that assume the absence of any kind of network effect. In this sense, the real question is whether their performance is satisfying enough, or can it be exceeded by different methods. For example Pitoski et al. 2021b tests the gravity model in the case of Croatian internal migration, and they find that gravity models structurally overestimate low migration links, the link-weight predictions of these models fit quite imperfectly to the observed distribution, and deployed models are only able to explain about half of the variance of real migration. This predictive shortcoming is one of the strongest argument for introducing network effects into internal migration analysis. Modeling network effects does not require explicit network models, it can be done by simply expanding gravity-type models with network variables. Sardadvar and Vakulenko 2020 tested a gravity-type specification with network effects for the internal migration system of Russia. They approach the idea of network effects from the tradition of analysing spatial spillover effects (when some process in a locality affects neighbouring localities) and competition effects (when some variables decreases in a locality if it increases in neighbouring localities). Based on these ideas, they define network effects destination and origin spillovers, and destination and origin competitions.

Including network effects in more conventional regression modeling frameworks is a research strategy that connects naturally to previous gravity-based frameworks. Despite excellent examples following this strategy, in my view a more promising approach is to focus on the whole system instead of its individual components (that are affected by network effects), and explicitly look at migration as a network. This way, network mechanisms do not need to be translated to variables useable in classical regression models, but the whole range and content of network theory can be applied easily. A recent example for this approach is the work of Pitoski et al 2021.b, who look at the internal migration system of Croatia at the inter-settlement level. They point out that settlement level analysis is rare, but can be seen as optimal granularity for internal migration research. They abstract the migration system as a weighted loop directed graph, a network definition that does not bring information loss (for example when migration streams get transformed to binary edges). Loops are important since intra-locality movements contribute to a large share of internal migration. The network statistics they use are: centralities (i.e. node strength, degree centrality), eigencentrality algorithms for the identification of migration hubs, and network structure measures (reciprocity, assortativity, transitivity), and statistics used for community detection and modularity. Through reciprocity analysis they show how internal migration contributes to urbanization and which regions suffer significantly losing population through migration; through community detection algorithms they show which settlements form clusters that tend to contain migrants; and with eigencentrality analysis they show which cities are the extreme attractors of migrants (Pitoski et al. 2021b). In my view, this study is a perfect reference for the application of network science to internal migration research. Other inspiring examples for different geographies are Charyyev and Gunes 2019 or Xu 2018 for the USA (the country with the most related research), and Li et al. 2017 and Sun and Pan 2014 for China. In the EU, this kind of research is still very rare. Pitoski et al. 2021b for Croatia or Pitoski et al. 2021c for Austria shows that such research for European geographies started to get published recently.

The theoretical usefulness of descriptive network statistics should not be taken as granted. While they will show us macro level patterns, the interpretation of these patterns is hard. It is a question whether we can use these patterns as empirical tests for sociological theories; or these new tools are merely a telescope – something that

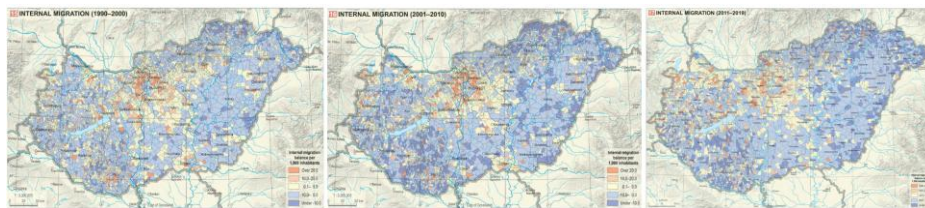
shows us new domains of a system, without telling us why that domain looks the way it does. Descriptive network statistics show us properties of the macro structure that can be interesting for the sociology of migration, but do not let us explicitly test hypotheses about causes. Cause here means a special kind of causality, namely generating mechanisms of network connections. These kind of causes are only visible through generative network models.

An important method for statistical inference in networks is the exponential random graph model (ERGM) which estimates the probability of an observed network conditional on a vector of network statistics. Beside network statistics, the model can contain exogenous covariates also (Holland and Leinhardt 1981). ERGM is restricted to only binary ties. This would not prohibit migration analysis (migration counts can be easily transformed some kind of binary structure), but in the case of internal migration, binary edges mean information loss compared to valued edges networks. Desmarais and Cranmer 2012 tackled this problem by developing the Generalized ERGM model (GERGM), which they applied to internal migration of the USA. They model (the change of) interstate migration from 2006 to 2007, and include dependence terms like transitive triads, reciprocity, etc. and they also include node-level covariates like population and unemployment rate. This model specification is a good example how GERGMs are able to infer the probability of an observed migration network conditional on both interdependence structure and classical migration theory factors at the same time. ERGMs were utilized for analysing the international network of refugees (Abramski et al. 2020), the network of global migration (Windzio 2018), the internal migration of the US (Desmarais and Cranmer 2012), intra-EU migration flows (Windzio et al. 2021). Block et al. 2022 developed a slightly different version of ERGMs for weighted networks than GERGMs. They argue that their model is useful for every kind of mobility table based analysis, and they show its usefulness for academic hiring and mobility processes. This shows that the idea of interdependency will be utilized for mobility independently from geographical space in the future.

1.3. The Hungarian internal migration system after the regime change

Massey describes one of the greatest divides in migration research as a divide between historical and abstract explanations. Foundational scholars like Wallerstein or Petras theorize migration as something so fundamentally historical and context-dependent, that it cannot be abstracted into general mechanisms (Massey 1990). The three articles that make up this dissertation, as well as the theoretical and research tradition they aim to connect to, treat migration as a social process that can be modelled generally, as a result of context-independent mechanisms. Nevertheless, all of the articles in this thesis work with Hungarian data, from a certain time period,

Figure 2.: Internal migration rates in municipalities of Hungary, for three different decades.



Source: The National Atlas of Hungary.

thus this introduction has to cover the most important aspects of the national context. This is especially important, since according to common wisdom, Hungarian (and Eastern-European) internal migration is special for multiple reasons: lower migration rates than Western countries, higher share of owner-occupied housing, large scale mobility-related changes in recent history like the regime change, etc. These “special” qualities are mostly referring to deviation from Western internal migration systems, based on which most of traditional (internal) migration theory and research was built. While I try to cover the Hungarian internal migration context after the regime change, I aim to show that by the application of the new network perspective, unique Eastern European and Hungarian internal migration processes can also be treated as a system with generalizable patterns – might those differ from patterns observed in Western societies. The post-socialist contexts provide an ideal case for studying migration system transformation, as rapid institutional, social, and economic change makes path dependencies and emergent structures more visible.

The brief history of internal migration after the regime change can be summarized into few distinct periods. 1990s-early 2000s brought strong suburbanization, Budapest and county seats lost population. Migration balances became more balanced across settlement types by the 2000s; briefly around 2008-2010 Budapest had positive migration balance; between 2016–2019 Budapest again showed growing negative balance county seats also lost population; and after 2019 smaller cities and villages are net gainers (Bálint and Obádovics 2021). It is important to point out, that with the Covid pandemic, internal migration became highly irregular (strong volatility between years, both in directions and rates). The three studies making up this dissertation all focus on different periods of the 2010s decade.

Figure 2. shows the internal migration rates in all Hungarian municipalities for the three decades after the regime change. It is strikingly visible that the main source and sink regions of internal migration streams are almost the same across the 30 years: Budapest, the Budapest agglomeration, and North-West Hungary – including the Balaton. North-West Hungary became a more dominant sink region for population redistribution across the years. Based on this map, and on the literature (Bálint and Obádovics 2021, Péntes 2024), the two main processes are population centralization, and an east-west movement.

The key to understand the internal migration system of Hungary, and the centralization process that defines it, is to understand that Budapest is an extremely dominant city in its country compared to other nations. Before presenting the numbers, a definitional clarification has to be made, since different statistics rely on different spatial delimitations, and these delimitations are not interchangeable. Budapest and Pest county are administrative units, and they do not automatically constitute the metropolitan area. The official Hungarian nomenclature delimits the Budapest agglomeration as a distinct set of settlements (KSH 2014), while international comparisons rely on the functional urban area concept, which defines metropolitan areas through commuting relations (Dijkstra et al. 2019). In this dissertation, different definitions are used for different purposes. Some descriptive statistics are reported through county or region administrative units. During the citation of international sources, the functional urban area numbers are reported, as they are used by the international research community. But when I refer to the “Budapest agglomeration” in the three articles of my own, it always means the

official Hungarian definition by HCSO. Possible criticism against this decision is that the settlements in this area represent an inner circle, with a longer history of suburbanization processes, and not the whole of the agglomeration. But the debate about the best possible definition is permanent in the literature and policy circles, so any decision can be criticised from one aspect or another.

Census statistics are only available for country and regional units, and according to the 2001 census, 28% of all Hungarians lived in Budapest or Pest county, and according to the 2022 census, this number grew to 31.4%¹. While there are other European nations with large capital cities compared to the population, one key statistic to understand the primacy of Budapest is the ratio of the population of the capital metropolitan area relative to the largest non-capital metropolitan area (functional urban area). In Hungary, this ratio is larger than 9, placing Hungary the 13th in the global ranking of the OECD, with only African, Asian, and Latin American countries making up the first 12. The next European country is Lithuania in the 31st place, with a ratio around 1.3 (OECD 2020).

The suburbanization processes around the capital that enabled this relative dominance of the Budapest agglomeration area were not entirely special in the Central-Eastern-European region. While suburbanization processes started way before, in all countries of the region urban decentralization accelerated after the regime change, the urban periphery grew much faster than the metropolitan core (Sýkora and Stanilov 2014). The destinations of suburbanizing populations followed similar patterns in other post-socialist capitals, like Tallinn (Leetmaa and Tammaru 2007). New, sporadic zones of newly built detached houses and conversion of datcha-like real estate emerged on the edges of towns and villages in the metropolitan periphery (Stanilov and Sýkora 2014). The missing financial foundations for planned, centralized development, and harsh institutional changes led to urban development defined by private and market actors (Sýkora and Bouzarovski 2012). Hirt 2012 describes this general post-socialist condition as privatism: private actors filling the vacuum left by the retreating planning state. In Hungary, this suburbanization was uneven from the very beginning, both socially and spatially (Timár and Váradi 2001), and the population inflow into the official agglomeration

¹ <https://nepszamlalas2022.ksh.hu/adatbazis/>

was already measurable during the 1990s (Kok and Kovács 1999), together with the land conversion that accompanied it (Kovács et al. 2019). Many times, pre-existing smaller towns and villages on the urban periphery became sinks for urban population outflow, and rural population inflow to the metropolitan area, rather than new, large-scale developments (Stanilov 2007).

Because of both the missing central planning authority, and the pre-existing fabric of sink settlements, new population many times appeared rather spontaneously, without proper infrastructure. Scholarly work on these infrastructural problems exists, but in a fragmented form: on the impacts of suburbanization and sprawl (Kovács and Tosics 2014), on governance and metropolitan development, and on infrastructure stress in single suburban towns. Kocsis and Tomay 2024 for example frame sprawl-related infrastructural deficits and the traffic they generate as central concerns of Hungarian urban development, and their longitudinal case study of Törökbálint documents how post-transition agglomeration development translated into rising car dependence, with the railway infrastructure poorly matched to residential commuting needs - even in a town whose strong fiscal position allowed it to keep basic utilities well-supplied. Framing intra-metropolitan migration in such circumstances with classic push-pull models is hard, since pull factors have to logically precede new population. If infrastructure (or amenities) only follow population redistribution, then directions of migration flows have to be explained by other mechanisms. This is an important example where network methods can help model Hungarian and generally Eastern European internal migration behaviour that does not resemble residential migration processes seen in Western European and North American countries. The third article of this dissertation is centered around this idea.

This spontaneity is only a partially defining element of general Hungarian, and particularly Budapest-centred suburbanization and agglomeration processes. Under state socialism, location within the metropolitan periphery carried little economic weight, and the zone around Budapest remained relatively undifferentiated. The transition re-monetized locational advantage almost overnight: settlements with favorable positions pulled ahead, and the periphery became increasingly unequal (Szabó et al. 2014). At the same time, the spatial structure of the Budapest Metropolitan Region shifted away from strict monocentricity. Commuting, once overwhelmingly directed toward the core, became increasingly reciprocal as

employment concentrations formed outside the city; analyzing census commuting data from 1990, 2001, and 2011, Szabó et al. 2014 show that these new functions were redistributed among already existing settlement centers - a trajectory closer to the European pattern of post-suburbanization than to US-style dispersed sprawl. The Budapest Metropolitan Region has evolved from a monocentric to a more polycentric spatial structure since 1990, with increasing reciprocal commuting flows and the emergence of job centers beyond the core city (Szabó et al. 2014). Recent evidence shows that this polycentric transformation is still ongoing: new subcenters are emerging in the inner metropolitan ring, and their development potential depends on accessibility, and on historical and infrastructural foundations – another sign of path dependency in the system (Kocsis and Varga 2025). Network modeling is a framework that enables us to understand and characterize the nature of polycentric development happening in Hungarian urban areas properly.

One of the defining factors of internal migration is housing. The Hungarian housing system is special in Europe, because of its unusually large share of owner occupation. It has to be noted that this is not solely the product of the regime change: Hungarian housing was already predominantly privately owned under state socialism, outside the central urban areas and the large housing estates (Hegedüs and Tosics 1996), and high homeownership rates are common across Southern and Eastern Europe (Stephens, Lux and Sunega 2015). The ‘giveaway privatisation’ of the state housing stock after the regime change is one contributing factor, which pushed the ownership share to its current extreme levels (Kováts and Kohl 2025). Homeownership generally reduces migration propensities (Mulder 2006). Based on this, the generally low internal migration rates compared to western countries are understandable, but the effect of housing structure on migration flow directions in Eastern Europe is still understudied (Bloze 2009). Network methods can help us understand this relationship better. The third essay of this thesis is about this topic.

1.4. The network perspective: a possible synthesis

Through this introduction chapter, I tried to argue that the traces of network thinking can be found through the development of internal migration research, from the foundations to contemporary theories, and even in adjacent fields, like urbanization

research. However, a formally theorized unification of all these network threads is still missing. In an article based dissertation, such a unification is a too ambitious task to handle. But the three articles making up this thesis aim to be examples of research that is theoretically founded in interdependence based foundational works, while being empirically enriched by state of the art network science methods.

It is important to point out, that examples of interdependence-based thinking differ from each other in important details. Ravenstein with his laws for example does not provide general, abstract mechanisms when he introduces the effect that lead to network-like flow structures. He points out two main forces, absorption and dispersion that emerge from the growth of centers of commerce and industry, and which ripple through all settlements in a country. This physical way of thinking can only be generalized to countries with growing centers, and hints of more abstract dependencies are missing. The only exception is his description of mutuality in migration flows. The system he describes generally does not work on flow-dependence (when flows in the system define each other), flow structures and volumes are emerging from relative weights of settlements. His model is in this way a secondary network model, and primarily a location-relation model. When contemporary researchers apply models like ERGMs (like in the second article of this thesis), those explicitly model edge-dependence, and in this way, are in odds with the origins of the foundations of flow network theories.

As it was mentioned earlier in the introduction of migration system theory, and the work of Mabogunje, migration system theory is hard to treat as a standalone tradition, it was used either as a broad background for empirics, or got integrated into other theoretical frameworks. Network-approach migration research likes to cite MST as an origin for network thinking, however this connection is also not self-explanatory. In his seminal paper, Mabogunje builds his explanation around abstract system theory, with a special focus on feedback loops and information. In his explanation, a migration system is exactly a set of negative and positive feedbacks affecting each other. Feedback loops in his model are social institutions, like the family, or the urban job market, and not mere flow structures. Flow structures in his work also are a secondary pattern, as in the work of Ravenstein. Just in the case of Mabogunje, the relation is not between settlements, but between societal subsystems.

If we look at network-approach research, explicit mentions of feedback loops are missing, and meso-level, Mertonian mechanisms are not described and tested which can stand behind network structures and patterns that are observed in flow networks. The most important element of network explanations which is explicitly connected to MST is path-dependency. But path-dependency itself is not usually analysed; it is treated as a given (mostly statistical) nature of migration, rather than something that has to be explained in itself. Usually, network-approach research tries to explain migration network patterns that cannot be explained by path-dependency alone. We can easily see the contradiction with Mabogunje: he explicitly treated path-dependency as something that captures information flow - induced feedback mechanisms for migration. Hein de Haas, who is probably the most influential contemporary migration theorist working with a system heuristic summarises this theoretical problem as a neglect from researchers towards positive and negative feedback mechanisms that operate through migration and change the initial conditions under which migration takes place. He also points out that contemporary research cannot differentiate why some migration events lead to network migration and migration system formation, while some moves do not. Maybe his most important criticism is that not only positive feedback loops exist, but negative ones also - the mechanisms that undermine migration on the long run. As de Haas writes: migration studies has experienced "*theoretical regression rather than progress*" (de Haas 2010).

I hope that the three articles in this thesis can provide small, incremental steps towards theoretical progression, with different methodological innovations. The first applies multi-scale community detection, and a radiation-model predicted synthetic migration network as a comparison to show how internal migration flow structure is organized on different spatial scales, and which elements of the stream structure connects those scales. The second paper uses a count-ERGM framework to model the inter-district flow network of Hungary, to ask the question whether a house price boom in a dominant capital city can lead to immediate flow network reorganization across the migration system, or whether path dependency is stronger on the short timescale. The third paper shows that in the absence of significant amounts of new housing, social networks between individuals, and especially links to tightly knit communities can partially explain migration flow volumes in the agglomeration of

Budapest. All of the articles work with different scales and aggregations, ask different questions about the Hungarian migration system, and try to use different network science methods to understand internal migration better.

1st ARTICLE:
RURAL-URBAN FLOWS DETERMINE
INTERNAL MIGRATION STRUCTURE
ACROSS SCALES

Rural-urban flows determine internal migration structure across scales

Gergely Mónus, László Lőrincz

Published: Cities 2025

<https://doi.org/10.1016/j.cities.2025.105992>

Abstract:

Historically, internal migration happened through a centralization process upwards the urban hierarchy, but in recent decades, a shift towards network-based hierarchies has occurred in developed economies. This shift has led to the development of inter-urban mobility networks that reveal how different connections move populations between cities. Yet, most studies concentrate only on major cities and their networks, often overlooking the impact of smaller-scale connections. In our analysis, we show that inter-settlement migration patterns are heterogeneous across spatial scales, but a strong hierarchical structure persists across them. Results for a multi-scale community detection algorithm show that most of internal migration happens in local, clique-like groups of smaller settlements, and these cliques form larger units through connections to towns with center roles. Surprisingly, towns with similar center roles on different scales can differ by population and administrative rank. By applying the radiation model of migration as an analytic null model, we identify network patterns that cannot be explained by the population of settlements and the distance between them. We find that the radiation model fails to predict population flows from local cliques to centers, a characteristic type of observed flow that defines the hierarchic network structure.

1. Introduction

Internal migration is the long-term residential relocation of people from one defined locality to another one within a country. In European countries with low rates of natural population increase, internal migration has already replaced fertility and mortality as the primary demographic process shaping the spatial distribution of populations (Rowe and Patias 2020). Traditionally, internal migration was modelled by considering localities as independent observations. The most widespread theoretical framework, push-pull theory presumes that migration behaviour in a settlement only depends on the attractive and repulsive qualities of that given settlement (Poot et al. 2016). Recent research questions the notion that migration research should use the individual locality as the basic unit of analysis. In the age of information society, globalization and the availability of functional telecommunication techniques led to a new geography of centers and peripheries. The space of places and space of flows interact with each other, and this interaction defines a new hierarchy (Castells 2020).

The notion of modeling internal migration as a system of interdependent flows already appeared at the very beginning of modern migration research, but it did not get particular attention during the last century. The first modern social scientific work on migration, *Laws of Migration* by Ernst Ravenstein modelled the internal migration system of the UK. Ravenstein describes that most people move to the centers of commerce and industry, and the gaps they leave are filled by migrants from more distant areas (Ravenstein 1885), 2nd law. This absorption process influences the whole population, by producing currents of migration. By applying current as the basic theoretical building block, Ravenstein's work aims to explain migration as a process induced by relations of settlements. Consequently, Ravenstein's laws can easily be translated to a network model: localities are nodes, where agents reside, and the movement of agents produces the emergent macro-level structure of edges.

Ravenstein described the migration system of the UK during the time of the Second Industrial Revolution, an era of high levels of centralization and urbanization. On the level of individual behavior, this meant that the average migrant usually moved one step upwards the urban hierarchy (Plane et al. 2005). This step-migration model

serves as an analytic null model for sociologists and demographers and can be extended to general urban hierarchy models. Urban hierarchy models are useful to identify turning points in internal migration (Mu et al. 2022), to understand how missing flows between levels of a hierarchy can explain developmental differences between regions (Lerch 2016), or connect urban development research to different migration theories, such as the life-course models (Plane et al. 2005). Traditionally, urban hierarchy models focused only on the size or administrative rank of different cities, but recent trends in urban development have led to a shift towards network-based hierarchies (Neal 2011).

This shift towards network-based urban and internal migration systems happened because continued urbanization led to the birth of a new urban form: polycentric urban regions (Meijers 2007). Through differentiation processes, urban functions evolve through center-subcenter-periphery structures (Vasanen 2012). Systems of central places change to network systems through the change from centrality to nodality, from vertical accessibility to horizontal accessibility, and from the tendency towards primacy to the tendency towards complementarity (Batten 1995). The problem with polycentric development theory is that it has been adopted so widely, that it is used in very different contexts, and very different spatial scales (Liu et al. 2016, van Meeteren et al. 2016).

Meanwhile, researchers have mapped different interurban networks, such as transportation (Dong et al. 2016), telecommunication (Krings et al. 2009), corporate (Zhao et al. 2017) systems as networks, as well as online social connections (Lengyel et al. 2015). It is also shown that online information flow between cities and offline population flow between them is highly correlated (Hu et al. 2020, Lőrincz and Németh 2022). Interurban network models have also proven useful to map population flow characteristics, for example, delineating source and sink regions of migration (Wang et al. 2019), identifying temporally coherent subgroups of settlements in internal migration systems (Zhang et al. 2020), or the rich-club characteristic of internal migration (Zhu et al. 2021). The network perspective also enabled mapping the structure of urban regions and their polycentric development (De Goei et al. 2010, Liu et al. 2016), to differentiate central cities and hinterlands of megacity regions (Hu et al. 2024).

Research on inter-urban networks has also shown, how network effects relate to size-based scaling effects. The superlinear scaling in wages between cities can be explained more through network effects (Keuschnigg 2019), and mobility connections have a greater impact on cities' exports than population-related factors such as GDP or consumption (Lei et al. 2021). Analysis of intermediacy in intra-urban road networks also show that both small cities with high number of connections and large cities can achieve similar network characteristics (Prieto-Curiel et al. 2022). Untangling the effects of population and connectivity is a fresh direction in urban research, thus the number of observed geographies and network types is still limited (Liang et al. 2024). While there are modeling frameworks to understand how spatial interactions of the population lead to city hierarchies (Batty 2013), current literature mostly focuses on dissecting size effects and network effects.

Previous research, however, mostly focused on networks of larger cities. For example in the articles identified by the systematic meta-analysis of Li and Neal 2024 the sample size in terms of cities varied between 9 and 1575, while only our middle-sized case country, Hungary contains more than 3000 settlements. The approach of focusing only on larger cities overlooks the role of the connections between villages, small towns, and regional centers, which is problematic for multiple reasons. First, delineating rural and urban areas is not straightforward, and different delineation methods often obscure certain social phenomena because their criteria may label areas as rural, even if these areas share characteristics or exhibit dynamics typical of urban environments (Tacoli 1998). Second, rural-rural, urban-urban, rural-urban, and urban-rural migration are usually studied separately, while they are interdependent, and their relative influence in the same geography can vary over time (Selod and Shilpi 2021). The network approach can systematically integrate these different directions and show their interplay. Third, while inequalities between urban and rural regions are growing, previous research has shown that small towns have a vital role in the reduction of rural dependency on city centers, enhancing the development of rural areas (Najarsadeghi and Dorostkar 2022). Network modeling of flows on the micro- and mesoscale of urban hierarchies can shed light on how smaller towns broker relations between small villages and main cities.

Distinguishing network effects is particularly important for migration and mobility research in under-studied geographies. For example, internal migration in the wider context of our case study, Central- and Eastern Europe, is usually only understood through the notion of centralization. Historically, countries in this area experienced strong urbanization trends with larger population gains in main urban areas (Rowe et al. 2019). However, the accompanying suburbanization processes were different from the ones observed in Western countries. Suburbanization in Western societies happened through strong population outflow from cities, while in less industrial Eastern European countries migration flows from rural origins stopped before the city border, growing satellite settlements in the urban periphery (Stanilov and Sýkora 2014). After the regime change wave - the most influential event in the development of the region - the area has been connected to the globalized economy at a rapid pace. Therefore, economic and social centrality became less tied to geography, but rather determined by inclusion in global networks (Castells 2020). The new network hierarchies led to a shift towards polycentric metropolitan regions, with functions being reshuffled between existing centers (Szabó et al. 2014). Counter-urbanization and subsequent rural gentrification trends also started to emerge, especially aiming at rural communities with significant cultural heritage (Dragan et al. 2024). It has also been shown that while migration happens by economic motives, regional identity and stereotyping play an important role in the exclusion versus integration of migrants (O'Brien et al. 2023).

Looking at the previous literature, we found only two studies using a systematic network approach to understand intra-settlement migration focusing on countries in the region, Pitoski et al. 2021b for Croatia, and Pitoski et al. 2021a. While these two examples offer valuable network descriptions of migration systems of the region, they do not identify system-defining elements and mechanisms. Therefore, a key question remains: how do migration flows structure connections across different scales of the urban hierarchy? While past research has documented broad urbanization patterns and shifting migration mechanisms in the region, there is still little understanding of how individual settlements – ranging from small villages to major cities – integrate into the internal migration network. Our main research question therefore is:

Research Question 1

What elements of the flow structure connect the different scales of an internal migration network?

To answer this question, through a case study of the internal migration system in Hungary, we investigate how settlements with different sizes and administrative ranks connect to other settlements in a residential migration network. We build the network we analyse from all the municipalities of the country. To be able to map migration on all resolution levels, we define migrants in the highest granularity available, as any person who permanently changes residence across municipality borders.

In the analysis, it is important to differentiate between connections pointing up and down the urban hierarchy. This way the inter-dependent nature of different migration directions between urban and rural areas can be treated systematically. Through multi-scale community detection with the map equation framework, we are able to show different spatial scales of the internal migration system of Hungary, from smaller, local, dense groups of settlements, to polycentric urban regions containing more than one large town or city. We show that the attracting and brokering influence of population size is different between flows connecting different levels of the urban hierarchy.

To be able to understand this varying effect of different settlement sizes, we have to distinguish between effects that are simple emergent mechanisms of the geographical layout of population and the effects that are the consequences of more complex societal factors. Our second research question, therefore:

Research Question 2

Is population size and geographic distance sufficient to model connections between different scales in the network?

To address this question, we use the radiation model (Simini et al. 2012) as a null model for comparison to the observed network. By comparing observed migration behavior to the radiation model predicted migration flows, we can differentiate which network characteristics emerge from non-demographic aspects of the city system inside Hungary.

Our findings indicate that in the observed migration network, module formation is driven through urban agglomeration formation, while in the radiation network migration modules do not correspond to existing urban areas. Moreover, our results show that while the radiation model is relatively accurate in modeling outflow from urban centers, it fails to predict rural-urban flows, a characteristic type of observed flow that defines the hierarchic network structure.

2. Data and methods

2.1. Data

For our analysis, we use permanent residential migration records collected by the Hungarian Central Statistical Office. These records show every officially administered migration event in the country, with origin and destination locations recorded on the municipality level (which equals to settlements outside the capital, and the districts of Budapest). The data, in aggregated form, is accessible through different HCSO publications (like the Statistical Yearbook of Hungary), and in granular form upon online request from HCSO, and in the data rooms managed by HCSO.

We constructed a settlement-level network of migration events between all settlements in 2015–2019. For constructing the network, we use only changes in the permanent address and exclude temporary addresses, because, in Hungary, temporary addresses are mostly used by young people during short-term mobility (like university students, young tenants in larger towns, or guest workers). On the other hand, permanent address change usually refers to a life cycle event (marriage, establishing a new household, having children, etc.). While some studies apply thresholds for the minimum population size of settlements, we include every Hungarian municipality. As we will show, our results indicate that there are settlements that have a higher influence on the migration network than their size or administrative status would predict intuitively. Also, by including even the smallest settlements, we became able to track the influence of large towns on their surroundings more precisely. Because spatial granularity requires longer timescales,

we aggregate the data over five years to exclude short-term local shocks and fluctuations.

The constructed settlement-level migration network is both weighted and directed. The weight of every connection is simply the total migration between the two settlements in the examined five-year period.

2.2. Methods

To uncover the hidden structure of the migration system, first, we describe the network from the aspect of urban hierarchy. Specifically, we analyse the concentration of migration flows, by considering the share of the largest inflow and outflow by size and settlement type.

Next, we apply multi-scale community detection. Using community detection we identify the organizing mechanisms of the migration system, taking into account our key interests, distance, and urban hierarchy. With multi-scale community detection, we can observe whether the organization of communities is different across scales.

Network characteristics, however, could simply be a consequence of the population distribution of the country, and the geographical layout of settlements. To disentangle the effect of these factors, we apply a radiation model as an analytic null model for a comparison network. Through the steps of our analysis, we compare the observed and radiation-simulated migration networks.

2.2.1. *The radiation model*

The radiation model has been widely used as a null model to analyse migration and other interactions between cities. The radiation model builds on the classical gravity model, in which the magnitude of migrants between two cities depends positively on the population of cities, while the distance between them affects it negatively according to a decay factor (Poot et al. 2016). The gravity model however has a strong limitation, namely, that it can only predict undirected flows. This limitation prevents it from being useful for more granular analysis, where flows from large cities can be very different than flows to these cities. The radiation model overcomes this limitation by introducing the effect of neighbouring settlements on every city's

migration behaviour. According to the model, the volume of a flow between locality i and locality j is:

$$T_{ij} = T_i \cdot \frac{P_i P_j}{(P_i + s_{ij})(P_i + P_j + s_{ij})}$$

where P_i and P_j are the populations in the localities, and s_{ij} is the population within a given distance to locality i. This way, flows became directed and asymmetric, based on the geometry of population distribution and distances. The resulting model predicts mobility patterns in good agreement with mobility and transport patterns observed in a wide range of phenomena, from long-term migration patterns to communication volume between different regions (Simini et al. 2012). Radiation models have been used to analyse many different problems of human mobility, such as commuting patterns (McNeill et al. 2017) or the spread of infectious disease (Luca et al. 2022). The application of such models to migration systems has also been proven successful. Li et al. 2017 modelled migration between 337 cities in China, while Zheng et al. 2020 applied such models to analyse the Yangtze River Middle Reaches urban agglomeration. For European countries, such studies are more scarce, but Pitoski et al. 2021b used the radiation model to analyse internal migration inside Croatia, a neighbouring state of Hungary.

While there are many extensions and generalizations of the original radiation model, we choose to use the original one as a null model for our analysis based on two main reasons. One is that it is not data-intensive and only needs information about the locations and populations of settlements, leading to results based purely on the effects of demographics and spatial layout of cities. The other reason follows the first one, which is that the model's simplicity makes our research more comparable to studies conducted in other countries.

It must be noted that the radiation model was also criticized; foremost on the basis that because population size is used as a proxy of the economic activity of cities, it falls short of predicting mobility in societies where migration events happen for non-economic reasons, for example enhancing life quality, which is not necessarily related to larger populations (Alis et al. 2021). This limitation, however, does not limit its utility to be used as an analytical null-model, for which we use it. Another criticism is that it predicts long-range moves worse than short-range ones (Kluge et

al., 2022). This can affect its applicability as a null model, but as our results show, did not result in significant problems in the case of this analysis.

2.2.2. *The map equation and Infomap with Markov time scaling*

To be able to find characteristic clusters in the migration system, we have to be able to recognize groups of settlements that are strongly connected through migration flows. Community detection algorithms are the most important tools to analyse strongly connected sub-units in networks, called modules or communities. Previous studies shown the utility of this approach both for national (Pitoski et al. 2024) and international migration systems (Deutschmann 2022). Available community detection methods are numerous, with distinct advantages and disadvantages for all of them. In our analysis, we are focusing on the migration process happening in Hungary, and for this, we look for a community detection algorithm that aims to uncover network structure based on dynamics happening in the network, not its topology.

Out of the different options, we apply the map equation, and its search algorithm Infomap for two reasons. First, it is a flow-based community detection method, aiming to find modules exactly based on network flows (Rosvall et al. 2009) that corresponds directly to our research subject of migration between localities. Second, the map equation, in contrast to several other methods, treats directed and weighted networks naturally, which also corresponds to the migration network we consider. For the map equation, the proxy for network flows is a random walker, and it utilizes the duality between network partitioning and the description length of random walks on the network. The partition with the shortest description length can be seen as the optimal partition. A description is compressed if the network has clusters where a random walker tends to stuck for a long time. For module partition M , the average description length of a single step is given by:

$$L(M) = q_{\sim} H(Q) + \sum_{i=1}^m p_{\cup}^i H(P^i)$$

The equation contains two terms, the entropy of the movement between modules, and the entropy of movements within modules. $q_{\sim}$ is the probability that the random walk switches modules on any given step. $H(q)$ is the entropy of the module names, $H(P^i)$ is the entropy of the within-module movements. The weight $p_{\cup}$ is the fraction

of within-module movements that occur in module i , plus the probability of exiting module (Rosvall and Bergstrom 2008).

The map equation and Infomap have shortcomings, the most important for networks like the migration network of Hungary is the field-of-view limit. The field-of-view limit is an upper scale on the communities the algorithm can detect, which means long-range communities tend to be overpartitioned (Schaub et al. 2012). This limitation would be a strong constraint on our analysis because it is not realistic to assume that settlements are only connected in local clique-like (all-to-all) groups, longer-range connections to distant settlements also play an important part in population redistribution. Introducing Markov time scaling to the map equation overcomes this limitation (Kheirkhahzadeh et al. 2016). The original map equation works with a Markov time one. Markov times lower than that mean that the same node may be visited by the walker multiple times, because it moves slower, resulting in smaller modules. Markov times over one on the other hand mean that not every node on the trajectory will be encoded, because the walker moves faster, resulting in larger modules (Smiljanić et al. 2026). Markov times scaling offers a dynamic zooming tool across scales, which enables us to analyse module formation in the migration network for every kind of spatial scale, and every kind of migration behaviour.

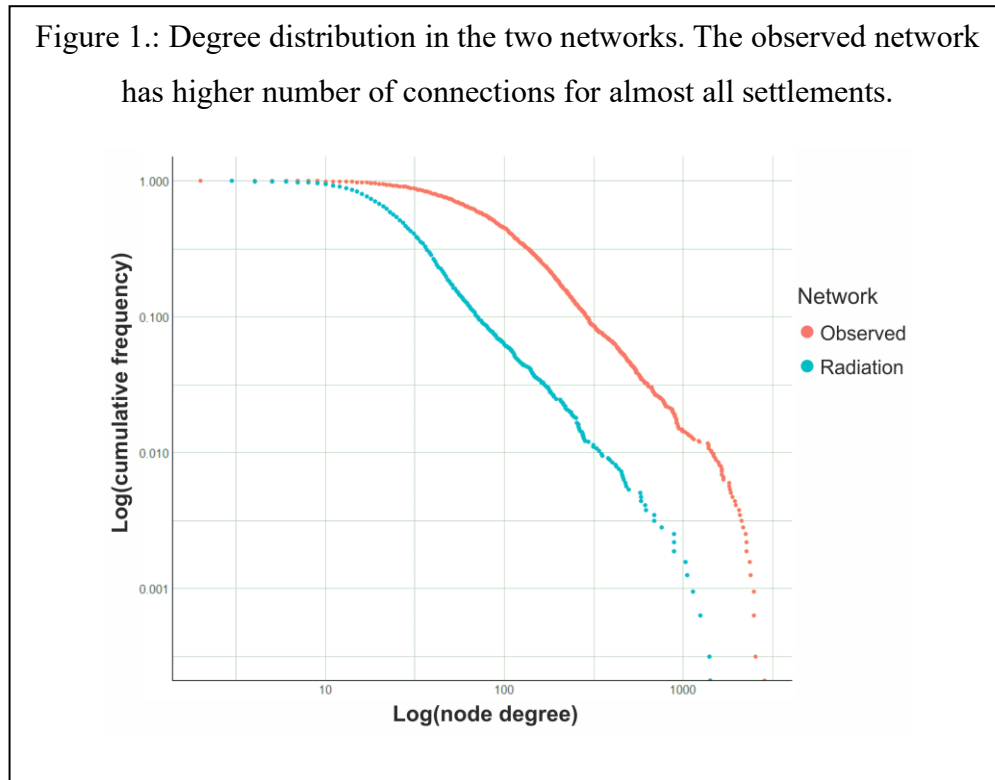
3. Results

3.1. Comparison of the two networks

3.1.1. Simple topological properties of the two networks

Both the observed migration network and the predicted radiation network are built from all 3177 Hungarian municipalities (23 districts of the capital, regional centers, other towns, and villages). Figure 1. presents the degree distribution of the two networks. It is apparent that while the shape of the distributions is similar, the observed network has a much higher number of flows compared to the radiation network. We get similar conclusions from Table 1, where we present further network statistics for these networks. In the observed network, there are 244,314 edges, resulting in a mean total degree of 153.8. Settlements tend to have similar amounts of

incoming and outgoing migration streams, the median in-degree is 43, and the median out-degree is 46. The radiation model however underestimates the number of migration streams settlements have, predicting a mean degree of 44.13, less than a third of the observed amount. Furthermore, in the radiation network, incoming and outgoing connections are asymmetric, the median in-degree is 70% higher than the mean out-degree.

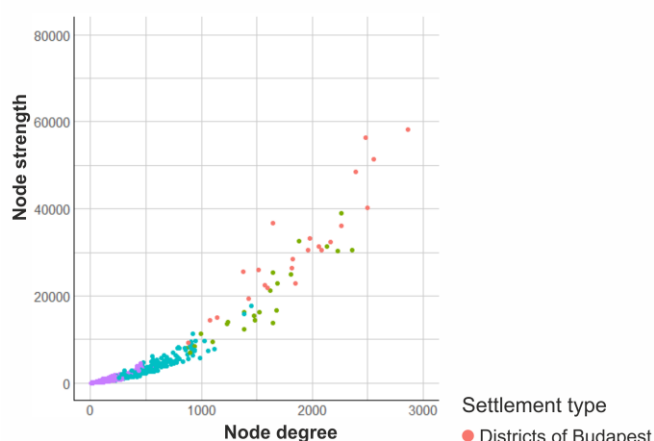
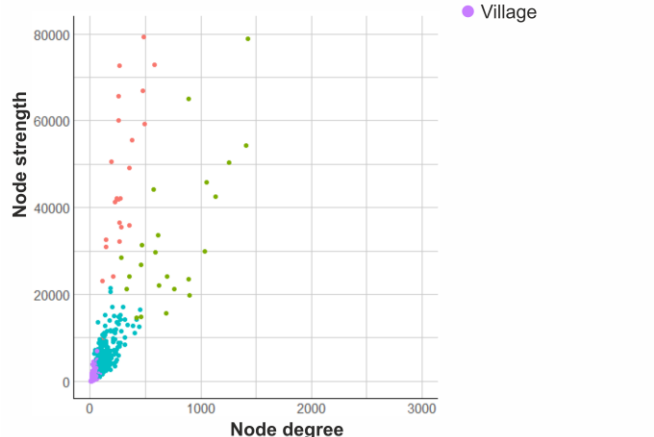


The average edge represents 6.09 migrants, meaning most connections do not move large volume of migrants. The mean degree strength is 936.5. As with the case of degree, node strength also has a very skewed distribution. Strength coming from incoming and outgoing connections is symmetric. Corresponding to the lower degree, the radiation model predicts significantly higher edge weights and node strength. The mean edge weight according to the model is 38.5, the mean node strength is 1699.8.

Table 1. Descriptive statistics for the two networks		
Network statistic	Observed network	Radiation network
Settlements	3177	3177
Edges	244314	70093
Mean edge weight	6.1	38.5
Mean total degree	153.8	44.1
Median total degree	89	27
Median in-degree	43	17
Median out-degree	46	10
Global cycle CC	0.47	0.57
Global middleman CC	0.48	0.82
Global in CC	0.47	0.75
Global out CC	0.49	0.69
Global total CC	0.47	0.71

Compared to the observed network, the number of connections is lower in the radiation world, but these flows move a larger amount of migrants on average. Figure 2 shows the association between node degree and node strength by administrative rank of settlements. For the observed network, a super-linear scaling is apparent, which is not the case for the radiation network. In Hungary, the only region with high population growth during the last decades was the central region, the capital, and its functional urban area. The central municipalities of Budapest thus have the highest degree (reaching over 1000 connections) in the observed network, and only the largest towns can compare to them. This feature, however, is not reflected in the network predicted by the radiation model. While the districts of the capital are comparable to regional centers in node strength in the predicted model, they have fewer connections. This difference between the observed and predicted networks indicates that districts of the capital attract migrants from a larger number of settlements than we would expect based only on population size; possibly for social and economic reasons. The radiation network predicts a comparable population redistribution through flows, but with less numerous, diverse, and distant connections.

Figure. 2. Node degree and node strength by settlement type.

A. Observed network**B. Radiation network**

A: In the case of the observed network, strength scales superlinearly with number of connections. B: The radiation network is unable to predict the superlinear relationship, and heavily differentiates between the districts of Budapest and other towns.

understand group formation with simpler tools than community detection, we analysed the clustering behaviour for both networks (also shown in Table 1). Traditional clustering coefficients are not suited for weighted and directed networks, so we used five different local clustering coefficients developed by Clemente and Grassi 2018: the cycle, the middleman, in-, out- and total clustering coefficients. These metrics all account for different types of triangles in which a settlement can take part. For the observed network, all coefficients are of similar size, around 0.5. The radiation network is clustered more strongly according to every metric, but interestingly, there is a high heterogeneity between the coefficients. The cycle

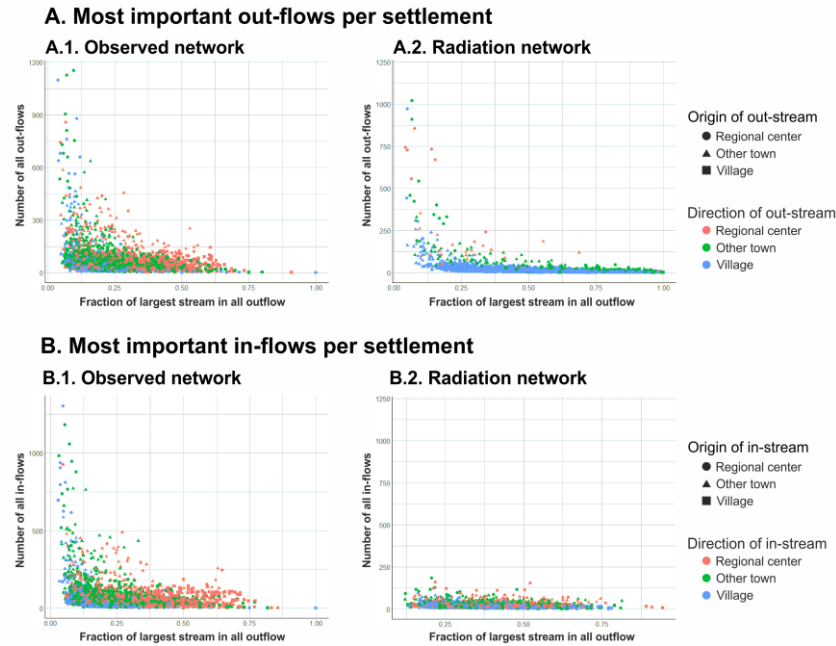
clustering coefficient is the lowest, falling close to the coefficient of the observed network. The cycle clustering coefficient accounts for triangles of settlements (i.e. settlements A, B, C), where there is a flow pointing from A to B, from B to C, and from C to A, or vice versa. The coefficient with the highest value is the middleman clustering coefficient, which is almost twice as in the case of the observed network. This metric accounts for triangles where there is a flow pointing from A to B, and from B to C, but the last flow is not finishing the cycle, but pointing from A to C (making B the middleman). This shows that the lower number of stronger edges not only leads to higher clustering, but the patterns of connections are very different compared to the observed network.

3.1.2. Analysis of the most important migration connection of settlements

In order to better understand the microstructure of the network, in the next step, we looked at the most important inflow and outflow of every settlement, whether they account for a significant share of inflows and outflows. This is relevant for the community detection, as for weighted networks, the map equation algorithm treats edge weights as probabilities for the random walker to choose a direction, thus strong connections will influence the multi-scale structure of network communities. Here we also compare our results to the null model based on the radiation model that takes into account these effects.

Figure 3 presents the relationship between the largest migration streams and the total number of streams a settlement has (for in- and out-flows, and for different administrative ranks). For the observed network, both the largest inflows and outflows follow a skewed distribution, with a mean of around 20%. For a large share of settlements, the strongest migration connections represent over 30% of both in and outflow. This result indicates that for most settlements, residential migration streams are not uniformly important, and in many cases, there is a dominant partner.

Figure 3. Migration flows with the largest volume for every settlement for the observed and the radiation network.



While the radiation model is able to predict the two-dimensional distribution, in the observed network rural-to-urban outflows mix population across the urban hierarchy. A: Most important out-flows per settlement, by origin and direction. B: Most important in-flows per settlement, by direction and origin. The radiation model is unable to predict the magnitude of in-flows of regional centers and other towns.

In the radiation-model predicted network the volume of the most important outflows is highly overestimated. Instead of a heavily skewed distribution, the model predicts a more concentrated one, with a mean near 50%. According to the radiation model, more than half of the population inflow is coming from a single settlement in around half of the settlements. This error of the model is the strongest in the case of non-regional center towns, which shows that the by simple radiation laws, agglomeration formation around smaller towns is less predictable based only on size and distance. In the case of in-flows, the observed and radiation-predicted share of the strongest flows show a similar distribution, with a similar mean.

Further, as indicated by the distribution of different shapes in Figure 3. B1, the number of inflows is the highest in the case of regional centers, it spreads between 100 and 700 in the case of other towns and remains under 200 in the case of villages.

In the case of the radiation model however, the number of inflows is independent from the settlement category (Figure 3 B2). Considering the settlement types of the origins and destinations (colours and shapes together), we can see a strong mixing effect across the urban hierarchy. By mixing effect, we mean that outflows show diversity in the status of origin-destination pairs. Typically, for strong in-flows, with a share higher than 50 %, this mixing has a hierarchical nature, the origin is typically a village, and the destinations are regional centers or other towns. This pattern is present both in the observed and the predicted network, thus the radiation model is able to predict directions across the urban hierarchy.

In the case of out-flows, the radiation model fails to predict the structure of observed streams. In the case of regional centers, the number of out-streams is similar in the predicted network. In the case of smaller settlements, predicted total stream numbers are not as high as observed. Most importantly, observed population out-flows also have a mixing behaviour: streams typically originate from villages, and go to regional centers or other towns, but with significant variability. According to the radiation model, this is not the case: settlements typically send population out-flow to the same kind of settlement. There is some mixing between regional centers and other towns, but most villages have their strongest connection with another village. Thus, based purely on size, villages would not send their population to smaller towns nearby, and more migration would happen locally between smaller settlements in dyadic-triadic structures. This divergence can lead to different community structures according to the map equation, if the observed, but unpredictable mixing across the urban hierarchy (the population brokerage role of smaller towns) is driving community formation in the observed migration networks.

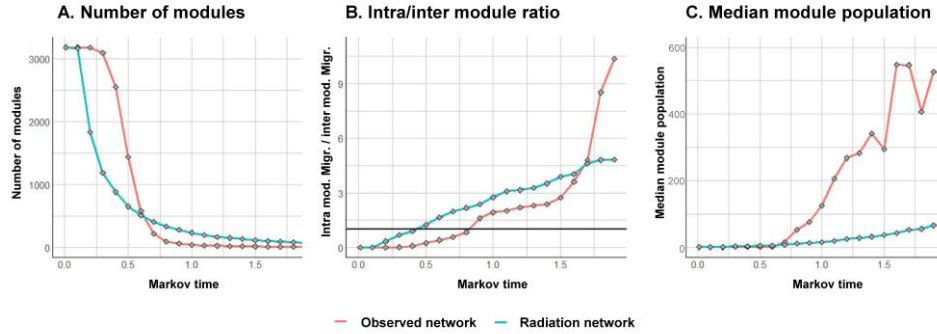
3.2. Multiscale community structure

3.2.1. *Results of the map equation and Infomap for the observed network*

Figure 4 indicates how community structure evolves at different Markov times in terms of the number of modules, the ratio of intra-module and inter-module moves, and the median population of modules. In the number of modules (Figure 4 A) a reversed S pattern is observable. The algorithm starts with every settlement being its community. This initial state starts to change around Markov time 0.2, after which a

very steep fall happens until Markov time 0.8. After that the decrease is also

Figure 4. Results for multi-scale community detection with the map equation and Infomap.



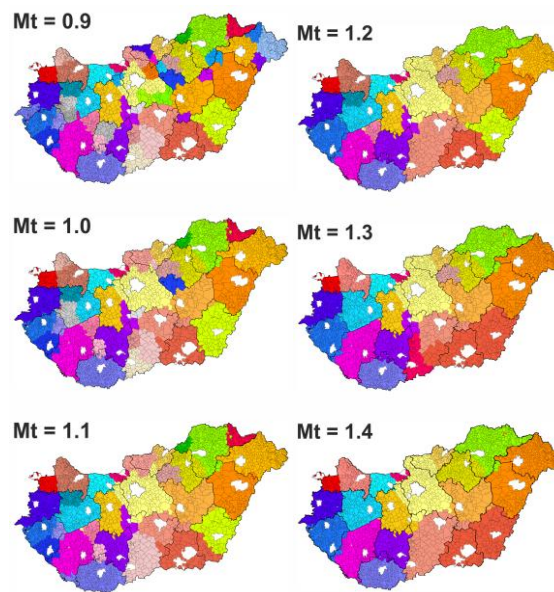
A: Number of modules at different resolution parameters. B: The ratio of intra-module migration and inter-module migration at different resolution parameters. If the ratio is 1, the same number of moves happen inside modules and between them. C: Median population of modules at different resolution parameters. In the observed network, compared to the radiation model, settlements start to agglomerate to modules slower, but they converge to large agglomeration sizes quicker at higher scales.

monotone until there is one module, but in a much slower pace. Between scaling parameters 0.2–0.8 the algorithm starts to connect nodes to their neighbours with the strongest connections. However, as the module population graph (Figure 4 C) shows, most of the modules remain as single settlements for a long time. This is because the algorithm recognizes the strong dyads first (that we analysed in the previous section), which typically occur in population- and migration-wise dense areas, mostly the agglomerations of cities. It considers the partition optimized by connecting only these settlements and leaving alone low migration, smaller settlements, which are further away from urban areas.

We can recognize three very different phases in Figure 4 B, considering the intra/inter-module migration ratio, with 0.8 and 1.5 Markov times as boundaries. The ratio reaches one around Markov time 0.8, indicating that in this first phase, more movements happen between modules than within them. This is an interesting finding because modules are still very small at this point, the median cluster size is 24

settlements and around 50,000 population. At this point, almost every settlement is a part of a community of such size, only very low migration settlements stay in their own module. What happens after this, in the second phase, we also illustrate in maps of Figure 5. At this point, the dense, clique-like parts of agglomeration zones are classified as one module. Based on population and migration density, these agglomerations can be quite large geographically (like in the eastern part of the country), but in the case of Budapest and western cities, they remain relatively small. In the maps, different colours represent different modules (regional centers are marked with white, the largest white town in the center, marked with a black circle is Budapest).

Figure 5. Multi-scale community structure in the observed network.
Examples of spatial layout of modules at a few different resolution parameters

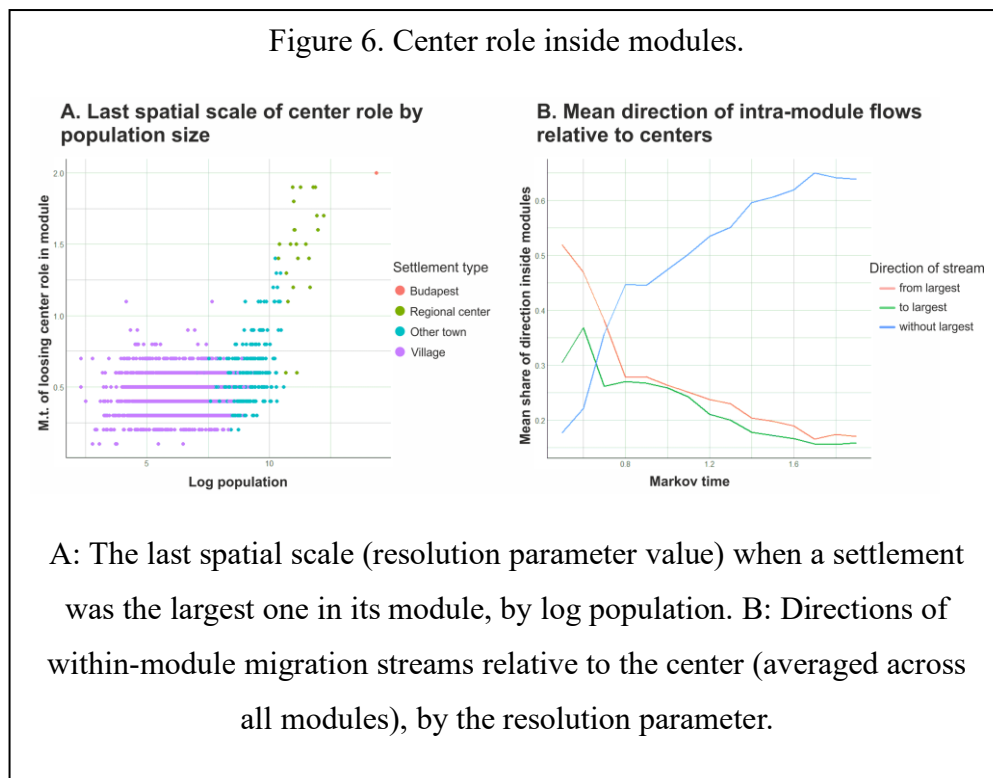


Different colors represent different modules. Thin lines represent borders of settlements, thick lines represent borders of counties (NUTS-3). Regional centers and Budapest are marked with white (and they always belong to the module around them).

After Markov time 1, a plateau-like slow, almost linear growth trend can be observed in the intra-inter migration ratio (Figure 4 B). Module populations are growing until Markov time 1.2, after which they also plateau until the next phase. From 0.9 Markov time, the process that we observe in Figure 5 is simple: agglomeration zones

expand. The modules of larger towns swallow the modules of smaller, less dense, less populous towns. Only a few modules without regional centers survive this phase until relatively high Markov times. However, the fact that a few such modules exist at high parameter values shows that certain towns can get the role of a regional sub-center without being populous or having the administrative rank of a regional center. The settlement structure also has a characteristic mark on the multi-scale structure: in the eastern part of the country, modules with two or more regional sub-centers occur more frequently, while in the west it is rare for relatively high Markov times. This shows that settlement structure strongly affects migration density, and thus the relative importance of given towns in the internal migration system. The settlement structure also interacts with the administrative system of the country: in the eastern part of the country, modules quickly converge to counties (NUTS 3 level units), and at high scales, they correspond to regions (NUTS 2); while in the western half, this convergence is much slower, and stops at the county level.

Figure 6. Center role inside modules.



The process of agglomeration zones absorbing each other changes after Markov time 1.5, from when the intra-inter ratio starts to skyrocket: from this point, the sizes of modules stay the same, and the large central cluster starts to absorb every other agglomeration module. At Markov time 2, only the module of Budapest remains. The results of the community detection algorithm itself can be thought of as a way to

operationalize the urban hierarchy. As Figure 5 shows, the community of the capital starts to absorb the modules of smaller regional centers as we increase our resolution parameter, which shows that it is the city attracting the most long-distance moves in Hungary, but it is unclear whether this phenomenon coming from the fact that it is the most populous city, or because of its administrative role.

The process of modules of larger towns absorbing the modules of smaller towns raises the question of hierarchical community structure across scales. By hierarchical structure, we mean the fact that if a settlement belongs to a given module (defined by the other members of the community), it will belong to the same module at every higher resolution parameter. The maps in Figure 5 show that the migration network of Hungary is indeed strongly hierarchical across scales. First, smaller settlements do not change modules, they stay in the same clique they belong around Markov time 0.9–1. This shows that settlements give and receive population in a well-defined local neighbourhood, and even if we move the random walker faster, these neighbourhoods stay well defined. Second, these cliques typically stay attached to the same large center across scales. This is however not perfectly true. For example in the central-southern part of the country - the agglomeration zone of regional center Baja-, modules on different scales have different shapes (between Markov times 1.2–1.4). It is interesting to point out that Baja has only been promoted to the administrative rank of regional center recently, so the urbanization process is not at the same stage there as in case of other, more important centers. But apart from these few local examples, there is a strong hierarchy across scales in cliques connecting to larger level units through the local center town. This hierarchy is not imposed by the algorithm, which is a valuable property of Markov time scaling with the map equation. Local migration cliques are connected to one center, and the star shapes emerging from these connections define a multi-scale community structure.

In order to understand the association between population size, administrative rank, and the influence of different settlements, Figure 6 shows the relationship between the log population of a settlement and the last Markov time value where it was the largest member of its community. The plot shows two different regimes, one of villages, and one of towns and cities. In the case of villages, there is no clear trend. There are villages with relatively large populations, that lose their central role in their

communities at low resolution parameters, while there are villages with low populations and their own community at high values of the zooming parameter.

The case of towns and regional centers is more clear, a linear trend between log population and Markov time of incorporation can be observed. Regional centers, except for two outliers, which both fall very close geographically to a much larger city, outperform other towns, signifying that administrative ranks indeed affect (or reflect) influence in the internal migration system. This underlines the path-dependent nature of migration and urban development: these cities are historical centers of their regions.

The fact that one settlement is the most populous in its community does not necessarily mean that has a central role. Figure 6 shows the mean value of different aggregated flows at every Markov time parameter: the flow from the largest, and to the largest settlement, and the sum of flows not connecting to the largest member of a community. We can observe that the aggregated flow from the centers is higher than the flow to these settlements at every parameter value.

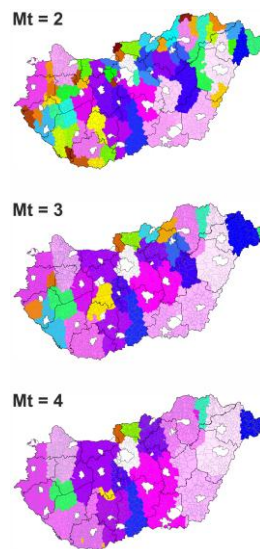
Flows connecting to the largest settlement are larger on average under Markov time 1.1 than flows between other settlements. At large parameter values, this latter flow type represents over 60% of the total migration within the modules. This indicates that as long as every community has only one large city, the flows connecting to that city represent the majority of movement inside communities, thus, the most populous settlements indeed have a central role in the distribution of migrants in their area of influence. At higher resolution values, however, we can observe communities based on the connections between these small communities.

3.2.2. Results of the map equation and Infomap for the radiation network

The multi-scale community detection on the migration network predicted by the radiation model brings surprising results. As apparent in Figure 4, scaling is a much slower process on this network, with smaller, clique-like communities remaining for high parameter values. Figure 4 A illustrates that although the number of modules starts to fall faster than in the case of the observed network, this agglomeration slows down, generating a higher number of modules throughout the range of Markov times higher than 0.6. At parameter values where we have seen the formation of urban

agglomeration zones in the observed network (0.9–1.1), the radiation network does not show signs of that. As we can see from the descriptive results, the radiation model predicts stronger flows, but pointing in less versatile directions, which explains why smaller modules get connected into larger structures much slower. The number of modules only reaches one at Markov time 5.8, out of the range of Figure 4, almost three times the resolution value than in the case of the observed network. Such a high parameter value shows that the map equation can not coerce the radiation network into one large module in any meaningful sense. At very high Markov times, volatility appears in community statistics, both in term of number of modules, population of modules, and inter-intra module migration ratio.

Figure 7. Multi-scale community structure in the radiation network.



Examples of spatial layout of modules at a few different resolution parameters. Different colors represent different modules. Thin lines represent borders of settlements, thick lines represent borders of counties (NUTS-3). Regional centers and Budapest are marked with white (and they always belong to the module around them).

The change of intra-inter module migration ratio (Figure 4 B) reaches one much faster than in the case of the observed network (at Markov time 0.5). Therefore, the radiation model predicts the creation of cohesive groups on a much smaller scale, than the scale of regional centers, where it happens in reality. Furthermore, the sudden changes that we observed are also missing here. In the observed network, the different phases of cliques sending population to large centers in the first phase, and

large centers connecting many cliques into zones of urban influence afterward can be distinguished, which is missing when we consider the simulated network.

The maps on Figure 7 show that multi-scale structure according to the radiation model does not only fail to predict urban agglomeration zones, but also misses other structures. As we raise the resolution parameters from Markov time 2 (which roughly corresponds to module sizes of Markov time 1 on the real networks), it is visible that the module structure does not converge to county borders. This shows that the high influence of the administrative structure does not come purely from demographic factors and the spatial layout of the settlement structure, but the path-dependency of social processes, upon which the historic administrative system is built. The structure across scales in the radiation network also shows a weaker hierarchy compared to the observed one. While local cliques tend to stay together across scales in this case also, cliques can belong to different higher-order structures on different scales. This underlines that while local migration mechanisms are predictable with the model, the connections between local groups and regional centers are not.

4. Discussion

Our results relate to multiple important topics in previous research on urban population processes and internal migration. The finding that the observed multi-scale community structure is strongly hierarchic, without the algorithm imposing it, points to the fact that migration decisions on all scales are pre-determined to a large extent. There are not only local clusters defined by short-distance moves, but most small settlements in these local clusters send and receive population with the same local center in medium-distanced moves, and the same large city in long-distance moves. The comparison with the radiation model makes this finding more interesting: in the radiation model, a lower number of streams move higher populations, so theoretically, network communities should be more strongly knit together. However, we find that up from the scales defined by medium- and long-distance moves, the multi-scale structure is less hierarchic in the radiation network. This shows that path dependency induced by social factors is a stronger influence on characteristic directions than forces of large population and geographical distance. This connects to qualitative research that shows that local identities and local

heritage define mobility decisions, making mesoscale units preserve themselves through time (Dragan et al. 2024).

Further, this path dependency, we observe, is not fully explained by the administrative system of the country. As the maps in Figure 5 show, network communities eventually converge to county borders and then regions, however, this convergence is imperfect. Communities often flow over county borders, especially when the influence of a large town is present. Additionally, the history of urban development is highly uneven across Hungary. The eastern half used to be less urbanized, relying more on agriculture (Tóth 1994). Here, the convergence from small communities to large ones is much faster. At the same resolution, clusters in the eastern half resemble the size of regions, while in the western half, they are the size of counties or even smaller. While it is true that the average size of settlements is smaller in the western half than in the southeastern part, they are not smaller than in the northeast. The length and history of urbanization processes in the western half may be a factor explaining network communities remaining smaller on larger scales. The role of smaller towns in the development of rural areas is an important question in regional research (Najarsadeghi and Dorostkar, 2022), and our results show that this influence varies based on different societal and historical factors and not just on population size.

In Figure 6 A, we also see that some towns lose their central role in their respective communities at much lower scales than some villages. Moreover, some important cities with the administrative rank of a regional center lose their own communities sooner than some villages and smaller towns. This shows that, with respect to internal migration influence, there is a mismatch between official administrative status and actual influence. Thus, in line with previous research that pointed out the problematic nature of simple rural-urban delineations (Tacoli 1998), our results indicate that approaches that focus only on large towns, and select samples based on predefined urban areas can overlook places and mechanisms that are influential in migration systems. Therefore, analysing the polycentric nature of different regions should be done by first identifying different subcenters through data-based methods.

Further, while we identified large polycentric network communities at higher resolutions where multiple administratively important towns and cities are present (in

line with classical polycentric development theory (Meijers, 2007), the polycentric nature of most identified communities means one administratively important town and one or more influential settlements without an administrative title. Our findings show that polycentricity can work on different scales, even across scales. This underlines the problem raised by van Meeteren et al. 2016: adaptation of polycentric development theory should rely on well-defined definitions.

Budapest is the largest and most economically important city in Hungary as its capital. Population growth through immigration is the largest in the agglomeration zone of Budapest, with almost a third of the country living in the central region (Kovács et al., 2019). Our results also show that most long-distance migration is directed towards the capital, so the cluster around it starts to incorporate all the other clusters above a resolution. This aligns with the classical step-migration model, where most migrants tend to move upwards in the urban hierarchy in one or two steps (Plane et al., 2005). However, the scale at which the module of a large city is incorporated into the module of the capital shows that the step-migration model is a strong generalization. If the module of a city gets incorporated fast, at lower scales, it suggests that more migration happens towards the capital relative to other towns and cities that retain their modules longer. This allows us to measure how resilient a city can be against the economic pulling force of the largest metropolis in the migration system, an important aspect of possible inequalities in urban hierarchies (Lerch, 2016). The difference between the radiation model and the observed network is also the most striking in the case of Budapest. In the radiation network, we see that the module of the capital grows more slowly as we raise our resolution parameter. Some of this difference can be explained by the fact that the radiation model predicts more distant connections with lower accuracy (Kluge et al., 2022).

However, the main difference lies in our finding that the model can predict urban-rural flows relatively accurately but cannot predict rural-urban flows as they work in reality. Failing to predict this defining type of flow shows that the classical understanding of urban agglomeration zones – the idea that population is simply pouring out from large centers to available locations – is flawed. While population outflow can be explained as a simple consequence of large population concentrations and nearby available receiving settlements, the backward direction has to be

understood more as a social phenomenon. This is not only important in the case of Budapest but also in the case of smaller and mid-sized towns across the country.

5. Summary and conclusion

In this article, we aimed to map and understand the multi-scale structure of internal migration networks. Traditionally, internal migration analysis focused on the individual locations as the base unit of explanation. However, in the age of ever-growing interconnectedness, defined by polycentric urban regions, analysis based on individual units has limited explanatory power. Therefore, our analysis joins the growing literature based on the idea of inter-urban networks. However, the literature on inter-urban mobility networks usually considers connections between larger cities with great economic importance. Such analysis can show ranks and relations of large urban centers but is unable to describe how single centers can keep their central role in their local geographical and societal context. We contribute to this literature by looking at settlements of every size and urbanization level, being able to analyse on different spatial scales the emergent structure originating from different migration flows. We modelled internal migration in Hungary between 2015 and 19 as a weighted and directed network, where nodes are settlements, and edges between the nodes are migration flows. To understand the structure on different scales, we used community detection with the map equation with Markov time scaling. We use this approach as a multi-focal lens, with which we are able to zoom from small, dense, local groups of small settlements to large, long-distance polycentric urban regions.

By simple descriptive analysis and comparison of the real and radiation-predicted networks, we see that in migration is happening through a higher number of streams that are on average weaker (representing fewer migrants) compared to what we would expect based on population and distance. The clustering of the observed network is also lower than that of our null model. The real migration network is less concentrated in terms of strength but exhibits a more hierarchical (star-like) structure. Further, the relationship between the number of connections and the strength of these connections shows clear super-linear scaling across the urban hierarchy in the case of the observed network. Larger towns with regional center roles not only have more connections than smaller towns and villages, but the number of migrants moving

through these connections is even higher, meaning connections of larger localities are more efficient. This is not the case of the radiation network, where the relationship is more close to a linear one. We also analysed the strongest connections for every settlement. In the case of the observed migration network, we see that the share of the most important migration connection in all migrations of a given settlement is relatively high, around 25% on average. For smaller villages, this share is even higher, reaching over 50%, meaning that half of the total population flow connects them to only one settlement.

Our first research question asked: What elements of the flow structure connect the different scales of an internal migration network? We find that migration is structured by a hierarchical but interconnected system where small settlements cluster into local communities, and these local communities connect to towns with local center roles.

In particular, we observe that the strongest migration links tend to bridge smaller settlements to urban centers, rather than forming exclusively within settlement categories. The radiation model is able to predict this mixing in the case of in-flows but fails to predict it in the case of out-flows. This means that if only population and distance defined migration processes, people would stay inside the settlement category of their origin, and the strongest flows would connect village to village, and town to town.

Further, this mixing effect of rural-urban flows is also reflected in a strong multi-scale community structure defined by large urban areas. When we set our multi-focal lens to look at a very granular picture, we find dense, clique-like (all-to-all) communities of a few settlements. As we set the lens to find higher distance units, we see a process of these cliques connecting to each other through their migration connections to local towns and cities with center roles. As we lower the resolution, communities cluster around the most influential towns and regional centers. The administrative structure of the country emerges from the community formation at high parameter values: communities first converge into counties, and after that, regions. At the lowest resolutions, the community of the capital incorporates all other communities, because it is the destination of most long-distance migration in the country. At a high enough parameter value, the whole country becomes one migration community. The importance of the regional centers connecting the small

communities is indicated by the hierarchical nature of the community structure, that is once the area of a smaller town joined to a center, they remain together at the further, higher scales. This feature is also not explained by the radiation model.

Our second research question investigated whether population size and geographic distance alone are sufficient to model connections between different scales in the network. By comparing our observed migration network to the radiation model, we find that these factors alone cannot fully account for migration patterns. While the radiation model captures some broad trends, it systematically fails to predict migration flows from local cliques to urban centers. This suggests that additional factors – such as historical ties, economic linkages, and institutional influences – play a role in shaping migration dynamics beyond simple demographic mechanisms.

The multi-scale structure of the radiation network shows very different characteristics from that of the real migration network. Local clique formation happens there too, but the convergence process of local cliques connecting to larger communities is much slower. This slower convergence to higher-level units leads to another difference: communities do not connect into urban agglomeration zones, counties, or regions in any part of the country. This finding shows that the relations of large urban agglomeration zones are not predictable purely based on population and distance. Thus, the process of how these urban zones circulate populations between each other is not only demographic but also social and economic in nature.

2nd ARTICLE:

INTERNAL MIGRATION FLOW NETWORK
REORGANIZATION DURING A HOUSE PRICE
SHOCK IN A CAPITAL CITY

Internal migration flow network reorganization during a house price shock in a capital city

Gergely Mónus

Abstract:

Existing research establishes that house prices affect migration propensities and directions for individual locations. Most studies treat bilateral flows as independent, omitting the systematic nature of migration networks. We lack evidence on how entire flow systems respond when a dominant hub becomes unaffordable. Using Hungary's inter-district migration network during Hungary's 2010s housing boom, we apply exponential random graph models for count-valued networks to test whether rising central city prices trigger immediate flow network reorganization. Comparing three periods with distinctly different price dynamics between 2010 and 2018, we find network expansion rather than structural reorganization: edge counts grew 27% and migration volume increased 38%, while Budapest's share of flows declined only modestly. Triangles with Budapest as middleman multiplied without increasing direct bypass flows. However, ERGM results reveal differentiation invisible in descriptive statistics: regional centers and smaller towns showed significant effects as both population retainers and migration destinations, even when controlling for network mechanisms.

1. Introduction

Internal migration is the long-term residential relocation of people from one locality to another within a country. In European countries with low natural population increase, internal migration has replaced fertility and mortality as the primary demographic process shaping spatial distribution of populations (Rowe and Patias 2020). Despite this, internal migration remains understudied compared to international migration, especially regarding its systematic properties. Traditionally, internal migration is attributed to the attraction force of economic opportunity, that large centers of commerce and industry pulling populations from other locations (Ravenstein 1885). Long distance moves are typically explained by employment factors, while short distance moves (residential mobility) are explained by housing opportunities (Niedomysl 2011). Housing thus has a central position in internal migration theory, but its role is usually analysed at the level of individual decisions, not at the level of entire migration systems.

The first modern social scientific work on migration, *Laws of Migration* by Ernst Ravenstein, already conceptualized internal migration as a system of interconnected flows. Ravenstein observed that people move towards centers of commerce and industry, and the vacancies they create are subsequently filled by migrants from more remote areas, producing what he called "currents" of migration (Ravenstein 1885). This notion led to the emergence of push-pull thinking in internal migration research, originating from Lee 1966. Lee's seminal paper is typically cited as the canonical foundation of push-pull theory, but contemporary usage tends to reduce his framework to a plus-minus calculus of economic and social factors at origins and destinations. This interpretation misses that Lee explicitly embedded push-pull dynamics within flow structures, "Migration tends to take place largely within well-defined streams", and "For every major migration stream, a counter-stream develops". His plus-minus calculus operates through streams (Lee 1966). Mabogunje 1970 developed this relational thinking further through his systems approach, conceptualizing rural-urban migration as a circular, self-modifying system where origins and destinations are linked through feedback mechanisms.

Surprisingly, the systematic elements of early foundational internal migration research like Ravenstein, Lee, or Mabogunje mostly disappeared from empirical research on the area. The dominant empirical approach has been to model migration as a function of origin and destination characteristics, what might be termed an 'individual unit' approach, or to estimate bilateral flows between origin-destination pairs through gravity models, what we can call an 'individual dyad' approach. Comprehensive reviews of this literature (Greenwood 1975, 1985, 1997, Lucas 1997; Cushing and Poot 2004, Molloy et al. 2011, Etzo 2008) organize the field around identifying determinants of migration at the settlement, county, or regional level. Individual unit analyses usually model gross or net flow as a function of local characteristics, independent from other locations. Individual dyad analyses model net flow between two locations, or gross flow in each direction, and it acknowledges spatial structure through distance decay. But it treats each dyad independent from all other dyads, estimating migration as if the flow from one region to another was unaffected by other flows in the same migration system.

The relational nature of migration systems translates naturally to network representation, where localities are nodes and flows are edges. Researchers mapped

various interurban networks during recent years – transportation (Dong et al. 2016), telecommunication (Krings et al. 2009), corporate ties (Zhao et al. 2017), online social connections (Lengyel et al. 2015). It is also shown that online information flow between cities and population flow is highly correlated (Hu et al. 2020; Lőrincz and Németh 2022). Network models proved useful to map many population flow characteristics: delineating source and sink regions (Wang et al. 2019), identifying temporally coherent settlement subgroups (Zhang et al. 2020), revealing rich-club characteristics where well-connected cities preferentially exchange population (Zhu et al. 2021). The network perspective enabled analysis of polycentric urban regions (De Goei et al. 2010; Liu et al. 2016), differentiating central cities from hinterlands based on actual flows rather than administrative boundaries (Hu et al. 2024). For Hungary, Mónus and Lőrincz 2025 show that multi-scale community detection reveals hierarchical structure in migration networks.

Network science provides tools to operationalize the relational ideas of classical migration theory that disappeared from empirical research. The shift from central place theory to network-based hierarchies in urban studies (Neal 2011) gives theoretical grounding for this. Systems of central places transform to network systems through transition from centrality to nodality, from vertical to horizontal accessibility, from primacy to complementarity (Batten 1995). But more importantly, network methods enable modeling of edge dependencies, the fact that one flow's presence and magnitude affects others in the system, reframing the effect of pushing and pulling migration factors as the original formulation of Lee 1966. Standard regression approaches assume these dependencies away. Statistical models for networks, especially exponential random graph models extended to valued networks (Krivitsky 2012), allow estimation of flow determinants while controlling for network structures like reciprocity and heterogeneous node activity. As Topa and Zenou 2015 summarize, very few empirical studies test interactions between urban space and flow structure. Our analysis aims to address this.

The new inter-urban network literature aims to model population flow networks systematically. This approach enables us to revisit classical push-pull migration factors - not as isolated forces but as elements operating within interconnected spatial systems. Among these factors, housing stands out as particularly important for internal migration. Housing and internal migration have a long established bi-

directional relationship, mainly emerging from the fact that housing availability acts as both an attractor and a barrier to migration (Mulder 2006). According to the literature, one of the most important aspect of housing that makes a mark on mobility behaviours is the sheer value of a home, and how it changes in certain market conditions. The ‘lock-in-effect’, as originally formulated by Quigley, refers to the mobility reducing effect of rising mortgage interest rates (Quigley 2002). Generally, most housing related financial frictions can have a similar effect: lock-in from down payment constraints, lock-in from equity constraints, or lock-in from loss aversion (Stein 1995, Chan 2001, Engelhardt 2003). This effect can be even more influential during housing bust periods, as demonstrated by Henley 1998 for the UK during the early 1990s or by Ferreira et al 2011 for post-2008 USA. The question whether rising house prices have the opposite effect by enhancing the probability of migrating had less attention from researchers. However, for the USA between 2002-2010 Goetz 2013 shows that a 20% increase in house prices raised the probability of migrating between metropolitan areas by 7.5%. Similarly, Ermisch and Washbrook 2012 found that in the UK between 1992-2008 increasing housing equity encouraged residential mobility. Another important aspect of the effect of house prices on internal migration structure is temporality. Sources cited above focus on heterogenous timescales, with long-term panel studies being the most common – but differences in housing price effects in different timescales, and especially possible short-term reactions in migration systems to market shocks remain relatively unclear. For example, Peng and Tsai 2019 shows for Taiwan between 1994-2016 that prices and migration are cointegrated with positive long-run relationships, but short-term effects can be asymmetric - negative at low migration volumes and positive only at high volumes.

The well-documented lock-in effect, or the effect of rising house prices both only refer to migration propensity, and not the direction of migration events, i.e. migration from rural areas to urban areas, or migration between given geographical regions. The effect of housing on direction of migration streams is relatively understudied, which is surprising, since housing markets and opportunities are strikingly different across the urban hierarchy, and between geographical regions with different economic contexts (Partridge et al 2009, Gyourko et al. 2013). Area-level housing price measures often offer counter-intuitive positive effect for migration directions, where higher prices attract migration. This can be attributed to the fact that

population attractor locations tend to have high prices, such as large cities or capitals. But after accounting for rental and ownership decisions, the costs of rental and owned properties, and the costs of holding housing capital, models predicting individual behaviour show that higher housing costs reduce the probability that a metropolitan area is selected (Plantinga 2013). Also, modeling price change effects instead of price level effects show that housing costs operate as a deterrent to destination choice, and this effect is stronger than the price push factor from origins, -- this asymmetric effect was documented for both Germany (Stawarz et al 2021), the USA (Olney et al 2024), and Australia (Siriban et al 2025). However, the nature of this asymmetry can be heterogeneous across countries. Cavalleri et al. 2021 apply a gravity framework to model inter-regional migration flows in 14 OECD countries, and find that house prices in destinations act as a barrier for migration everywhere, but effect of higher prices in origins varies between countries, it is either a push or a retaining effect.

Research on the effect of house prices on direction of migration flows mostly focuses on bilateral flow models, this way assuming independency of flows, omitting the systematic, interdependent nature of inter-urban flow structure. Our analysis aims to overcome this by applying statistical models explicitly treating edge-dependence, count ERGMs. This way we are able to estimate the effect of house prices and other contextual effects controlling for flow mechanisms like mutuality. Such methods enable us to ask questions that previously were not in the focus of the literature, for example how changes in one part of the migration system affect the whole of it. We propose Hungary during the 2010's decade as a case study. Hungary is special from internal migration and housing aspect in that period. Budapest, the capital is an outstandingly dominant city in the country measured both in population and economic activity. Most of internal migration events are directed either to the capital or its broader agglomeration. Internal migration peaked during this decade, while the country experienced an extreme house price boom, originating from Budapest. Our systematic network approach enables us to analyse how the flow structure (re)organized during this period, following a price shock in the dominant central city.

After regime change in 1990, internal migration in Hungary was defined by suburbanization processes, mostly around Budapest. After the 2008 crisis, Budapest briefly had a positive migration balance, but after 2016, the capital and large county

seats started to lose population again (Bálint and Obádovics 2021). Generally between the regime change and the early 2020s, population concentration in major cities and their regions increased from 54.62% to 57.19%, but the share of non-Budapest major city regions only grew slightly (Pénzes 2024). According to the Hungarian Central Statistical Office during the 2022 census, 17.5% of Hungarian lived in the capital, and 31.4% lived in the capital and Pest county. The ratio of the population of the capital metropolitan area of Hungary relative to the largest non-capital metropolitan area is exceptionally high in Europe. The Budapest metropolitan area is more than 9 times bigger than the second one, which is the 13th largest in the global ranking of the OECD (with only Asian, Latin-American, and African countries in the first 12) (OECD 2020).

While inter-regional migration in Eastern European transition countries has been researched to some extent, the effect of housing on internal migration in this geography is relatively understudied (Bloze 2009). The Hungarian housing system is also an outlier in Europe. After the regime change, Hungary implemented a ‘give-away privatization’ strategy of the previously state-owned rental housing stock, leading to a homeownership rate around 90%, one of the highest in the EU/OECD (Kováts and Kohl 2025). Mulder 2006 shows that homeownership reduces internal migration probability. International comparisons generally underline that high homeownership nations in Europe are less migratory than nations with high renter ratios. However, this only refers to baseline internal migration rates, and does not reveal possible differences in the weight of housing affordability in different internal migration systems. For example panel analysis of inter-county migration of United States home owners and renters shows that home owners are more responsive to housing cost burdens than renters when making moving decisions, and that they avoid high-cost areas both as origins and destinations (Xu 2024). What makes the question of owner-occupation in Hungary special is the price boom in the country, and especially Budapest after 2014, so the middle of our analysis period. Between 2014 and 2024, house prices in the EU grew on average by 50%, but Hungary is the only country where nominal house price growth exceeded 300%. Between 2014-2024, Hungary had also the 4th largest increase in price-to-income ratio among EU nations (European Commission 2025). The house price shock started in Budapest,

and while by time it rippled through the urban hierarchy, house price growth in the capital exceeds other locations significantly still today.

The 9 years covered by our analysis were strongly affected by the 2008 crisis. Hungary was also heavily affected by the global financial crisis, which was reflected both on house price dynamics and internal migration rates. The effect of housing busts on internal mobility is widely documented, mostly for western geographies (see for example strong evidence for post-boom lock-in effects by Henley 1998 for the UK, or by Ferreira et al 2010 and Ferreira et al 2011 for the US). Rapid decrease in price and mobility in the beginning of our study period, and rapid increase in both in the last third makes our analysis stronger, because it introduces heterogeneity in migration behaviours even on the shortest timescale.

Our analysis addresses whether internal migration networks respond to dominant central city price shocks in real-time through rapid structural adjustment, or whether hub-centered patterns persist despite substantial affordability deterioration. Hungary provides an exceptional case for examining this question. Budapest's metropolitan area is more than nine times larger than the country's second city, representing a level of urban dominance found otherwise only in Asian, Latin American, and African countries. This extraordinarily hub-centered system experienced a severe housing price shock: between 2014 and 2024, Hungarian nominal house prices grew by over 300%, with the boom originating in and most intensely concentrated in Budapest. Our analysis work on a very short time scale: it covers only 9 years. In the literature there is evidence for long-run relationships, but we aim to answer whether an abrupt, swift price shock in a dominant, central city is followed by immediate reorganization across the whole migration system, or some parts of it. Existing research shows house prices affect migration propensities and directions, but most studies rely on gravity models that treat bilateral flows as independent from one another. These methods omit the systematic, interdependent nature of migration networks that classical theorists like Ravenstein and Lee emphasized. We lack evidence on how entire flow systems reorganize when a dominant hub becomes unaffordable. We ask the research question:

Do internal migration networks respond to central city price shocks in real-time through rapid structural adjustment, or do hub-centered patterns persist despite substantial affordability deterioration?

To address this gap, we employ exponential random graph models for count-valued networks, which allow us to estimate the effects of housing prices and other contextual factors while explicitly modeling network mechanisms. Examining three periods between 2010 and 2018 that are distinctly different in house price dynamics, we find that despite a 162% deterioration in the capital's affordability ratio, the migration network expanded rather than reorganized, and while Budapest kept its hub dominance, some differentiation across the urban hierarchy indeed happened.

2. Data and Methods

2.1. Data

We use administrative data sources for all steps in our analysis. We construct a flow network for Hungary on the district level. Districts are the equivalent of local administrative units in the European Union nomenclature (LAU1, formerly NUTS4), containing on the average around 20 settlements each (3155 settlements, 174 districts without Budapest). In the official nomenclature, Budapest consists 23 districts, which we treat as one district (to omit inter-Budapest residential mobility), resulting in a network of 175 nodes. Inter-district migration is hard to differentiate from residential mobility, because while it may include people crossing municipality borders, geographically these moves may be very short distance, and many times does not require changing jobs or general context of everyday life. The directed, valued edges in our networks are migration flows, computed from permanent residential address changes. These address changes are collected by Hungarian Central Statistical Office (HCSO). We use only permanent migration events because temporal migration is prone to definition changes in the legal system, and can behave irregularly in different time periods (Bálint and Obádovics 2021). Internal migration data is aggregated to different geographical levels, some available openly online, and district level data is available upon request. HCSO also collects information on house prices, based on official transaction contract data from the National Tax and Customs

Administration. Transaction contract data is aggregated to the settlement level, and it is openly available online². We aggregate settlement level data to the district level.

As controls, we use different district measures: population, amenity and economic indexes, and an urban hierarchy rank, all coming from HCSO databases on municipalities (“Tájékoztatási adatbázis”). Amenity index is made as a principal component of retail stores, kindergartens, primary schools, and general medical practitioners; and the economic index is made as a principal component from enterprise built dwellings (yearly), total income tax, and registered job seekers. All of these variables were computed per capita for the settlements inside the districts, and averaged across the settlements. Urban hierarchy ranks can be: Budapest, agglomeration center (largest towns of the Budapest agglomeration), regional centers (towns of county rank in official nomenclatures, outside of the Budapest agglomeration), and other towns. A district gets the rank of the highest ranked settlement in it, because districts of larger towns tend to contain mostly them and their agglomeration. Agglomeration centers serve as the reference category for our estimates, representing a mid-level position in the urban hierarchy that allows comparison both upward (to Budapest) and downward (to regional centers and other towns).

2.2. Modeling strategy

To be able to model the Hungarian inter-district migration flow network systematically, we have to use a modeling framework that explicitly incorporates dependence structure between flows. Exponential random graph models provide a flexible statistical framework for this, being able to handle both node (county) characteristics like house prices or populations, edge attributes (like distance), and network mechanisms (like mutuality). The approach originates from Holland and Leinhardt 1981, who introduced an exponential family of probability distributions for directed graphs that could simultaneously estimate parameters for reciprocation tendencies and differential node attractiveness. This framework was subsequently extended to capture more complex dependencies (Snijders et al. 2006).

² <https://www.ksh.hu/s/ingatlanadattar/>

Classical formulation of ERGMs are restricted to binary network data, where ties only refer to a connection being present or absent. For flow networks, where most of the information is not in the presence of ties, but the volume that moves through them, this limits their applicability. It is possible to dichotomize ties according to some threshold in flow volume, but this omits information about the flow structure. Krivitsky 2012 addressed this limitation by introducing a generalized ERGM framework for valued networks, particularly for count-valued ties. The extension lies in the introduction of a reference measure ($h(y)$), that defines the baseline shape of edge value distribution. The formulation of the generalized ERGM models follows the equation:

$$\Pr_{\theta; h, \eta, g}(Y = y) = \frac{h(y) \exp(\eta(\theta) \cdot g(y))}{\kappa_{h, \eta, g}(\theta)}$$

Where the normalizing constant is:

$$\kappa_{h, \eta, g}(\theta) = \sum_{y \in \mathcal{Y}} h(y) \exp(\eta(\theta) \cdot g(y))$$

Y is the random network variable and y its realization; θ is the model parameter vector; η maps model parameters to canonical parameters; $g(y)$ is a vector of sufficient statistics capturing network features of interest; $h(y)$ is the reference measure that defines the baseline distribution of edge values; and κ is the normalizing constant ensuring the probabilities sum to one. For a Poisson-reference ERGM, $h(y) = \prod (y_{i,j})^{-1}$, which provides a natural baseline for count data and, under dyadic independence, reduces to Poisson regression. To be able to use the Poisson as a reference distribution, we have to transform edge weights. First, we have to log-transform them, and after we have to round them to integers. This omits some information and detail in flow volumes, but differentiates between flows with minimal and large volumes properly.

Differentiating between level effects and change effects is crucial in understanding the effect of prices on internal migration flow structure. As Stawarz et al. 2021 demonstrate, cross-sectional models that capture level effects can show high house prices as a pull factor, because expensive destinations tend to be attractive places. Their approach is to use panel models isolating change effects. Panel specifications in the ERGM framework are absent. There are extended model formulations that

incorporate temporal effects, but with strong limitations, especially for valued networks (Krivitsky and Handcock, 2014, Leifeld et al., 2018, Huang and Butts, 2023). Instead, we fit cross-sectional models for three 3-year periods that are remarkably different in house price dynamics; and after we fit a cross sectional model for the 3rd period using past period flow structure as an edge attribute (analogous to a lagged dependent variable approach in panel models) modeling how node attributes and network mechanisms affect the change of flows between periods.

To understand what mechanisms shape the structure of migration flows, we fit two models for all three periods. The first models only contain structural terms (that model properties of the distribution of flows), district level attributes, and distance as an edge covariate. In other words, these models assume structural independence, and are equivalent to Poisson regressions. The second models include two network mechanisms. Comparing the models for the same periods shows us what are the real effects of push-pull factors after controlling for the flow structure, and comparing models across periods can reveal how the structure of flows changed through time. Following Huang and Butts 2023, we incorporate two main network terms: mutuality, and sender/receiver activity variance. The former accounts for the widely documented nature of migration flows that they tend to be reciprocal. The second measure accounts for the fact that some nodes are just inherently more active than others, and not accounting for it might contaminate other estimates. Huang and Butts 2023 uses a different measure for this (waypoint flow). The sender-receiver variance we propose follows this equation:

$$g_{\text{actor cov.}}(y) = \sum_{i \in N} \frac{1}{n-2} \sum_{j,k \in Y_i \wedge j < k} (\sqrt{y_{i,j}} - \sqrt{\bar{y}})(\sqrt{y_{i,k}} - \sqrt{\bar{y}}),$$

The formally called the actor covariance statistic quantifies the extent to which individual actors exhibit consistent patterns in the values they assign to their outgoing ties. The outer summation iterates over all actors i in the node set N , while the inner summation considers all unique pairs of alters (j, k) to whom actor i has directed ties. The term $\sqrt{y_{i,j}}$ represents the square-root transformation of the edge value from actor i to alter j . The quantity $\sqrt{\bar{y}}$ denotes the network-wide mean of these transformed edge values. The normalization factor $1/(n-2)$ adjusts for the number of possible alter pairs. Reciprocity as a network measure only works inside dyads, so

including it only does not achieve dependence between dyads. But our sender-receiver activity measures achieve dependence between dyads by accounting for covariance between out- and in-flows for a given node. Applying the two simultaneously thus makes our model specifications able to control for dependence between different dyads.

3. Results

3.1. Change in push-pull factors across the study periods

Table 1: Evolution of Budapest-to-country ratios for key migration determinants									
Period	Period 1			Period 2			Period 3		
Year	2010	2011	2012	2013	2014	2015	2016	2017	2018
Retail stores	2.141	2.144	2.222	2.272	2.181	2.116	2.031	1.981	1.961
General practitioners	1.204	1.196	1.202	1.21	1.186	1.194	1.2	1.201	1.206
Kindergartens	0.47	0.471	0.482	0.48	0.483	0.486	0.495	0.499	0.496
Primary schools	0.419	0.406	0.414	0.398	0.404	0.405	0.407	0.414	0.412
Dwellings built by enterprises	16.89	15.73	5.086	13.29	12.61	12.90	1.74	6.636	6.997
Total income tax	2.973	2.56	2.029	1.836	1.756	1.717	1.714	1.692	1.687
Registered job seekers	0.306	0.301	0.329	0.367	0.338	0.337	0.299	0.252	0.22
House price	2.756	2.809	2.821	2.812	2.943	3.308	3.704	3.854	4.091

Table 1. presents the evolution of Budapest-to-country ratios for key migration determinants from 2010 to 2018. The most pronounced change occurs in housing affordability. Budapest's house price ratio increased from 2.76 to 4.09 (48% increase), while the income tax per capita ratio – our proxy for income differentials – decreased from 2.97 to 1.69 (43% decrease). This affordability crisis intensified particularly after 2014, coinciding with the Period 2 and Period 3 observations in our analysis where Budapest's migration premium began declining despite maintaining hub dominance. It is important to note, that the house price boom happened within a specific spatial and temporal structure across the country: embedded in the urban

hierarchy. The price boom started in the capital, and rippled downwards in urban ranks. Appendix A.2. provides a more detailed picture of this spillover-type process.

Housing construction patterns present a potential confounding factor, since across the years the premium of Budapest shrunk in this measure. Budapest's enterprise-built dwelling ratio shows volatility, peaking at 16.9 in 2010 before collapsing to 5.1 in 2012, then fluctuating between 6.6 and 13.7 in subsequent years. These early-period spikes occurred during especially low absolute construction levels nationwide following the 2008 crisis. By 2017-2018, when construction activity surged throughout Hungary, Budapest's relative advantage had largely dissipated (ratio of 7.0), suggesting that housing supply responses were insufficient to offset price pressures. This pattern aligns with our finding that network expansion accelerated in Period 3 perhaps because of increased construction activity dispersed across the country rather than concentrated in the capital.

Other push-pull factors display relatively modest changes that fail to explain observed network dynamics. Budapest's unemployment ratio improved from 0.31 to 0.22, indicating strengthening labor market advantages that should theoretically increase in-migration. Infrastructure differentials remain stable: general practitioners (1.20), kindergartens (0.47 to 0.50), and primary schools (0.42) show minimal variation across the period. The retail store ratio declining from 2.14 to 1.96 suggests some convergence in consumer amenities. These stable or improving relative advantages, combined with worsening affordability, support our interpretation that Budapest retained its attractiveness despite price barriers. The persistence of migration to Budapest despite a 162% deterioration in affordability underscores that housing prices alone cannot overcome agglomeration advantages in hub-dominated systems.

3.2. The inter-district migration network across the study period

Table 2: Main statistics of the inter-district migration network of Hungary			
Period	2010-2012	2013-2015	2016-2018
Number of nodes	175	175	175
Number of edges	5074	5265	6458
Total migration	380804	404835	525217
Mean degree	35.634	37.441	46.4
Mean strength	4352.046	4626.686	6002.48
Mean weight per edge	61.069	62.127	64.948
Density	0.167	0.173	0.212
Share of Budapest in edges	0.068	0.066	0.054
Share of Budapest inflow in total	0.207	0.214	0.183
Share of Budapest outflow in total	0.19	0.187	0.207
Share of sum Budapest flow in total	0.396	0.401	0.389
Global cycle cc unweighted	0.6	0.606	0.601
Global middleman cc unweighted	0.628	0.632	0.644
Global in cc unweighted	0.616	0.622	0.622
Global out cc unweighted	0.631	0.635	0.648
Global total cc unweighted	0.616	0.62	0.624
Global cycle cc weighted	0.005	0.006	0.003
Global middleman cc weighted	0.006	0.006	0.004
Global in cc weighted	0.006	0.006	0.004
Global out cc weighted	0.006	0.006	0.004
Global total cc weighted	0.006	0.006	0.004

Table 2. presents the structural evolution of Hungary's internal migration network across three periods. The network maintains 175 nodes throughout all periods, allowing direct structural comparison. The most evident pattern is increasing connectivity. Edge counts increase from 5,074 to 6,458 (27% growth), while total migration volume grows from 380,804 to 525,217 individuals (38% increase). Mean degree rises from 35.6 to 46.4, indicating that the average district connects to 11 additional locations by the final period. Network density increases from 0.167 to 0.212, suggesting more realized connections within the potential migration space. Mean edge weight shows modest growth from 61.1 to 64.9 migrants per connection,

indicating that expansion occurs primarily through new edges rather than intensification of existing routes.

Budapest's position within the network displays gradual and modest erosion across multiple metrics. Its share of network edges declines from 6.8% to 5.4%, while the share of Budapest inflow in all internal migration decreases from 20.7% to 18.3%. Outflow share remains relatively stable around 19-20%, and total Budapest-related flows (inflows plus outflows) decrease slightly from 39.6% to 38.9% of all migration. These patterns indicate that while Budapest maintains its hub status in absolute terms, its relative dominance moderately weakens as peripheral connections multiply.

Clustering coefficients reveal stable triadic structures despite network expansion. Unweighted clustering coefficients remain consistently high across all variants (cycle, middleman, in, out) at approximately 0.60-0.65, indicating that triadic closure is a persistent network feature. The weighted clustering coefficients however, are extremely low (0.003-0.006) and decline slightly over time. This dramatic gap between unweighted and weighted measures suggests that while triangles are structurally common, they carry relatively little migration volume compared to hub flows. The slight decline in weighted clustering during Period 3. coincides with the period of maximum network expansion.

These descriptive patterns – network expansion through edge creation, Budapest's slightly declining but persistent centrality, and stable structural clustering with low weighted triangles – suggest a system adding connections between countryside districts while maintaining its fundamental hub-dominated structure.

3.3. Analysis of triangles and weight transitivity in the migration network

Table 3. examines the micro-structure of migration triangles where Budapest serves as the middleman node, connecting origin and destination sub-counties. This configuration represents the most likely context for observing housing price barrier effects: migrants from countryside origins seeking Budapest, but potentially diverted to alternative destinations due to cost constraints. If network reorganization was

occurring, we would expect increasing direct flow shares between origin and destination nodes, bypassing the expensive hub.

The aggregate patterns show remarkable stability in flow distribution despite substantial growth in triangle counts. Across all Budapest-middleman triangles, the direct flow share remains minimal and relatively stable (6.1%, 5.8%, 4.9%), with Budapest inflows and outflows each capturing approximately 45-50% of triangle volume. The number of such triangles increases from 4,717 to 6,110 (30% growth), yet the proportional flow distribution hardly changes. The pattern of more triangles with similar internal weight distributions reinforces the network expansion interpretation rather than structural reorganization.

Disaggregating by triangle location types reveals pronounced spatial heterogeneity. Budapest agglomeration triangles (both origin and destination in the agglomeration) maintain extremely low direct flow shares (1.8%, 1.7%, 1.8%) with minimal triangle growth (253 to 268), suggesting that in the agglomeration of the capital, every sub-county is in a migration connection only with the capital. In contrast, countryside triangles (both origin and destination in the countryside) show higher direct flows (8.3%, 7.8%, 6.9%) and substantial increase (3,265 to 4,065 triangles, 25% growth). Mixed triangles (either the origin or the destination in the agglomeration) display the strongest hub dependence: agglomeration-to-countryside triangles channel 80% of flows through Budapest inflows with merely 1% direct flows, while countryside-to-agglomeration triangles route 85% through Budapest outflows with even lower direct shares (0.6% by Period 3).

Table 3: Analysis of middleman-triangles with Budapest as middleman					
All triangles					
Period	Budapest inflow share	Budapest outflow share	Direct flow share	Number of triangles	Triangle growth between periods
2010-2012	0.513	0.426	0.061	4717	
2013-2015	0.513	0.429	0.058	4917	1.042
2016-2018	0.488	0.463	0.049	6110	1.295
Budapest agglomeration triangles					
Period	Budapest inflow share	Budapest outflow share	Direct flow share	Number of triangles	Triangle growth between periods
2010-2012	0.435	0.547	0.018	253	
2013-2015	0.450	0.532	0.017	256	1.012
2016-2018	0.379	0.604	0.018	268	1.059
Countryside triangles					
Period	Budapest inflow share	Budapest outflow share	Direct flow share	Number of triangles	Triangle growth between periods
2010-2012	0.545	0.317	0.083	3265	
2013-2015	0.547	0.375	0.078	3408	1.044
2016-2018	0.525	0.406	0.069	4065	1.245
Agglomeration - countryside triangles					
Period	Budapest inflow share	Budapest outflow share	Direct flow share	Number of triangles	Triangle growth between periods
2010-2012	0.810	0.177	0.012	494	
2013-2015	0.824	0.164	0.012	503	1.018
2016-2018	0.800	0.189	0.011	747	1.512
Countryside - agglomeration triangles					
Period	Budapest inflow share	Budapest outflow share	Direct flow share	Number of triangles	Triangle growth between periods
2010-2012	0.184	0.808	0.008	705	
2013-2015	0.177	0.815	0.008	750	1.064
2016-2018	0.142	0.852	0.006	1030	1.461

The evolution of Budapest's share within triangles shows modest shifts rather than dramatic change. In countryside triangles, Budapest's outflow share increases from 31.7% to 40.6%, suggesting the capital increasingly serves as a stepping stone from

rural areas. Conversely, in mixed triangles, Budapest's role intensifies rather than diminishes – capturing over 80% of flows in agglomeration-countryside connections and showing increased outflow shares in countryside-agglomeration patterns. These patterns contradict simple substitution hypotheses where peripheral locations would increasingly bypass the expensive hub.

The weighted triangle results align with the network-level descriptive statistics: while triangles are structurally common (particularly in countryside networks), they carry minimal migration volume relative to hub-spoke flows. Direct flows between origin and destination nodes remain below 10% even in countryside triangles and approach zero in agglomeration-related configurations. But neither global network statistics, nor the triangle analysis can answer the question whether the growth of connectivity with constant (or slightly diminishing) hub status of the capital happened according to some concrete pattern, structural differentiation, or whether it happened evenly across all districts. To be able to answer this, we use ERGM modeling.

3.4. ERGM results

While structural terms are usually only considered as technicalities, they can reveal interesting findings about the migration system of Hungary. The baseline migration parameter (sum) shifts from negative to positive when network effects are introduced, in all periods. This indicates that the negative baseline in non-network models partially reflect unmodeled heterogeneity in destination attractiveness and reciprocity patterns. Once these network structures are accounted for, the underlying baseline migration rate appears higher, with variation attributed to systematic network processes rather than suppressed through a lower intercept. Interestingly, for the model with network effect for period 3, intercept is not significant. Intuitively this indicates that the included covariates and network effects explain flow structure to a higher extent, there is little residual baseline migration not attributable to modelled factors. While the non-significant intercept in Period 3 suggests the model captures most systematic variation, the persistent negative CMP parameter indicates that migration flows remain underdispersed across all specifications. This suggests additional unobserved regularities remain even after controlling for spatial, socioeconomic, and network effects, possibly reflecting institutional constraints or established migration channels, that are not fully captured by the modelled

covariates. The significant negative nonzero parameter for all models indicates zero-inflation in migration flows: there are more dyads with zero migration than a Poisson distribution would predict. This is expected given that most origin-destination pairs have no migration between them.

Classic gravity variables show an interesting picture. Distance and destination population both have an expected effect in all models and periods: longer distances reduce migration flows, and larger populations lead to larger flow volumes. However, origin population also shows a positive effect on flow volume, showing that districts with larger populations send larger flows. Because we control for economic activity, administrative rank, and even for sender activity variance, this finding indicates that larger districts actually have an active role in sending populations to other locations (such finding could possibly be connected to agglomeration formation).

An affordability hypothesis for the effect of house prices would dictate that high origin house prices make people leave, and high destination house prices deter new population. Contrary to this, in all periods, and both models with and without network effects, origin house prices have a negative parameter, suggesting that people do not leave expensive places on average. Destination house prices have a positive effect, also in all models. This either shows, that even after controlling for amenity and economic factors, rising house prices proxy economic opportunities for internal migrants; or simply that districts (or just Budapest) did not become expensive enough that would mean a barrier for internal migration. Or maybe that through reverse causality migration increases house prices, leading to a positive correlation in cross-sections.

The economic and amenity indexes we include show a mixed picture. For origins, they remain positive and significant for all periods and models, meaning that economically active districts, with more local amenities send more population than poorer, less developed districts. This finding connects to the finding about origin populations. On the other hand, destination economic and amenity indexes are the only variables besides the urban hierarchy categories that become insignificant after controlling for network effects, in multiple periods. This shows that the attractive

force of these pull-factors disappear after we consider the relative places of districts in the flow structure.

Table 4: ERGM results for the three periods												
	Period 1				Period 2				Period 3			
	Without network effects		With network effects		Without network effects		With network effects		Without network effects		With network effects	
Variable	Est.	Sig	Est.	Sig	Est.	Sig	Est.	Sig	Est.	Sig	Est.	Sig
Sum	-0.987	*	1.389	***	-1.755	***	0.749	*	-2.938	***	0.322	
Non zero	-11.98	***	-9.831	***	-11.683	***	-9.712	***	-11.451	***	-9.59	***
CMP (dispersion)	-4.877	***	-3.545	***	-4.789	***	-3.535	***	-4.744	***	-3.53	***
Distance	-1.007	***	-0.687	***	-1.012	***	-0.706	***	-0.973	***	-0.696	***
O: Capital	0.523	***	0.194		0.635	***	0.299	*	0.669	***	0.376	**
O: Regional centers	-0.092	*	0.008		-0.103	**	-0.05		-0.122	***	-0.137	***
O: Other towns	-0.16	***	0.023		-0.175	***	-0.056		-0.152	***	-0.148	***
D: Capital	0.944	***	1.399	***	0.968	***	1.405	***	0.92	***	1.377	***
D: Regional centers	-0.213	***	-0.107	**	-0.175	***	-0.038		-0.086	**	0.046	
D: Other towns	-0.338	***	-0.169	***	-0.293	***	-0.095	**	-0.096	**	0.061	*
Same region	0.517	***	0.369	***	0.521	***	0.379	***	0.525	***	0.396	***
O: House price	-0.329	***	-0.513	***	-0.311	***	-0.499	***	-0.296	***	-0.46	***
O: Population	0.792	***	0.68	***	0.783	***	0.698	***	0.825	***	0.714	***
O: Economy index	0.178	***	0.158	***	0.209	***	0.181	***	0.188	***	0.104	***
O: Amenity index	0.144	***	0.113	***	0.132	***	0.133	***	0.069	***	0.09	***
D: House price	0.082	***	0.337	***	0.127	***	0.348	***	0.157	***	0.329	***
D: Population	0.675	***	0.132	***	0.678	***	0.156	***	0.678	***	0.167	***
D: Economy index	0.11	***	-0.029		0.142	***	-0.014		0.214	***	0.064	**
D: Amenity index	0.139	***	0.074	***	0.089	***	0.026		0.011		-0.009	
Mutuality			1.34	***			1.271	***			1.129	***
Receiver activity variance			0.858	***			0.839	***			0.729	***
Sender activity variance			-0.123				-0.215	**			-0.055	
Significance stars: *** p<0.001, ** p<0.01, * p<0.05												
O: origin variables, D: destination variables												

Directions across the urban hierarchy offer an interesting modeling result underlining the importance of network specifications. For origin districts, administrative levels in the middle and lower urban hierarchy show a similar picture for the first two periods: after we control for network effects, they lose their significant effect. However, in

period 3, they do not, the capital has a significant and positive effect, and regional centers and other towns both have a significant and negative effect. This

means, that even after controlling for reciprocity and sender activity variance, in period 3, certain levels of the urban hierarchy started to show additional effect as origins of migration flows. The negative coefficient for regional centers and other towns is especially important, because it shows that while Budapest was still the major hub, they were able to retain their populations more optimally. Administrative ranks for destination districts show a different picture: besides regional centers, they have a significant effect for all periods, after controlling for network effects. But for period 3, non-regional center towns show a positive effect, while they had a negative coefficient in periods 1 and 2. This shows that in period 3, they became a stronger receiver of migration in-flows. Budapest shows parameters both as a sender and a receiver that are in line with its hub-status, it receives and sends larger flows, even after controlling for network effects. Our results show, that even if descriptive analysis was unable to show strong reorganization in the flow structure across our periods, with ERGM tools, we can see that complex reorganization indeed happened, mainly in the middle of the urban hierarchy, on the level of regional centers and other towns.

To be able to understand changes in the flow structure between period 2 and 3, we fit similar models, with past migration flows included as predictors (edge covariates). This way our models explain the change in flow structures between period 2 and 3. Origin house prices have a significant and negative effect after controlling for network effects. This means, that expensive districts retained migrants in period three over the amount they retained in period 2. Destination house prices also have a significant, but positive parameter, so expensive districts attracted more flow volume, than they did in period 2. Economy and amenity indexes show a mixed picture. After including network effects, origin economy index remains positive and significant, but origin amenity index is non-significant even before network controls. Destination economy index is non-significant after including network mechanisms, but amenity index becomes significant and negative.

Table 5: ERGM results with past flows as controls				
	Period 3			
	Without network effects		With network effects	
Variable	Estimate	Sig	Estimate	Sig
Sum	-0.766		2.939	***
Non zero	-12.601	***	-10.974	***
CMP (dispersion)	-6.026	***	-4.881	***
Distance	-0.736	***	-0.539	***
Past flows	0.407	***	0.362	***
O: Capital	0.511	***	0.354	**
O: Regional centers	-0.125	**	-0.101	*
O: Other towns	-0.121	**	-0.077	
D: Capital	0.719	***	1.24	***
D: Regional centers	-0.029		0.084	*
D: Other towns	0.027		0.165	***
Same region	0.4	***	0.326	***
O: House price	-0.248	***	-0.368	***
O: Population	0.661	***	0.509	***
O: Economy index	0.14	***	0.052	*
O: Amenity index	0.012		0.029	
D: House price	0.17	***	0.303	***
D: Population	0.535	***	0.072	*
D: Economy index	0.144	***	-0.012	
D: Amenity index	-0.035		-0.036	*
Mutuality			0.943	***
Receiver activity variance			0.797	***
Sender activity variance			0.231	**
Significance: *** p<0.001, ** p<0.01, * p<0.05				
O: origin variables, D: destination variables				

Regional clustering has a positive effect after controlling for past flows, showing that in period 3, when general internal migration behaviour enhanced in the country, it happened inside regions to a higher extent. The urban hierarchy levels show the same effects as previous models. After controlling for network effects, origin capital has a significant and positive effect, showing that Budapest as a sender became more influential. But regional centers and other towns have a negative effect, showing that compared to period 2 flow structure, they were able to send lower amount of migrants through flows originating from them. However, the effect of other towns is not significant after controlling for network effects. As destinations, Budapest, regional centers, and other towns have significant effects, but only after controlling

for network mechanisms, showing that they attracted larger flow volumes than in period 2.

Network mechanisms show a very coherent picture across all models: with and without past flows, in all periods. Mutuality is always significant and positive, showing evidence for the fundamental internal migration mechanism, that flows tend to have their counter-flows. Receiver and sender activity variance show a more complex picture. Receiver activity variance is positive and significant in all models, showing that there are influential attractor districts in the Hungarian migration system, with higher activity in receiving high flows than it could be predicted by economic or amenity factors, house prices, or the levels of the urban hierarchy. Sender activity variance however is non-significant in models without past flows for periods 1 and 3, and it has a negative and significant parameter only in period 2.

4. Discussion

Our results demonstrate that omitting dependence structure biases parameter estimates in migration flow models. While magnitude of parameter estimates are not directly comparable before and after including network effects, we can see that after including them, some variables loose and some gain significance. This shows that individual unit and individual dyad approaches can lead to misleading model estimates. The application of the ERGM framework shows that the original formulation of push-pull mechanisms by Lee 1966 that embeds the push-pull calculus in the pre-existing flow structure can be included in contemporary statistical modeling of migration flows. Our network mechanisms were significant in all models, showing for example that mutuality, an internal migration effect documented as early as Ravenstein 1885, is indeed in effect. Even for future research not applying systematic network modeling frameworks, a simple mutuality term could be included in some form.

The pattern of which variable lost or gained significance in different periods is mixed however. Our amenity and economic indexes do not tend to lose their effect for origins, but for multiple periods they lose it for destinations. This shows that economic opportunities and amenity level can have an asymmetric effect. House prices have a robust effect, they do not lose their significance for any period, either as a destination or an origin effect. Positive destination house price effects persist

even with network controls, suggesting prices proxy opportunity rather than acting as barriers in Hungary's context, or that Budapest hasn't reached the affordability threshold that would deter migration. This contradicts results for the USA (Plantinga et al. 2013, Olney 2024), or Australia (Siriban et al. 2025), or even Germany (Stawarz et al. 2021). We are not certain about the source of this deviation from previous results, but one possible explanation is the unique urban hierarchy in Hungary, the internationally rare dominance of the capital. The USA, Australia, or Germany are all countries where the urban hierarchy is much more balanced, thus destination opportunities are also more evenly distributed. In a country where destination opportunities largely cluster in one central location, the barrier effect of prices might not work. This might also apply to origin price effects. Our negative origin coefficients suggest that in Hungary, expensive origins retain populations – rising prices function as an anchor rather than a push factor. This finding connects to Cavalleri et al. 2021, who showed that in some OECD countries, high origin prices operate as a retention factor.

Our finding about the significance of urban hierarchy rank categories show that urban hierarchy effects differentiate only after controlling for network mechanisms, suggesting hierarchy operates through but is not reducible to network position. This finding generally underlines polycentric urban development literature, according to which a shift from central place to network-based hierarchies happened in developed economies (Batten 1995, Neal 2011). Our results also connect to important empirical results. Applying descriptive analysis of the Hungarian migration system Bálint and Obádovics 2021 finds that after 2016, the capital and large county seats started to loose population. Our parameter estimates for period 3 origin effects for regional centers and other town is significant and negative after controlling for network effects, showing that the descriptive population loss can be attributed to internal flow network mechanisms to some extent.

One methodological criticism against interpreting our urban hierarchy finding could be that these variables can be seen as proxies for positions in the flow structure, and this way, whether the model finds network variables or urban hierarchy rank variables significant is merely random. This is not the case, for three reasons. First, based on both AIC and BIC, our models improved after including the network variables, so information-wise they were not redundant (model fit statistics for all

specifications are reported in Appendix A.1., Table A.1.1.). Second, if multicollinearity between urban hierarchy and network terms produced spurious significance patterns, we would expect erratic variation in coefficient signs and significance across periods and specifications. Instead, destination effects show complete stability across all periods (significant in all six model specifications), while origin effects display stability in Periods 1-2 followed by a coherent regime shift in Period 3 – precisely when our descriptive analysis identifies structural network reorganization. This systematic pattern is inconsistent with random collinearity artifacts. Third, such multicollinearity would mean model degeneracy in ERGMs, which we didn't find: even at low, experimental MCMC control parameters, our models converged fast and efficient, with the same result as with high parameters. In appendix section A.1. we also provide detailed ERGM-diagnostics showing that our results cannot be interpreted as a result of model misspecification or model degeneracy.

5. Summary

This study examines how internal migration flow networks respond to house price shocks in dominant central city. Internal migration research shows that house prices affect migration propensities and directions. Rising destination prices generally deter in-migration (Plantinga et al. 2013; Stawarz et al. 2021; Olney et al. 2024), while origin price effects vary across countries, operating as either push or retention factors (Cavalleri et al. 2021). However, most studies rely on gravity models treating flow dyads as independent, omitting the systematic nature of migration networks that classical theorists emphasized (Ravenstein 1885, Lee 1966). Recent inter-urban network literature addresses this gap, modeling flow characteristics with network methods (Wang et al. 2019, Zhang et al. 2020, Zhu et al. 2021). Yet we lack evidence on how entire flow systems respond when dominant hub becomes unaffordable.

Hungary during the 2010s is a proper case for examining this question. Budapest's metropolitan area is more than nine times larger than the country's second city, this hub-centered system experienced severe housing price shock: between 2014 and 2024, Hungarian nominal house prices grew by over 300%, the highest growth in the EU, with the boom originating in Budapest. Our analysis covers nine years between 2010 and 2018, divided into three periods with distinctly different price dynamics. We apply exponential random graph models for count-valued networks (Krivitsky

2012), which allow estimation of flow determinants while explicitly modeling network mechanisms like reciprocity and heterogeneous node activity. This approach enables us to estimate effect of house prices and other contextual factors controlling for the interdependent nature of migration flows.

Descriptive findings show network expansion rather than structural reorganization. Edge counts grew 27% and migration volume increased 38% across the study period, while Budapest's share of flows declined only modestly. Triangles with Budapest as middleman multiplied, without increasing direct bypass flows between origin and destination nodes. Share of direct flows in these triangles remained below 10% even in countryside configurations. These patterns suggest that the migration system added connections between countryside districts while maintaining its hub-dominated structure. Despite 162% deterioration in capital's affordability ratio, Budapest retained its central position in the flow network.

ERGM results reveal differentiation invisible in descriptive statistics. After controlling for network mechanisms, urban hierarchy effects showed significant changes across periods. In period 3, regional centers and other towns showed significant negative effects as origins, indicating enhanced population retention compared to earlier periods. House price effects were counter-intuitive in all periods: origin house prices had negative parameters (expensive districts retained population), while destination house prices had positive effects (expensive districts attracted migrants), even after controlling for economic and amenity factors. Network mechanisms showed consistent patterns, mutuality was always significant and positive, confirming that flows tend to have their counter-flows, and receiver activity variance was significant, showing influential attractor districts in Hungarian migration system.

Results show that internal migration networks do not respond to central city price shocks through rapid structural reorganization. Hub-centered patterns persist despite substantial affordability deterioration. The network expanded with more edges, more migration volume, more triangular configurations, but Budapest maintained its dominant position. The differentiation we observe happened in the middle of urban hierarchy, on the level of regional centers and smaller towns, which showed enhanced population retention and attraction effects in the period of highest price

growth. This suggests that while housing prices matter for migration propensities and directions, architecture of hub-dominated migration systems is resistant to short-term price shocks, at least on the timescale of our analysis. The interdependence of flows might buffer the system against rapid change even when affordability conditions shift dramatically.

Appendix

A.1. Model diagnostics

This appendix documents the estimation procedure, convergence behaviour, degeneracy diagnostics, and goodness-of-fit assessment for all count-ERGM specifications reported in Article 2 (Tables 4 and 5). The diagnostics follow recommendations for explicit reporting of degeneracy diagnostics and MCMC convergence procedures (Desmarais and Cranmer, 2012).

A.1.1. Estimation protocol

All models were estimated by Monte Carlo maximum likelihood (MCMLE) using the `ergm` and `ergm.count` packages (Krivitsky, 2012) with a Poisson reference measure and the transformed edge weights described in the main text. Estimation used four parallel MCMC chains (PSOCK cluster). MCMC control parameters were set at the package defaults: burn-in of 16,384 proposals, sampling interval of 1,024 proposals, and a nominal sample size of 1,024 networks per chain, with a ceiling of 75 MCMLE iterations. The MCMLE stopping rule certifies convergence when the estimating equations are statistically indistinguishable from zero at 99% confidence (Hummel et al., 2012); where the certification test lacks Monte Carlo precision, the algorithm adaptively increases the sample size rather than terminating early, so the final-round samples on which the reported estimates rest meet or exceed the nominal per-chain sample size.

Models without network terms are dyad-independent specifications; under the Poisson reference these reduce to Poisson-regression models for which MCMC degeneracy concerns are not valid. Full graphical diagnostics are therefore presented for the specifications containing dependence terms.

A.1.2 Convergence behaviour

Every model converged with 99% confidence long before the iteration ceiling (Table A1, column "MCMLE iter."). Estimation for the network term models followed the profile characteristic of a well-behaved likelihood surface: an initial phase of dampened step lengths while the sampler approached the region of the observed

sufficient statistics, followed by a sustained sequence of full-length steps (step length = 1.000) with monotonically shrinking log-likelihood improvements until certification. At no point in any estimation did step lengths collapse toward zero, oscillate, or fail to recover — the signatures of near-degeneracy described by Hummel et al. 2012. Final-iteration step length was 1.000 for all ten models.

The parameter estimates themselves proved robust to the estimation configuration: models re-estimated under different MCMC control parameters - including during early experimentation runs with considerably shorter chains - reproduced the reported coefficients to the second or third decimal, with all signs and significance levels preserved. The invariance of the estimates across independent MCMC trajectories and chain-length regimes rules out the possibility that the reported results are artifacts of sampling error or chain configuration.

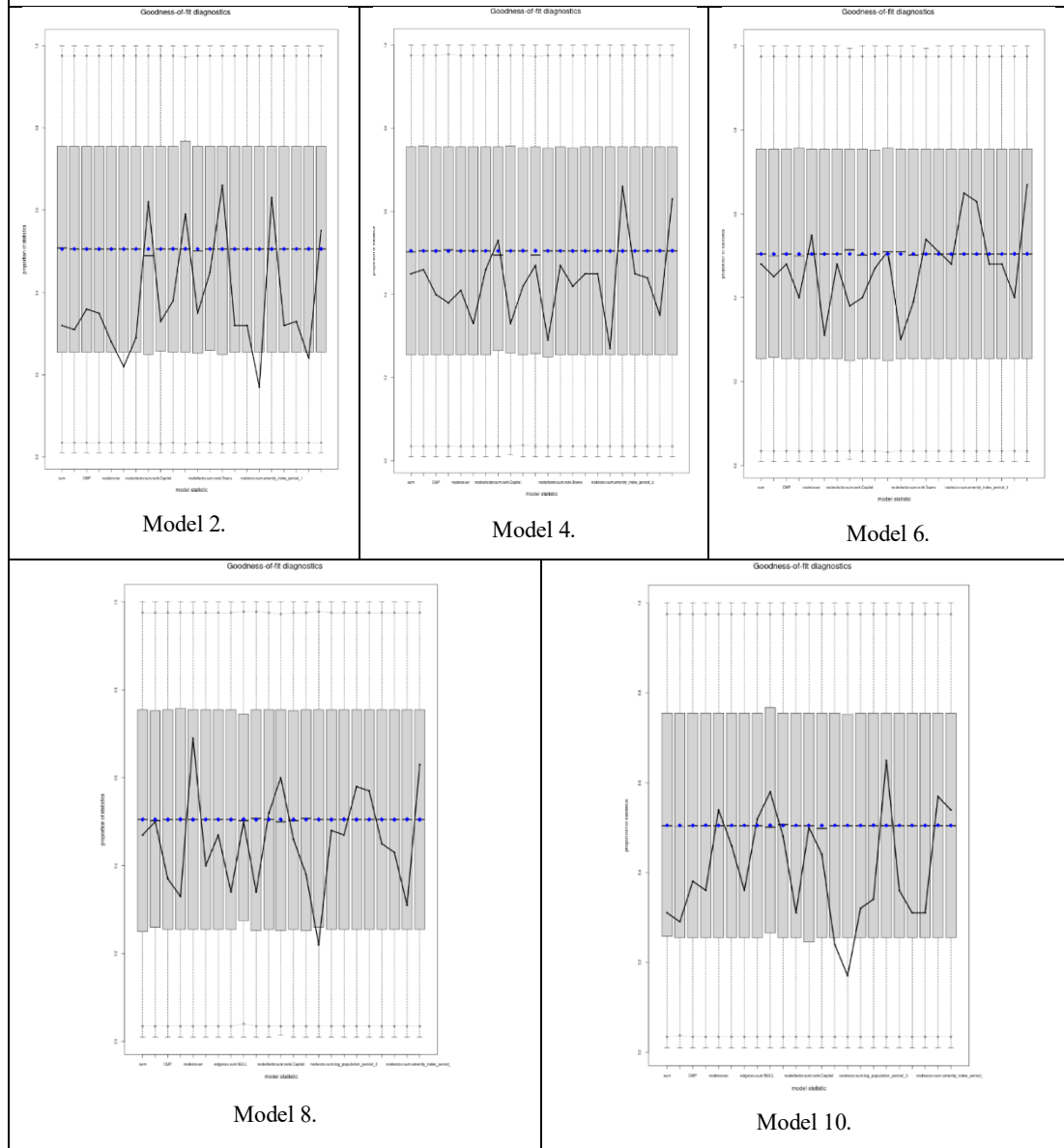
Table A.1.1. count-ERGM model fit statistics							
Model	Period	Specification	MCMC LE iter.	AIC	BIC	GOF min p	GOF stats with p <0.05
1	1	Levels, without network effects	22	-61,531	-61,372	0.64	0 of 19
2	1	Levels, with network effects	47	-63,651	-63,468	0.22	0 of 22
3	2	Levels, without network effects	22	-61,848	-61,690	0.66	0 of 19
4	2	Levels, with network effects	43	-63,728	-63,545	0.62	0 of 22
5	3	Levels, without network effects	18	-62,739	-62,581	0.42	0 of 19
6	3	Levels, with network effects	38	-64,649	-64,466	0.34	0 of 22
7	2 (change)	Past flows, without network effects	22	-65,439	-65,272	0.34	0 of 20
8	2 (change)	Past flows, with network effects	33	-66,669	-66,478	0.58	0 of 23
9	3 (change)	Past flows, without network effects	25	-66,682	-66,516	0.54	0 of 20
10	3 (change)	Past flows, with network effects	30	-67,898	-67,706	0.62	0 of 23
Period: the period the given model is fitted on.							

A.1.3. Degeneracy assessment: goodness-of-fit at final estimates

The most direct test of model degeneracy is whether networks simulated from the fitted model reproduce the observed sufficient statistics: a degenerate specification places its probability mass on empty or complete graphs and cannot reproduce the observed statistics by construction (Handcock, 2003). For each model, networks were simulated at the final parameter estimates and the observed value of every model statistic was compared against its simulated distribution.

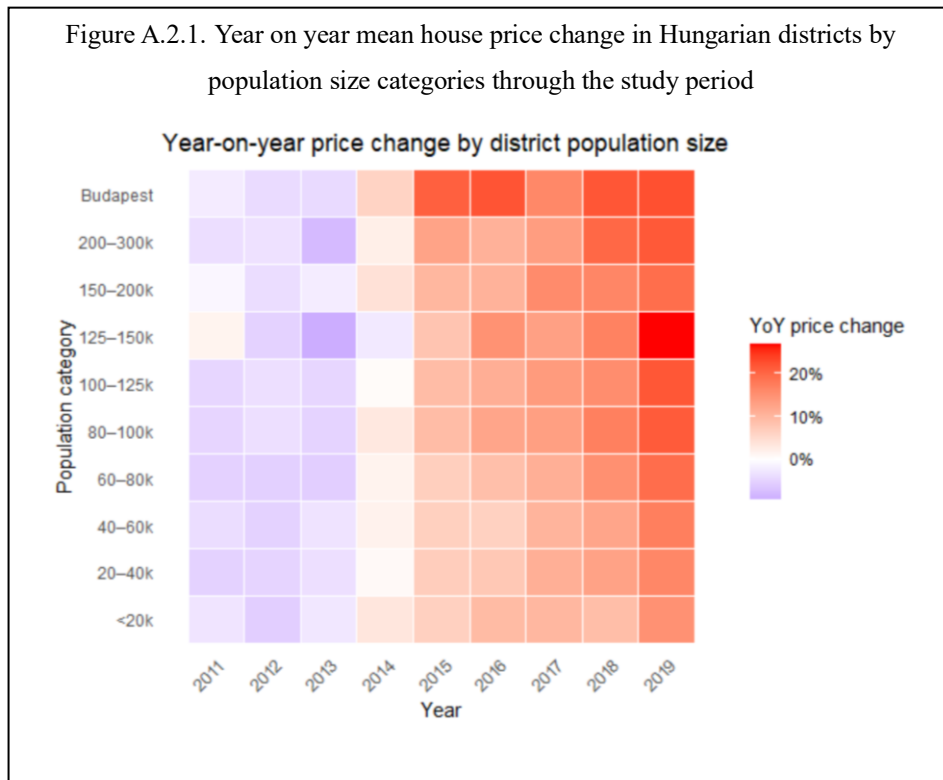
The results are uniform across all ten specifications: not a single model statistic, in any model, falls outside its simulated distribution at the 5% level (Table A1, final column). Minimum Monte Carlo p-values range from 0.22 to 0.66 across models, and observed statistics sit near the centers of their simulated distributions throughout. Figure A1. displays the goodness-of-fit plots for the five network-type specifications (Models 2, 4, 6, 8, 10); observed statistics lie at or near the medians of the simulated distributions for all statistics, including the network terms where degeneracy concerns concentrate. No specification exhibits degeneracy.

Figure A.1.1. Distributions of model statistics across networks simulated from the fitted models



Grey boxes: distributions of each model statistic across networks simulated from the fitted model at its final estimates; blue diamonds: observed values in the empirical network; black line: quantile of the observed value within each simulated distribution (values pinned at 0 or 1 would indicate non-reproduced statistics; mid-range variation is the expected pattern under a correctly specified model).

A.2. Structure of the house price boom in Hungary



Our identification of the house price shock rests on a simple premise: the boom of the 2010s was a single, top-down price shock originating in Budapest, and not a mosaic of independent local booms. If price dynamics had emerged independently across districts - driven by local policies, local labour markets, or local housing supply - then the price effects estimated in our models could be confounded by those local factors. If, however, the boom was one shock propagating outward from the capital, no such local confounding structure exists, and the price variables measure positions in a single national price wave.

Figure A.2.1. presents year-on-year house price changes between 2011 and 2019 for districts grouped into population size categories, from Budapest at the top to districts below 20 thousand inhabitants at the bottom. Population size categories serve here as a finer-grained version of the four-level urban hierarchy classification used in our models. Similarly, while our analysis compares three periods of distinct price dynamics, the figure resolves these periods into nine yearly transitions (with 2019 as an extra year included to show the full scale of price dynamics).

The pattern the figure shows is a wedge. In the first third of the observed window price changes are negative or near zero across every category: the post-crisis bust was general, and no part of the settlement system escaped it. The turn comes from the top. Budapest is the first to shift into strong positive growth, in 2014, while the categories below it remain near 0% change for another year or two. Growth then appears in the larger town categories, and reaches the smallest settlement categories - predominantly located in the peripheries of the country - last. Remarkably, by the end of the window the strongest single year of appreciation belongs not to Budapest but to the middle of the size distribution: the catch-up phase, in which the crest of the price wave has already passed the capital and moved down the hierarchy. This leader–follower sequence is precisely what hierarchical diffusion looks like, and what a collection of independent local shocks would not look like.

We note that the diffusion was hierarchical rather than strictly geographical. Regional centers such as Debrecen followed Budapest early in the boom despite their considerable physical distance from the capital, while many settlements geographically closer to Budapest, but lower in the settlement hierarchy, turned later. The price wave thus travelled through the urban size distribution first, and spread into hinterlands second — which is why we describe the spatial structure of the boom through the hierarchy, and why our models operationalize position in the boom through hierarchy rank and prices rather than through distance from the capital alone.

Two limits of this exercise should be stated. The figure is descriptive, as it documents the spatial-temporal structure of the boom, and does not by itself identify the causal channel of propagation. And the figure extends to 2019 for context, one year beyond our analysis window, to show the catch-up phase completing; the models themselves use only the 2010–2018 data. Neither limit affects the conclusion this appendix exists to support: the boom entered the Hungarian settlement system at a single point, its dominant hub, and propagated downward through the urban hierarchy. The price variation our models exploit is therefore variation within one national shock.

3rd ARTICLE:

SOCIAL CONNECTIONS AND LINKS TO
COMMUNITIES DRIVE INTRA-
METROPOLITAN MIGRATION WHEN
HOUSING IS SCARCE

Social Connections and Links to Communities Drive Intra-Metropolitan Migration when Housing is Scarce

Gergely Mónus

Abstract:

Social network theories complement economic models in explaining migration, yet social connections' role is usually analysed as an unconditional effect rather than in interaction with the societal and economic context shaping mobility. We examine how social connections drive migration between settlements in the agglomeration area of a large European capital, testing whether network effects vary with housing availability. Combining online social network data with administrative migration records enables settlement level analysis of network topology and subsequent flows. Beyond confirming that connections shape migration direction, we find that the density of friend cliques in destinations also has migration facilitating effect. Critically, both total connections and density of friend groups show a substitution effect: their migration-facilitating power intensifies when destination housing is scarce, revealing how social ties function as informal infrastructure compensating for inadequate housing supply.

1. Introduction

Internal migration is the residential relocation of people within a country. Traditionally, the internal migration processes are attributed to the attraction force of economic opportunity – the fact that large centers of commerce and industry pull populations from other locations (Ravenstein 1885). Push-pull theories of migration articulate internal migration dynamics as an interplay between attracting forces in destinations and pushing factors in the area of origin (Lee 1966). In this traditional model, different factors are thought to drive long and short distance internal migration events: long distance moves are defined by the pulling forces of better work opportunities, while short distance moves – residential mobility – are defined by better housing opportunities (Niedomysl 2011). Classic push-pull models, however, tend to assume economic centralization, and thus do not apply to all historical and geographical contexts. In post-industrial societies, for example, where economic centralization has already occurred, the predominant geographical mobility

is the "*movement of migrants from city to city and within individual urban agglomerations, especially within a highly elaborated lattice of major and minor metropolises*" (Zelinsky 1971). The explanatory power of classic push-pull models for such dynamics is limited: while economic opportunity undoubtedly plays a large role in migration decisions, these models are unable to explain the patterned, geographically clustered morphology of migration that links particular origins to particular destinations (de Haas 2010). To account for this structure, the idea that social networks play a key role in defining migration patterns emerged as a complementary explanation to economic models.

Social networks can provide specific content to idiosyncratic preferences and help explain why similar people make different choices (Büchel et al 2020). Migrant networks are sets of interpersonal ties that link migrants and former migrants, and people in the origin and destination communities. Migrant networks include ties of kinship, friendship, and shared origin. Their function has long been recognized in structuring migration decision making. Each act of migration creates social capital for others to whom the migrant is related, reducing the costs and risks of movement. As the number of network connections in an origin area reaches a critical threshold, migration becomes self-perpetuating because each act of migration creates the social structure to sustain additional movement (Massey 1990, Massey et al. 1993). While migrant network models and theories are usually applied for international migration, the idea of information sharing between previous and former migrants in rural-urban migration is foundational for systematic models of internal migration also (Mabogunje 1970).

The original formation of social capital of Bourdieu defines it as the aggregate of actual and potential resources that are linked to the possession of a set of connections, in other word, a network. This way, Bourdieu differentiates between the connections and the means that can be accessed through the connections - this distinctions mostly disappeared in later works on social capital (de Haas 2010). Research on the effect of migration networks also usually only consider the number of social connections in origins and destinations, and thus conflates the many different aspects of how social capital can effect a migration process. Jackson differentiates between seven forms of social capital: information, brokerage, coordination, bridging, support, reputation, and community. In internal migration,

two of these can be of greater importance, information and support capital. The former refers to the ability to acquire or spread valuable information, and the latter to the ability of exchanging favours through network position (Jackson 2020). Social capital also has to be differentiated between relative locations: local social capital refers to a households resources that can be utilized and arise from social connections with people living in physical proximity (Kan 2007).

The evidence on the mobility reducing effect of local contacts, and the mobility enhancing effect of contacts in destinations is substantial across different measurement strategies. Kan 2007 found for the USA that the availability of nearby emergency assistance significantly deters household residential mobility, with this deterrent effect being stronger for long-distance moves than short-distance moves. Also, in the USA, family ties explain a high share of variance in racial interstate migration patterns, because they have a stronger negative effect on migration for Black Americans than White Americans (Spilimbergo and Ubeda 2004). For Great Britain, Ermisch and Mulder 2019 finds that living far from parents increases longer distance mobility, while seeing parents weekly and more interactions with neighbours reduce longer distance mobility. Similarly, living close to parents significantly reduces the likelihood of internal migration in Germany, and this relationship is partially mediated by instrumental support exchanges between generations (Hünteler and Mulder 2020). In China, household expenses on wedding gifts significantly reduce migration probability in rural areas (Zhao and Yao 2017). In Australia, both homeownership and frequency and level of neighbourhood cooperation significantly reduce migration probability, with these factors increasing model explanatory power by almost a third beyond standard demographic controls (Clark and Lisowski 2019). It is important to note, that based on the literature, the effect of social connections is only complementary to classical demographic and economic factors. Comparing the weight of these effects in the migration process is still a question of emerging new research. But for example Michaelides 2011 finds for the USA that workers with strong local social networks (locally born or married to locally born spouses) are significantly less likely to migrate between metropolitan areas, with effects substantially larger than wage or housing cost differentials.

Literature summarized above mostly test the effect of the sheer number of social connections in origin and destination locations. While this is a proper approach to

operationalize social capital defined in total connections, heterogeneities across different social network structures remain hidden. Social network theory fundamentally differentiates between strong and weak ties (Granovetter 1973). Family ties, workplace ties, and friendships can all have different effects on migration decision. New evidence on the effect of different social network configurations on internal migration started to emerge during recent years. For large-scale phone contact networks, there internal migration in multiple countries shows associations with social networks. Büchel et al. 2020 finds for Switzerland that a higher share of phone contacts in a district makes choosing that district as destination of migration more likely, and that share of direct and second order phone contacts in 10 min proximity reduces the probability of moving. Using algorithmic classification of relationship types show that friends matter more for younger ages, and family ties are more influential for the older ages. Similarly, Blumenstock et al. 2025 finds evidence for Rwanda that phone contacts in destination areas makes choosing given destination more likely, and contacts in the origin location reduce the probability of migration. Also, Rwandan migrants are more likely to move to destinations where their social networks are interconnected (friends know each other) rather than extensive (friends have many other friends). This indicates that support capital might be more important for Rwandan internal migrants than information capital.

Big data sources, like mobile phone networks or online social networks have their own limitations too. Most importantly, they are anonymized, so if a migration event cannot be explicitly tracked, then it is impossible to connect individuals to other sources of migration records. One way to overcome this limitation is to use aggregate level social connection measures between locations, between which migration is observable from other data sources. Aggregate-level studies for the effect of social capital on migration are more rare than individual, since it is not straightforward to aggregate individual connections to the level of a geographical unit. Traditional survey tools have multiple proxies for social capital, like social support, trust, cultural activities, social contacts. For example in Finland, municipalities with higher levels of social capital measured through such tools have higher net migration rates (Jokela et al. 2025). Social capital measured through associational density, civic participation, and nonprofit organizational presence is effected positively by immigration and negatively by out-migration in US counties (LeSage and Ha 2012).

The aggregated effect of social networks on internal migration has not only been measured for singular countries, but in a comparative manner also. David et al. 2010 utilizes the European Community Household Panel, and show that local social capital (strong ties to friends/neighbours) reduces geographical mobility while general social capital (club membership) may facilitate it, and this relationship helps explain the North-South divide in European mobility rates. These examples all share the same design, where the effect of aggregated local social capital is tested on local migration rates. This approach omits the direction of migration flows – what is the effect of social networks on the exact flow structure connecting locations? As Massey writes, on the individual level, while migrant networks are a form of location specific social capital, their special quality is that they give access to resources in other locations also, and this is why they reduce different costs of transitioning between the origin and destination locations (Massey 1990). Accounting for flow structure can reveal how social connections move migrants between locations, and whether such moving effects depend on local characteristics. An important example for our study is Lőrincz and Németh 2022. They presents analyses for Hungary that uses online social network data to measure three types of social capital on the level of communities: bridging (between them), local (within them), and bonding (locally closed structures). They analyse how these three types of social capital affect migration between Hungarian communities, and only find significant effect for bridging social capital, which is positively associated with migration flow.

In most studies, the effect of social connections on internal migration and residential mobility is observed as an independent main effect. However, it would be unreasonable to think that social networks matter for everyone to the same extent and through the same mechanism. It is a widely observed social effect that when market income and institutional support fall short, social networks have an enhanced role in supporting households, working as a safety net and a “*third source of welfare*” especially for low income families (Hill et al. 2021). Similarly, there is evidence that the effect of social networks matter more in internal migration and residential mobility for those who suffer some form of scarcity. For example in the USA, local kinship ties and children's social networks significantly deter inter-neighbourhood residential mobility of families with children, with this effect being even stronger for low-income families (Dawkins 2006). An other example is Italy, where provincial

social capital significantly reduces migration, and this effect is stronger for those with basic education who rely more on family and local networks (D'Ingiullo et al. 2023). We theorize that the effect of social connections on moves has a special effect in scarcity of requirements of mobility, especially housing. To test this interaction, we need a case context with significant internal mobility but imperfect housing markets and housing supply. Such context can be an area where internal migration is attracted towards places without significant pre-emptive housing and amenity development.

The Budapest agglomeration area (BAA) offers an ideal context to test the interaction between social networks and housing scarcity in migration. The BAA represents a paradigmatic case of post-socialist Central and Eastern European urban transformation. A common pattern in CEE urbanization processes after the regime change wave is that because of the economic costs of the transition, the influence of central planning and development decreased significantly (Stanilov and Sýkora 2014). In post-socialist Central and Eastern Europe, urban transformation occurred through private- and market-driven rather than state-planned processes, with urban transformations often left to the operation of the free market, while *"ad hoc decisions have been preferred over the creation of policies embedded in a long-term plans"*, creating a system where housing development followed rather than preceded population settlement patterns (Sýkora & Bouzarovski, 2012). Post-transition development is characterized as largely unplanned and driven by local interests rather than coordinated metropolitan planning (Kovács et al 2019). Many of the former peripheral villages surrounding CEE metropolitan areas served as seeds of post-socialist suburbanization, suggesting existing population settlements became the foundation for suburban expansion rather than new planned developments (Stanilov 2007). In the BAA, the previously homogeneous structure of the metropolitan periphery experienced significant changes in the course of the transformation, as the advantages of location, which had been insignificant during the state-socialist period, once again became truly beneficial which resulted in growing disparities in the metropolitan periphery (Szabó et al. 2014). This transformation attracted substantial population inflows that continue to represent the predominant internal migration pattern in Hungary, though these flows were distributed unevenly across the region.

The reactive, market-driven nature of development in the BAA suggests that population flows were not directed by pre-existing infrastructure or planned amenities, but rather that development followed existing settlement patterns. For example, regarding Telki, one of the wealthiest settlements in the BAA, Csizmady et al. 2022 demonstrate that after the regime change, lacking sufficient financial resources, the municipality converted and sold agricultural land en masse to generate revenue for basic infrastructure development. This land sales strategy enabled an inflow of high-status newcomers from Budapest. The superior social networks and organizational capacity of these newcomers allowed them to seize municipal control, subsequently directing policy according to their preferences while displacing original residents. Telki is a good example for how the spontaneous character of development, with areas changing function from holiday home areas or agriculture to residential use, created new forms of spatial differentiation and social stratification. While social networks and residential decisions in the BAA have not been explicitly tested by quantitative research, previous research shows that suburbanization around Budapest is spatially segregated by socioeconomic status, with higher-status households clustering in attractive northwestern areas while lower-status households move to eastern zones – and this clustering to a large extent can be attributed to the spatial connections to districts of the capital with similar status (Kok and Kovács 1999). Also, there is qualitative evidence about how due to poor infrastructure and limited public services in peri-urban allotments (poor areas of the BAA that changed function from small scale agriculture to residence in a spontaneous way), the costs of social reproduction fall disproportionately on households, and the cooperative networks and reciprocal support systems that historically enabled collective resource-sharing and made self-provisioning viable (Gagyí and Vigvári 2025). These patterns suggest that social connections and social capital played a significant role in directing population flows within the BAA, making it an ideal case study for examining the interaction between social networks and housing market constraints in migration decisions.

Our study contributes to the understanding of the effect of social connections in internal migration several ways. We treat social connections at the level of settlements, and not on the level of individuals. This way, in our analysis social capital is the property of places, which defines the inter-dependent structure of them.

Such analysis is already available in the literature, but as far as we know, this study is the first that aims to differentiate different configurations of social connections at the location-pair level. Such experiments were only performed with individual-level ego networks yet. We also aim to show, that the effect of different configurations of social connections between locations is not uniform, other major residential migration defining factors have a significant interaction with it. Such interactions could be especially important for contexts where material factors of migration decisions do not work as a pre-emptive, attracting force. Intra-metropolitan migration, which is highly defined by housing opportunities can be such a context, if housing development only lags behind population redistribution. Our case study, the Budapest agglomeration area in recent decades serves us as a location where these circumstances can be observed with an enhanced power. We ask three different research questions:

Research Question 1: Are past social ties between settlements in the Budapest agglomeration area affecting the direction of future residential migration between them?

Research Question 2: Do past ties to densely connected local networks matter more for residential migration?

Research Question 3: Is the effect of past social ties on future migration enhanced in the case of housing-scarce destination settlements?

Our findings for the first question show that the effect of social connections in internal migration, which have been widely documented for country level, intra-regional migration processes can also be observed in the intra-metropolitan context. Results for the second question show that the density of online friend cliques in destination locations have a significant and positive effect on residential migration decisions, which complements the effect of all online connections between settlements. Results for the third question shows that housing supply has a significant interaction both with all connections, and the density of known friend groups. Answers to this question shows that the effect of social connections is heterogenous across scenarios where material push and pull effects have varying power. As Topa and Zenou 2015 summarizes the state of the literature on neighbourhood and network effect, *„there are very few empirical studies that explicitly test the interactions between the urban space and the social space and their impact on the outcomes of individuals”*. With the answers to our questions, we aim to demonstrate

that on the level of locations and migration streams, the interactions between housing and social networks is clearly observable.

2. Data and methods

2.2. Data sources

2.2.1. *Online social connections*

Analysing the role of social networks in the migration process presents a difficulty for data collection, since both connections and migration events have to be recorded simultaneously (Blumenstock et al. 2025). According to Lőrincz and Németh 2022, applying detailed administrative data sources for migration records and combining them with online social network measures presents a solution. Our social network data comes from iWiW, which was Hungary's major online social network in the pre-Facebook era, founded in 2002 and operating until it was shut down in 2014. At its peak around 2009, iWiW served more than 3 million active users, representing approximately 40% of Hungary's adult population – at that time 5 million out of 8.5 million adults used the internet in the country. The scale of the platform is evidenced by the more than 300 million friendship ties that users had established by the end of 2011. All data was anonymized by the service provider in 2013 for research purposes (Wachs et al. 2019). Online social network data sources always raise the question of representability, and iWiW's coverage was indeed selective by urbanization level, with smaller rural areas having much lower adoption rates than highly urbanized areas (Lőrincz and Németh, 2022). However, we believe this limitation is not influential in the case of the Budapest Agglomeration Area (BAA), since populations in even the more rural, village-like settlements within the agglomeration typically commute to the capital for work and cultural consumption, making technological adoption more likely. In one of the few available comparisons of an online social network, and a population register based network Menyhért et al. 2025 provides evidence from the Netherlands that the online social network they tested is strongly correlated to the population register based network, both in connectivity patterns and community structure. The question of coverage across different social strata and age groups is also important, since older individuals were less likely to use such platforms, but they are also less likely to migrate between settlements. Previous

research has demonstrated that geographical factors explain registration rates on iWiW, and that administrative and geographical boundaries constitute spatial borders on the platform, while connections between settlements reflect historical administrative boundaries and geographical barriers (Tóth et al. 2021). For our analysis, we aggregate network data from the decade 2000-2009, meaning we use all friendship connections that were initiated and accepted in that period. One important shortcoming is that our data only contains the settlements of users, and not exact addresses. This, for smaller settlements, does not make identification problematic, but in the case of a large settlement, like for example Budapest, it does. This is one practical reason why we consider only moves between agglomeration settlements, and not moves between the agglomeration and Budapest.

2.2.2. Administrative data sources

Internal migration is a form of geographic mobility without clear definitions. Traditionally, internal migration is differentiated from general residential mobility based on whether a move displaces people's daily activity spaces. Short distance moves that do not displace these activity spaces are considered to be residential mobility, and medium and longer distance moves that change these are considered internal migration (Roseman 1971, Nedomysl 2011). However, contemporary research shows that there is no clear-cut distinction between the two categories, and studies using administrative geographies, or self-reported reasons can only achieve fuzzy differentiation (Coulter et al. 2016). Our case, as an agglomeration zone of a large city, is special for several reasons. Most people tend to work in the capital, so moving in the agglomeration does not change their workplace and livelihood. But medium-distance moves can happen between different sectors of the agglomeration (even on the scale of 30-50 km), and even shorter moves to a neighbouring settlement can change everyday life of households fundamentally (commuting patterns, retail, etc). Our data sources only contain information about the settlements of residents, not exact addresses, thus we only analyse moves that cross administrative municipality borders. This is why we frame our analysis as a intra-metropolitan residential migration.

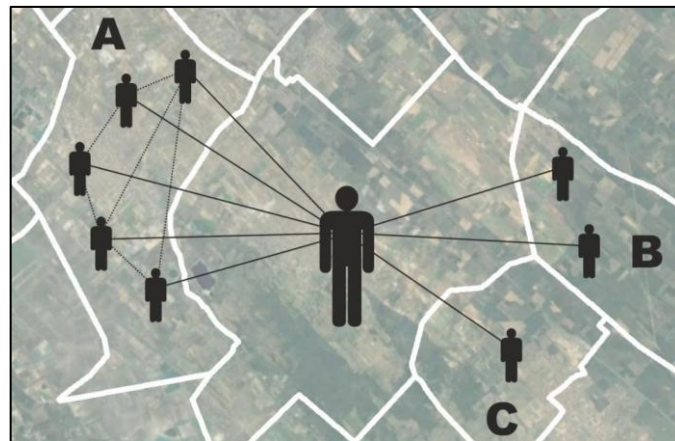
For migration data, we use residential migration records collected by the Hungarian Central Statistical Office, which capture every officially administered migration event in the country with origin and destination locations recorded at the municipality level (Mónus and Lőrincz, 2025). Following the approach established in previous research using this dataset, we construct our migration network from changes in permanent addresses only, excluding temporary addresses, as permanent address changes typically correspond to life cycle events such as marriage, establishing new households, or having children, while temporary addresses are predominantly used by students and young workers during short-term mobility. We aggregate migration flows over the period 2010-2019 to ensure sufficient statistical power for settlement-level analysis and to smooth out short-term fluctuations.

In our analysis, we also use administrative data on housing stock, and other urbanization factors. The source of this data is also the Central Statistical office of Hungary. Every variable is available for settlement – year pairs.

2.3. Methods

2.3.1. *Identification of connections to clique-like groups in destinations*

Figure 1. A: Schematic example for identifying connections to clique-like groups known by users of the online social networking site.



Online social network data suffers from significant limitations in migration analyses, most importantly that it is hard to assess the nature of given connections. Without information of names, family relations, school or workplace ties, the only possible solution for differentiating types of connections and social capital in the case of such

data is to analyze the topology of ego-networks. Blumenstock et al. 2025 computes the number of friends of friends (second-order social ties) as a proxy of information capital, and network support as proxy for support capital. Büchel et al. 2020 identifies pre-defined subgraphs in a user's network to measure family ties. Building on these examples, we measured strong connections of a user by computing the network statistics of their friend network in every agglomeration settlement that is not their home.

Figure 1. shows the example of a user, who has friends in 3 different settlements. In settlement C, the user has only one friend, and in settlement B, the user has two, but they do not know each other. But in settlement A, the user has 5 friends, and they form a clique-like network. We collapse the five connections into one connection, which this way is between a user and a settlement. We compute different network statistics for these connections: number of ties in the clique-like community, average degree, edge density, and global clustering. After these statistics are computed for all user-settlement pairs, we average them for every settlement pair. For example for edge density, and two settlements A and B, for our final measure we get the mean density of clique-like groups larger than 5 in B known by people in A, and vice versa. Correlations between computed network statistics are presented in the Appendix.

2.3.2. *Modeling migration flows with social connections*

Modeling migration with social connections presents fundamental endogeneity challenges, primarily through what Blumenstock et al. (2025) calls simultaneity, the fact that migration shapes social networks as much as networks shape migration. When aggregated to inter-settlement flow levels, this means settlements with histories of strong migration connections will have denser network connections. To address reverse causality, we use past connections to model future migration flow: social network measures are aggregated for 2000-2009, while migration events cover 2010-2019.

This decade-by-decade design offers several methodological advantages. First, iWiW's network evolution was non-linear, connection growth accelerated throughout the 2000s, with link deletion only emerging after 2010. Capturing the full decade

ensures we include both early and late adopters, maximizing demographic representativeness. Second, migration behavior requires sufficient temporal aggregation for stable statistical signal; our analyses indicate a minimum three-year window is necessary for reliable settlement-level origin-destination patterns.

While this design cannot eliminate reverse causality entirely, the Budapest agglomeration's evolving migration dynamics strengthen our identification. As documented in the results section, BAA internal migration represented an emerging form of intra-metropolitan mobility that grew substantially over our study period. This temporal growth pattern reduces the influence of pre-2000 migration on our network measures, making our case more suitable for isolating network effects.

We employ gravity-type regression models in log-log form, necessitated by the strongly skewed distributions of our variables, particularly migration flows. This specification allows parameter interpretation as elasticities – percent changes in migration corresponding to one-percent changes in explanatory variables. The gravity framework is fundamental given that population size and distance alone explain substantial migration variation, as we show in the results.

Sample selection requires careful consideration. Not all settlement pairs experience migration even when aggregating to the whole decade, and iWiW connections are even sparser. Most restrictively, directed settlement pairs where at least one person knows a five-person or larger community number approximately 4,500 out of 6,500 possible pairs. This is problematic for simple modeling reasons, but more importantly, it is conceptually problematic for our social network measures for the local communities. For example, network density can be zero (or around zero) substantially, if cliques have low amount of connections. If we would set absent settlement connections to zero in this instance, we would obscure the interpretation of our measures. Based on this, we choose to model only the settlement pairs where we can use our defined social connection measures. This gives a strongly limiting, but also meaningful framing to our results: they all refer to the effect of online connections and links to communities when those are present and measurable.

Our core gravity specification includes origin and destination populations and geographical distance between settlements. We build models incrementally: first adding social connection measures (total connections and mean clique density), then

housing variables (per capita housing in origin and destination), and finally urbanization controls. We compute the per capita level of housing and urbanization factors for every year during the decade after 2010, and then take the mean value as a predictor. We use three different principal components for urbanization factors based on the yearly per capita level of different variables: for economic factors (from variables local sum income tax, unemployment), schooling (number of nurseries and primary schools), and amenities (number of retail stores and restaurants-pubs) as control variables.

The interaction structure is critical to our theoretical framework. We include:

- Connections $\times$ Density: Tests whether connection quantity and quality reinforce each other
- Connections $\times$ Housing (origin and destination): Examines whether housing availability moderates connection effects
- Density $\times$ Housing (origin and destination): Tests whether housing scarcity amplifies the importance of tight-knit networks

Because of the log-log specification and interaction terms, main effects require careful interpretation. For instance, the connections coefficient in the full model represents the effect when density equals 1 (fully connected cliques) and housing per capita equals 1 in both locations – conditions not observed in our data. We therefore compute marginal effects at median values for substantive interpretation.

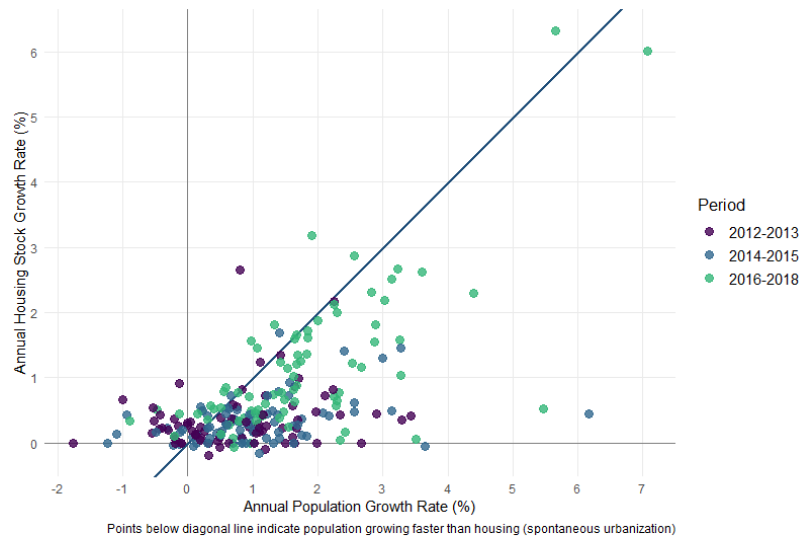
3. Results

3.1. Migration and housing development in the Budapest agglomeration area

Figure 2. shows growth rates of population and housing stock for settlements in the BAA, across our study period. We computed rates for three periods, not every year for visual clarity. This scatterplot provides evidence for the non-development-driven suburbanization in the BAA during the 2010s. The clear clustering of points below the diagonal reference line demonstrates that population growth consistently outpaced housing stock growth across all three periods, with settlements

experiencing population influxes before adequate housing infrastructure could be developed. Notably, the pattern intensifies over time: while the early post-crisis period (2012-2013, purple dots) shows modest growth rates clustered near the origin, the later periods reveal increasingly pronounced population growth, particularly during 2016-2018 (green dots) when several settlements experienced substantial population growth around 2-4% in that three years. This temporal progression illustrates how the Budapest agglomeration's suburbanization process was driven by population movements rather than planned development - people were moving to these settlements faster than the formal housing market could accommodate them, forcing at least some share of residents to rely on informal solutions like converted summer houses, extended family arrangements, or substandard housing. Even during the last period, when in many settlements housing construction ramped up, it was still slower than population growth. It is important to note however, that housing construction measured through statistical data sources can naturally lag to some extent, in situations where new residents build their own housing units, or generally when people inhabit dwellings without a put-to-use permit. While we cannot quantify the extent of such informal housing arrangements in our data, we consider this pattern of self-built or unpermitted housing as further evidence of spontaneous suburbanization processes.

Figure 2. Population growth vs housing growth by period across the 2010 decade.

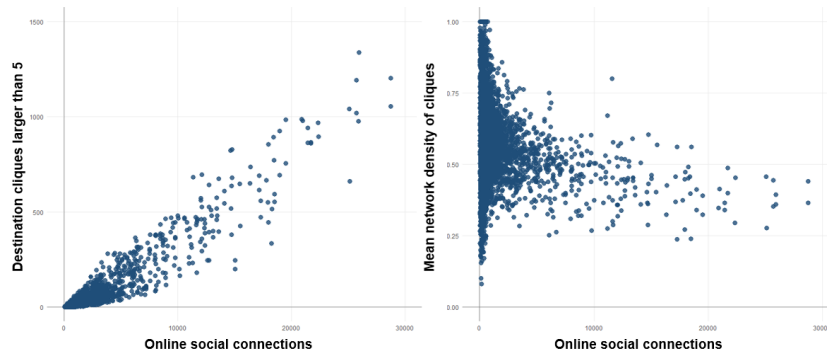


Every dot represent a settlement-period combination, different colours

3.2. Links to communities

One important test of our theoretical questions, and our statistical modeling is whether the structural properties of networks connecting settlements can be accounted simply to size, or whether there is heterogeneity between the topology of connections between settlements. The relationship between network quantity and quality in our data reveals both systematic measurement properties and meaningful structural variation that validates our analytical approach.

Figure 3. A: Number of online social connections between settlements vs number of cliques larger than 5 known in destination by origin users. B: Number of online social connections between settlements vs mean density of cliques larger than 5 in destination known by origin users.



In both graphs, each point represents a settlement pair in the Budapest agglomeration. To show patterns clearly, X axis on both graphs is limited to 30000 connections, omitting few pairs containing very large settlements.

Figure 3. A. demonstrates the expected linear relationship between total online connections and the number of clique-like communities we identify - more total connections naturally yield more identifiable cliques, reflecting the systematic extraction of meaningful network structures from the broader connection pool. However, Figure 3.B. reveals the crucial finding that justifies our analytical framework: substantial heterogeneity in network density across settlement pairs with similar connection counts. At very low number of connections, the density measure has extreme variability, spreading from minimal values to totally connected networks – this shows that our measurement can be noisy at low connection number. To avoid this, we would have to apply a more conservative minimal threshold, resulting in smaller population to analyse. However, at any given level of online connections - whether 5,000, 15,000, or 25,000 - settlement pairs exhibit dramatically different mean density values, with some showing tight-knit community structures (density around 0.7-0.8) while others display loose, sprawling networks (density around 0.3-0.4). This persistent heterogeneity across the full range of connection levels demonstrates that density captures genuine structural differences in how social networks are organized, not merely numerical artifacts of connection counts.

Correlation of clique network density and other networks statistics, especially size-driven ones, are presented in appendix section A.1. The fact that we can observe systematic differences in network density while controlling for network size provides the empirical foundation for our core finding that connection quality and quantity interact to shape migration flows in ways that neither measure could capture independently.

3.3. Regression modeling of migration flows

Table 1. Regression results for social connection measures

	Gravity	Connections	Mean Density	Interaction
Constant	-11.112*** (1.804)	-19.182*** (1.880)	-19.056*** (1.880)	-22.843*** (1.921)
Population origin	3.163*** (0.100)	1.528*** (0.162)	1.457*** (0.165)	1.218*** (0.166)
Population destination	2.932*** (0.101)	1.341*** (0.160)	1.389*** (0.162)	1.336*** (0.160)
Distance	-4.537*** (0.143)	-2.004*** (0.243)	-1.991*** (0.243)	-2.041*** (0.242)
Online connections		1.855*** (0.146)	1.911*** (0.148)	3.151*** (0.210)
Mean density			0.727* (0.358)	-8.967*** (1.225)
Connections x density				1.727*** (0.209)
R2	0.354	0.379	0.379	0.389
Adj. R2	0.354	0.378	0.379	0.388
Num. obs.	4145	4145	4145	4145
F statistic	757.634	630.749	505.803	439.772
RMSE	5.974	5.862	5.860	5.812

***p < 0.001; **p < 0.01; *p < 0.05; ·p < 0.1

Table 2. Regression results for social connection, housing, and urbanisation measures

	Housing	Housing x Conn	Full Interactions	Full model
Constant	-28.670*** (2.674)	1.451 (7.964)	-0.717 (8.019)	1.422 (7.966)
Population origin	1.242*** (0.167)	1.273*** (0.167)	1.291*** (0.167)	1.349*** (0.188)
Population destination	1.345*** (0.161)	1.347*** (0.161)	1.350*** (0.161)	1.422*** (0.181)
Distance	-2.047*** (0.241)	-1.987*** (0.241)	-1.979*** (0.241)	-1.543*** (0.243)
Online connections	3.135*** (0.210)	-1.962 (1.285)	-2.316· (1.342)	-2.450· (1.320)
Mean density	-8.842*** (1.227)	-8.491*** (1.230)	-14.935* (6.337)	-17.168** (6.249)
Conns x density	1.703*** (0.209)	1.636*** (0.210)	1.616*** (0.210)	1.656*** (0.208)
Housing from	-3.473** (1.306)	15.220** (5.466)	12.858* (5.536)	17.088** (5.492)
Housing to	-2.300· (1.343)	10.625· (5.778)	10.971· (5.791)	16.224** (5.765)
Conns x housing from		-3.087*** (0.880)	-3.695*** (0.910)	-3.808*** (0.898)
Conns x housing to		-2.102* (0.921)	-1.846· (0.978)	-2.151* (0.963)
Density x housing from			-10.330** (3.957)	-11.196** (3.894)
Density x housing to			3.614 (4.869)	2.486 (4.796)
Economic PCA from				0.032 (0.076)
Economic PCA to				0.037 (0.076)
Education PCA from				-0.079 (0.101)
Education PCA to				0.041 (0.101)
Amenities PCA from				-0.801*** (0.089)
Amenities PCA to				-0.748*** (0.089)
R2	0.391	0.393	0.394	0.416
Adj. R2	0.390	0.392	0.393	0.414
Num. obs.	4145	4145	4145	4145
F statistic	331.775	268.107	224.327	163.550
RMSE	5.807	5.796	5.792	5.691

***p < 0.001; **p < 0.01; *p < 0.05; ·p < 0.1

To be able to understand the effect of social connections on migration, we use gravity type regression model specifications. All the models are log-log form linear regression with a gravity specification. The log-log form is needed because almost all of our variables have a strongly skewed distribution, especially the response variable, migration. We include origin and destination populations, as well as simple geographical distance between origin and destination in all our models. We also compare model fit to the pure gravity model. We build our models gradually, first adding social connection measures, then housing, then urbanization controls.

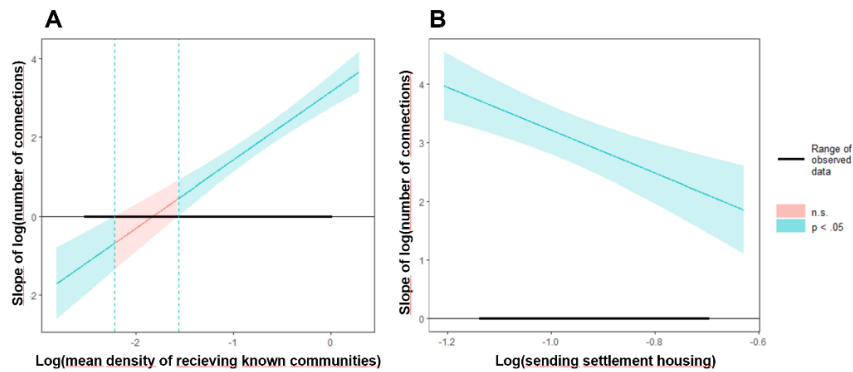
The pure gravity models shows expected results: population attracts migration, and distance decays it. Population effects (origin: 1.13-3.16; destination: 1.27-2.93) remain highly significant throughout all models, though their magnitude decreases substantially after controlling for social networks. Distance effect (-1.54 to -4.54) shows strong negative elasticity that diminishes when network variables are included. This confirms the basic gravity relationship but suggests social networks strongly mediate these traditional effects.

When included as the sole predictor besides gravity factors, the number of all online connections has a significant and positive effect: raising online connections by 1% corresponds to a 1.85% increase in migration. However, interpreting this variable's parameters becomes more nuanced in complex models due to the interaction terms. In the model with mean density and the interaction between connections and density, the main effect coefficient represents the scenario where mean density equals 1 (since $\log(1)=0$), meaning all cliques in the destination settlement known by people in the origin are fully connected. When interactions with both origin and destination housing are added, the main effect represents a scenario where density equals 1 and housing per capita in both origin and destination equals 1 – conditions not observed within the Budapest agglomeration. At the median values in our dataset (connections=498, density=0.58, housing from=0.372, housing to=0.372), the effect of connections is 2.30. This means that at typical values in the agglomeration, a 1% increase in connections is associated with a 2.3% increase in migration. After controlling for urbanization factors, the effect computed at median values becomes 2.55. These results show that online connections have a migration-inducing effect, but this effect varies substantially depending on network density and housing availability.

When included by itself in addition to gravity factors and all online connections, the mean density of communities in the destination known by people in the origin shows a moderate and significant effect: a 1% increase in density corresponds to approximately 0.73% increase in migration. Similar to connections, interpreting the density parameter requires consideration of interaction terms in more complex models. In the interaction model, the main effect coefficient for density represents the scenario where connections equal 1 (since $\log(1)=0$), indicating a minimal level of connectivity between settlements. When housing interactions are added, the main effect refers to a context where connections equal 1 and housing per capita in both origin and destination equals 1. At the median values in our dataset (connections=498, density=0.58, housing from=0.372, housing to=0.372), the effect of density is 1.77. This indicates that at typical values in the agglomeration, a 1% increase in network density is associated with a 1.77% increase in migration. After controlling for urbanization factors, the effect at median values remains stable at 1.76. Thus, network density has a substantial migration-enhancing effect at typical values, despite its negative main effect coefficient in complex models.

The interaction between connections and density provides robust and consistent finding across all model specifications. This interaction term is consistently positive and significant (1.62-1.73) even as additional variables and interactions are introduced. The positive coefficient indicates that the effects of connections and density mutually reinforce each other: increases in density enhance the migration-facilitating effects of connections, while more connections amplify the effects of density. This interaction remains significant even after controlling for housing and urbanization factors, suggesting the relationship between connection quantity and quality is fundamental to understanding migration patterns. The stability of this interaction coefficient across different model specifications further underlines its robustness as a predictor of migration flows.

Figure 4. Johnson-Neyman plots for interaction effects in different model specifications.



A: The slope for the log number of connections between settlements at different levels of mean density of known communities in receiving settlement. At low density, connections have no significant effect, and as density grows, the slope of connections also grows. B: The slope for the log number of connections at different levels of log per capita housing in the sending settlement. The interaction is negative, and significant across the whole range of observed housing level.

Housing variables show an intriguing pattern across our models. When first introduced without interactions, both origin and destination housing have negative effects on migration (-3.47 and -2.30 respectively), suggesting that higher amount of housing per capita reduces migration. However, when interactions with connection measures are included, the main effects become strongly positive (12.86 and 10.97), though these coefficients represent hypothetical scenarios where connections and density equal 1. At median values, the effect of origin housing is -4.54, while for destination housing -2.47. These negative effects at typical values indicate that better housing infrastructure reduces migration, with a stronger effect for origin housing. More importantly, the interactions between housing and network measures are consistently negative and significant. The negative interaction coefficients further reveal that housing and network measures function as substitutes: as housing availability decreases, the migration-facilitating effects of both connections and mean density of communities in destination become stronger.

The control variables for urbanization show varying levels of impact in our final models. Economic factors (represented by PCA components) show no significant effects on migration patterns after controlling for network and housing variables.

Educational factors show no significant effects for either origins or destinations, suggesting limited independent influence. In contrast, amenity factors emerge as strongly significant (-0.801 and -0.748) in the final model, indicating that settlements with more amenities experience less migration both as origins and destinations. This counterintuitive finding might reflect higher living costs or other constraints in amenity-rich areas. Importantly, our key network and housing interactions remain robust even after including these control variables, with only minor changes in coefficient magnitudes, further supporting the stability of our findings regarding social networks and spontaneous suburbanization.

The explained variance in our models ranges from 0.354 (pure gravity specification) to 0.416 (model controlling for all connection, housing, and urbanization effects). The R^2 of our model with only connection effects besides gravity variables is 0.379. This implies that while our social network and housing variables show significant, meaningful, and robust effects across multiple specifications, they provide incremental rather than transformative improvement in explanatory power after accounting for the gravity mechanism. This pattern is consistent with migration literature, where gravity variables typically explain a substantial portion of variation, and additional factors contribute incrementally to model fit. Nevertheless, the consistent significance and stability of our key interaction effects across specifications indicate that these relationships, though modest in overall explanatory contribution, represent genuine and important mechanisms in the migration process within the BAA.

The models presented above build on the idea that our specifications are able to differentiate the effect of social connections from other migration-inducing mechanisms, most importantly the simple gravity mechanism. However, a potential methodological concern that requires examination is that social connections themselves may be described by the gravity mechanism. Prior research confirms the intuition that social connections decay with distance, and larger populations attract more connections. This could create multicollinearity, and if our models cannot handle these correlations adequately, our estimates might be biased.

To investigate this, we examined the correlations between our social connection variables and gravity variables. In addition to origin and destination populations, we

calculated the sum of origin and destination populations. The correlation coefficients reveal that all online connections have a correlation with combined population, albeit not strong (0.333). Additionally, all online connections are negatively and moderately correlated with distance (-0.287). The mean density variable shows weaker correlations with combined population (-0.216) and distance (0.146). These patterns suggest that the structure of strong connection topology is less driven by population or - surprisingly - distance.

While these coefficients do not indicate strong relationships, modeling with online connections could potentially be misleading. To address this concern, we employed orthogonalization. Orthogonalization removes correlations between variables by using residuals from regressing predictors on control variables. We orthogonalized social connections against distance alone, and against all gravity variables, and used both new variables as predictors for migration. Remarkably, the parameter estimates for the original connections and the orthogonalized variables were identical up to three digits, and the gravity-orthogonalized variable also yielded very similar results. The R^2 values across models were also very similar. This demonstrates that our model specifications effectively differentiate the migration-inducing effects of different mechanisms from one another.

4. Discussion

Our results confirm that past social ties between settlements significantly affect future residential migration flows, extending network theory from international and inter-regional contexts to the intra-metropolitan scale. This finding validates the results of Massey et al. 1993 within agglomeration areas. The consistent positive effect of bridging connections demonstrates that social networks partially mediate traditional gravity effects of population and distance, addressing de Haas's critique that push-pull models cannot explain which specific origins connect to which destinations (de Haas 2010). Our approach responds directly to the gap identified in existing research: while most studies focus on individual-level effects or aggregate local social capital rates, we examine directed flows between location pairs, providing rare insight into how networks structure the morphology of metropolitan migration patterns.

The finding that clique density matters beyond sheer connection numbers represents a significant theoretical advance. Dense networks of friend communities in destinations facilitate migration independently of total connection counts, with effects remaining substantial even after controlling for housing and urbanization factors. This speaks directly to the distinction made by Jackson 2020 between information capital and support capital – if only connection quantity mattered, information transmission would be the primary mechanism. The density effect suggests support capital operates through tight-knit groups' ability to mobilize assistance. This parallels Blumenstock et al. 2025 about interconnected networks which facilitate migration more effectively than extensive but loose networks. The mechanism likely involves vouching systems, temporary accommodation, and information about scarce opportunities that is only available within cohesive groups. Our findings echo Lőrincz and Németh 2022 on Hungarian migration, but provide a more granular perspective – it's not simply bridging capital that matters, but the density of what you're bridging to.

The interaction between social connections and housing availability reveals the study's core theoretical contribution: networks matter conditionally, with their effects intensifying under resource scarcity. Connection quantity shows significant negative interaction with destination housing supply, indicating that social ties function as informal infrastructure when formal market mechanisms are inadequate. This connects to the "*third source of welfare*" literature (Hill et al. 2021), demonstrating how networks compensate when material factors are constrained. The substitution effect operates at the location level, similar to individual-level findings by Dawkins 2006 on kinship ties for low-income families and D'Ingiullo et al. 2023 on stronger social capital effects for those with basic education. The Budapest agglomeration context, where development followed rather than preceded settlement patterns (Sýkora and Bouzarovski, 2012), provides an ideal case for observing this housing-network interaction. This also connects to Kok and Kovács's 1999 documentation of status-based spatial clustering through social connections in the BAA, suggesting our quantitative findings have deeper qualitative foundations in how social reproduction occurs under infrastructural constraints.

5. Summary

This study examines how resource scarcity moderates the relationship between social network morphology and migration patterns, addressing a critical gap in understanding when and why network effects vary across internal residential migration contexts. While social networks are known to facilitate migration through information transmission and support provision, existing research has largely treated network effects as uniform rather than conditional on material constraints. Using social network data from iWiW (Hungary's pre-Facebook online platform) combined with administrative migration records for the Budapest agglomeration area from 2010-2019, we analyze how both the quantity and quality of social connections influence residential mobility when formal infrastructure is inadequate.

Our approach builds directly on recent methodological advances by Blumenstock et al. 2025 and Büchel et al. 2020, who pioneered the use of digital trace data to measure network effects on migration. However, while these studies operate at the individual level – examining how personal network characteristics affect individual migration decisions – we develop an aggregated measurement strategy that treats social capital as a property of location pairs. Rather than following the traditional social capital approach exemplified by Lőrincz and Németh 2022, which measures bridging, bonding, and local social capital as general community characteristics, we identify densely knit networks by computing clique density statistics for each origin-destination settlement pair. This involves calculating the mean density of friend groups larger than five people that users know in destination locations, capturing how tightly interconnected a migrant's potential support network would be at each possible destination. This represents a novel experiment for aggregated migration studies, which have typically relied on broad social capital measures rather than specific network topology indicators. By measuring the actual structural density of connections that link specific origins to specific destinations, we can test whether tight-knit community structures facilitate migration differently than loose, extensive networks – addressing Blumenstock et al.'s interconnectedness insight at the aggregate flow level rather than the individual choice level.

Our results demonstrate three key findings. First, social connections significantly direct migration flows, with a 1% increase in connections corresponding to a 1.85%

increase in migration. Second, the density of friend cliques in destinations matters beyond mere connection counts, with tight-knit networks facilitating migration independently of total connection volumes. This extends Blumenstock et al.'s interconnectedness finding to intra-metropolitan contexts, suggesting support capital operates alongside information transmission mechanisms. Third, and most significantly, social networks function as substitutes for formal housing infrastructure. Both connection quantity and clique density show stronger migration-facilitating effects when destination housing supply is constrained, revealing how informal social infrastructure compensates for inadequate market mechanisms.

The intra-metropolitan context provides an ideal setting for testing these dynamics, as the Budapest agglomeration experienced housing development that consistently lagged behind population redistribution during our study period. This scarcity condition, combined with the reactive, unplanned character of post-socialist suburban development, creates circumstances where material factors cannot serve as pre-emptive attracting forces. These findings answer our research questions directly: social ties affect migration direction, dense local networks matter more than loose connections, and network effects intensify when housing is scarce. The substitution mechanism demonstrates that networks matter conditionally rather than uniformly, shaping not just whether people move, but where they can successfully relocate when formal infrastructure proves inadequate.

Appendix

A.1. Measurement of clique cohesion - robustness of the density measure

Our analysis operationalizes the cohesion of clique-like communities in destination settlements through mean edge density. A methodological concern is that standard network measures may conflate tie strength with the sheer size or volume of connections, and that the selection of density among several candidate statistics is arbitrary. To address this concern, we examine multiple possible network statistics. As described in Section 2.3.1, for every connection to a clique-like group larger than five in another settlement, we computed the number of such connections (called here strong connections), the number of nodes, the number of edges, the average degree, the edge density, and the global clustering coefficient, and aggregated each to the settlement-pair level by both mean and median. If density were an idiosyncratic or opportunistic choice, we would expect it to behave as one measure among many interchangeable ones; if it captures a distinct construct - the internal knittedness of communities, as our theoretical framework requires - the measurement space should show a clear structure separating cohesion from volume. Table A.1.1. reports the pairwise correlations among all eleven measures.

Table A.1.1. Correlations among clique statistics and connection counts at the settlement-pair level												
#	Measure	1	2	3	4	5	6	7	8	9	10	11
1	Strong connections	—	0,378	0,055	0,424	-0,070	0,136	-0,189	-0,279	-0,313	-0,302	-0,223
2	Nodes (mean)	0,077	—	0,747	0,957	0,644	0,781	0,456	-0,397	-0,389	-0,299	-0,327
3	Nodes (median)	0,008	0,860	—	0,653	0,854	0,644	0,597	-0,330	-0,308	-0,200	-0,262
4	Edges (mean)	0,008	0,896	0,741	—	0,638	0,876	0,527	-0,201	-0,202	-0,155	-0,174
5	Edges (median)	-0,033	0,759	0,895	0,807	—	0,794	0,882	0,062	0,075	0,124	0,054
6	Degree (mean)	-0,026	0,849	0,809	0,843	0,790	—	0,806	0,155	0,159	0,163	0,116
7	Degree (median)	-0,069	0,742	0,871	0,714	0,859	0,926	—	0,390	0,395	0,392	0,326
8	Density (mean)	-0,229	-0,294	-0,222	-0,112	-0,081	0,040	0,082	—	0,968	0,797	0,773
9	Density (median)	-0,237	-0,272	-0,196	-0,098	-0,065	0,054	0,099	0,970	—	0,769	0,756
10	Clustering (mean)	-0,171	-0,169	-0,110	-0,058	-0,028	0,096	0,132	0,783	0,754	—	0,939
11	Clustering (median)	-0,161	-0,192	-0,143	-0,077	-0,048	0,064	0,097	0,757	0,740	0,957	—

The eleven measures do not form an undifferentiated pool. They separate into two internally coherent and mutually distinct blocks. First, the two cohesion statistics converge. Mean edge density and mean global clustering are strongly correlated ($r = 0.78$ Pearson, 0.80 Spearman), and the same holds for their median counterparts. The two statistics are computed differently - density as the proportion of realized ties among community members, clustering as the prevalence of closed triads - yet they agree on which settlement pairs are connected through tightly knit communities and which through loose ones. They behave as two indicators of one underlying construct.

Second, the size and volume statistics likewise converge among themselves. Node counts, edge counts, and average degree correlate at 0.71 – 0.93 (Pearson) within the block. They jointly measure how large the identified communities are. Third, the two blocks are nearly orthogonal, and where they relate, the relation is negative. Density correlates at -0.29 with mean community size, at -0.11 with mean edge count, and at 0.04 with average degree. If density were a disguised volume measure, these correlations would be strongly positive. The mild negative association between density and size is, in fact, exactly what a genuine cohesion measure is expected to show, since edge density mechanically tends to decline with size in empirical social networks.

The correlation structure shows that density is one of two strongly convergent indicators of a single cohesion construct, and not an isolated statistic that happens to perform well. Given this, the selection of density over clustering is a choice among closely related indicators of the same construct, and we made it on grounds that were fixed by theory and prior evidence before estimation, not by the results.

First, the theoretical mechanism of the article concerns support capital in the sense of Jackson 2020: the capacity of tight-knit groups to mobilize assistance, vouching, and temporary accommodation for arriving migrants. Support of this kind depends on the proportion of realized ties among the members of a community, which is precisely what edge density measures. The clustering coefficient captures triadic closure (whether one's friends know each other), which is a related but a conceptually narrower property. Density is the direct operationalization of tight-knittedness.

Second, density is bounded on the unit interval and normalized by community size by construction, which makes it comparable across cliques of widely varying sizes, an essential property in our data, where identified communities range from the minimal threshold upwards. The global clustering coefficient, by contrast, is estimated from the count of possible triangles, which makes it noisy in small graphs; and the raw size measures would require exactly the normalization that density already provides. Among the available statistics, density is the most stable in the regime of small structures where our measurement operates.

Third, precedent. Density is the standard operationalization of network cohesion in the social capital and network closure literature, from Coleman's closure argument onwards, and it is the default measure in ego-network and community-cohesion research. Our choice follows the established practice of the field rather than departing from it.

Density thus is the measure that theory and prior evidence predict should carry the cohesion construct, and the empirical structure of the measurement space confirms that it does: it converges with its theoretical sibling, it is uncontaminated by size, and it is independent of connection volume. We note the honest limit of this exercise: the table establishes the convergent and discriminant validity of the measures themselves, not the invariance of regression estimates under re-specification. But given a correlation of 0.78–0.80 between the cohesion indicators, the choice between them cannot be the source of the article's findings - the measures largely agree on where tightly knit communities are, and the results follow from that agreement.

4. SUMMARY AND REFLECTIONS

In the introduction, I condensed the main argument of this dissertation into an overarching research question: How do interdependent flow structures shape the spatial patterning of internal migration in Hungary – and what can network methods reveal about this structure that unit-based and dyadic models cannot? In this concluding chapter I aim to give concrete answers to this question, together with their limitations. Since the question has an empirical and a methodological component, and the three articles address it on three nested spatial scales, the answers can be summarized in a simple grid shown in Table 1. The rest of this chapter elaborates these answers, and reflects on their position in migration theory.

Table 1: Short answers to the overarching research question.		
	Empirical component: how do interdependent flow structures shape the spatial patterning of internal migration in Hungary?	Methodological component: what do network methods reveal that unit-based and dyadic models cannot?
Article 1 — national, multiscale	Flows organize into dense local cliques around centers whose maximal scales neither population nor administrative rank can predict exactly, while the lowest resolution still converges to one Budapest-centered cluster.	A radiation-model null shows that population and geographical layout alone cannot reproduce multiscale clustering - the deviations mark where social and economic structure operates.
Article 2 — inter-district, price shock	Despite a 162% deterioration in the capital's affordability ratio, the flow network expanded along pre-existing flows instead of reorganizing, with real differentiation in the middle of the urban hierarchy.	Count-ERGMs show that individual unit and individual dyad specifications can produce misleading push-pull estimates when reciprocity and sender-receiver heterogeneity are left uncontrolled.
Article 3 — intra-metropolitan	Besides gravity effects, past social ties, and especially ties to closely knit communities direct intra-metropolitan migration flows. The effect of ties intensifies as destination housing gets scarcer	Social connection measures improve prediction beyond pure gravity specifications, but only incrementally.

The positions of the three articles of this thesis in the context of internal migration research are complex on many levels, even fundamental ones, like the sheer definition of internal migration. Internal migration and residential mobility are traditionally differentiated from each other (Roseman 1971, Niedomysl 2011). The articles do not differentiate between internal migration and residential mobility in a strict sense, but they cover different aspects of it. As Coulter et al. 2015 points out, traditional categorizations can only achieve fuzzy differentiation, and there is no clear cut difference in motives and reasons between these spatial mobility types. This dissertation is also intended as a good reference of how the network framework and statistical network science methods can resolve the fuzzy differentiation of residential mobility and internal migration, constructing network models of migration systems that do not rely on pre-defined categories, but cover all migratory behaviour that is affected by social effects in the focus of a given research question. Administrative data sources used for all analysis only contain information about the origin and destination settlements (municipalities) of movers, thus shortest distance, intra-municipality residential mobility events are omitted, but short distance moves directed into neighbouring municipalities are analysed.

The first article is the most interesting for the mobility-migration distinction. By providing structural description on flow structure on multiple spatial scales at the same time, it tries to explain how short, medium, and long distance moves form a structure together that can be analysed with different scale zooming. By showing that there is a distinct multi-scale flow structure at place in the migration system of the country, the analysis provides evidence on how different forms of spatial residential mobility all happen in a connected system. The article on the relationship of house price differentials and flow structure omits moves that happen inside districts of the country, so mobility behaviour in small-medium, dense cliques of settlements. This way the effect of prices is only shown for medium and long distance moves, and the effect of a house price boom in the capital on flows in the whole country can be demonstrated. The article on the relationship between social connections and intra-metropolitan residential migration in the Budapest agglomeration has the smallest spatial scale. Even there, medium-distance moves can happen between different

sectors of the agglomeration, and even shorter moves to a neighbouring settlement can change everyday life of households.

Foundational macro-level migration theories like Ravenstein 1885 and Zelinsky 1971 show that while internal migration flow networks have internal mechanisms they are embedded in their own developmental context. Ravenstein 1885 shows that the absorbing and dispersing forces are set by the growth of large centers of commerce and industry. This growth pulls migrants from nearby settlements, and the gaps they leave are filled by more distant ones, and this process ripples through a country. Zelinsky and the migration transition hypothesis he formed in 1971 emphasizes that such absorbing forces of growing centers are only the system-defining element in developmental stages when strong centralization happens because of urbanization and industrialization, and in post-industrial, service oriented developmental stages, internal migration becomes more complex spatially, most migration happens in a complex lattice of cities and their agglomerations. Castells 2020 emphasized that such stages of development are heavily defined by information and telecommunication technologies, which lead to a new spatial hierarchy of centers and peripheries. This change in urban development was formalized in polycentric urban development theory. According to polycentricity theorists, during the last decades of the 20th century, and early decades of the 21st, in developed economies, a partial shift happened from hierarchies to networks through the change from a fixed number of spatial scales to variable number of spatial scales, and from the primacy of centrality to the primacy of nodality (Meijers 2007).

The first essay of this dissertation shows that internal migration flow structure in Hungary is defined by polycentricity, but also by the classical pull force of the growing capital city of its country. It is visible that on lower spatial scales, migration clusters around local flow centers, and a large share of internal migration happens in local, small, dense cliques of settlements. On larger spatial scales, these dense clusters form larger ones around influential regional centers. This in itself is no evidence for polycentricity. But we can see, that population of settlements is no single predictor for the maximal spatial scale they still behave as flow center, and neither is their administrative rank. There are villages that keep their flow clusters on lower resolutions as some towns, and there are less populous towns that keep theirs on a lower resolutions than larger cities. It is interesting, that the first analyses for

this paper happened before Baja and Esztergom became officially towns of county rank, and they keep their center roles in their respective clusters on similar spatial scales than other towns of county ranks. It was reassuring that the government officially gave them the title after the analyses, showing that the multi-scale community detection method had some predictive power on this. This also shows that polycentric development happens in a very complex manner, and many towns have some center role that is hidden from analysis working with classical approaches based on administrative nomenclatures. But multiscale community structure on the lowest resolution converges to one large cluster of the country, showing that the economic dominance of Budapest is working in similar ways that Ravenstein described for 19th century UK. However, the same multiscale community detection method for the radiation model predicted synthetic network showed that clustering around local flow centers or Budapest cannot be explained purely by populations and geographic population layout, underlining the focus of polycentricity theory on economic and amenity factors.

The second article is also strongly, but indirectly touching the subject of polycentric urban development, but from another perspective. It analyses the immediate restructuring of the migration flow network of Hungary during the 2010s internal migration surge, and the 2010s house price boom, which happened simultaneously. The count-ERGM modeling provides evidence on the fact that the hub dominance of Budapest was not weakened, even while affordability ratios deteriorated, but also that there was differentiation across the urban hierarchy: regional centers and other towns, in the sense that they were able to gain population retention effects, even after controlling for network mechanisms in the flow system. Such a population retention effect theoretically can lead to stronger centralities on the long run, but the study did not assess whether it remained after the 2010's. It is possible, that the internal migration irregularities brought by the covid pandemic made the findings of the second article obsolete on the long run, but that still would not negate that some polycentricity differentiation can happen during the times of economic boom cycles.

The second article provides important evidence also for how push-pull effects, like the effect of house prices should be modelled more according to foundational work on the subject. Lee 1966 emphasized that the push-pull calculus of migration is embedded in pre-existing flow structures. The second article shows that individual

unit and individual dyad approaches can lead to misleading estimates for different migration factors, if we do not control for flow network mechanisms. After controlling for reciprocity and individual sender and receiver heterogeneity, economic and amenity factors could lose or gain significant effects. But for important factors, like house prices, or population, or distance between districts, controlling for network mechanisms did not change significances or the positive-negative effect directions. This might indicate that migration factors such as housing work partly independently from flow structures, so individual unit and dyad estimates should not be looked over. But, in my view, a general methodological claim can be made for future non-network models on the topic. Including for example a simple reciprocity term is possible in almost all regression frameworks, and that might enhance the robustness and precision of the parameter estimates.

The third article about the effect of social networks in the absence of proper levels of housing in the agglomeration of Budapest also provides important evidence on push-pull models, albeit in special circumstances. It is intuitive that for a migration factor to be a pull effect, it has to be in place beforehand, and probably for a longer time period. The special post-socialist suburbanization processes that took place in Hungary were defined by suburban settlement development dominated by private and market actors, and not central planning (Stanilov 2007, Sýkora and Bouzarovski 2012). This gave development a spontaneous nature to some extent, exemplified by edge cases such as spontaneous population inflow to allotment gardens (Gagyí and Vígvári 2025). Hypothetically, in such circumstances, other factors could serve as a pull effect, like social networks. The third article, using online social network data, shows that indeed, online social connections in the past serves as a predictor of future inter-settlement migration in the Budapest agglomeration area, especially, when the connections have a certain quality to them, and they point to closely knit communities. By testing interaction terms, the analysis shows that the effect of social connections gets stronger as housing levels in a destination settlement get lower. But the models of the article also show that including these social connection measures to gravity-specified push-pull models, there is only an incremental step forward in explanatory power. This underlines that classic, material migration factors are the dominant forces behind population flow even in such special circumstances.

By the application of social connection measures, the third article strongly connects to systematic theories of (internal) migration, like the works of Massey and Mabogunje. Massey builds his systematic models on migrant networks, and shows that the more migration happens between two location, the more information can flow through personal connections, making the costs of mobility lower, and barriers easier to overcome (Massey 1990, Massey et al. 1993). His theories were first and foremost aimed at international migration, and applications to internal migration are still rare. For international migration, costs are higher, and barriers are harder to overcome, enhancing the effect of migrant networks, especially between countries historically connected by for example colonialization. But the third article shows, such forces can be in effect for short and medium scale residential migration also, even if in a weaker form. de Haas 2010 calls the simple idea that migration flows follows economic opportunity differences a truism, and points out that push-pull models are unable to explain the actual, geographically patterned flow structure of migration. By the application of social network methods, the third article shows that these factors can help such explanations even when included in classical gravity type push-pull model specifications.

The most important aspect generally of systematic theories is path-dependency. Mabogunje described a simple analytical model of urban-rural migration based on the idea that rural migrants provide information feedback to their origin villages, making migration flows between certain village-town pairs more robust, leading to path dependency in the system. Path dependency is a central notion in all three articles. The first paper, which is a multi-level structural analysis aimed at finding level-connecting structural elements, works with cross-sectional data, and finds district, county, and region borders in the clustering of migration flows, without using any geographical information. This is important evidence of path-dependency, and an indicator of the scales at which information feedbacks can work in the long term. But it does not explain why these are the scales that are appropriate for feedback loops to exist - that might require an analysis of the settlement history over the last 1000 years in the Carpathian Basin.

The second paper is built around analysing change in an important migration-inducing factor (house prices) in a system-defining location: the extremely dominant capital of Hungary. It shows that path-dependency is stronger, at least in the short

term, than a major housing boom, and the internal migration network does not experience major structural change; it expands along pre-existing flows and relations; with differentiations in the sending and receiving behaviour in the middle of the urban hierarchy. But the paper does not show how house prices could lead to these pre-existing flow patterns, and only works on an immediate timescale. The results do not show how migration flow structures would change if the affordability ratios were constant for a long time period, and whether that would break path-dependency. The third paper shows that, on the intra-metropolitan scale, social connections, and connections to dense friend groups between settlements can predict future residential migration between them. But it does not account for the fact that perhaps the same path-dependency connects both migration and social connections, or how migration between settlements that have connected cliques can enforce more migration between them.

Across the three studies, a consistent pattern emerges: migration flow structures exhibit path-dependency and interdependence that individual unit and dyad approaches cannot fully capture, yet classical migration factors retain their explanatory relevance when embedded in network specifications. This suggests that the network perspective and traditional push-pull frameworks are complementary rather than competing. Generally, I hope the three articles in this thesis will prove themselves as incremental, but useful experimental steps in a direction of internal migration research which builds simultaneously on foundational classic, systems-thinking social theory, and state of the art network science methodology.

REFERENCES

- Abramski, Katherine ; Katenka, Natallia and Hutchison, Marc (2020): A Network-Based Analysis of International Refugee Migration Patterns Using GERGMs. In Hocine Cherifi ; Sabrina Gaito ; José Fernando Mendes ; Esteban Moro and Luis Mateus Rocha (Eds). *Complex Networks and Their Applications VIII*. Springer International Publishing. doi:10.1007/978-3-030-36683-4_32
- Aksoy, Ozan and Yıldırım, Sinan (2024): A Model of Dynamic Flows: Explaining Turkey's Interprovincial Migration. *Sociological Methodology*, 54(1), pp. 52–91. doi:10.1177/00811750231184460
- Alis, Christian ; Legara, Erika Fille and Monterola, Christopher (2021): Generalized radiation model for human migration. *Scientific Reports*, 11(1), pp. 22707. doi:10.1038/s41598-021-02109-1
- Batten, David F (1995): Network Cities: Creative Urban Agglomerations for the 21st Century. *Urban Studies*, 32(2), pp. 313–327. doi:10.1080/00420989550013103
- Batty, Michael (2013): *The New Science of Cities*. The MIT Press, Cambridge, MA.
- Bálint, Lajos and Obádovics, Csilla (2021): Belföldi vándorlás. *Demográfiai portré*, pp. 211–226.
- Bell, M. ; Blake, M. ; Boyle, P. ; Duke-Williams, O. ; Rees, P. ; Stillwell, J. and Hugo, G (2002): Cross-National Comparison of Internal Migration: Issues and Measures. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 165(3), pp. 435–464. doi:10.1111/1467-985X.t01-1-00247
- Bell, Martin ; Charles-Edwards, Elin ; Ueffing, Philipp ; Stillwell, John ; Kupiszewski, Marek and Kupiszewska, Dorota (2015): Internal Migration and Development: Comparing Migration Intensities Around the World. *Population and Development Review*, 41(1), pp. 33–58. doi:10.1111/j.1728-4457.2015.00025.x
- Berry, Brian J. L. (1964): Cities as systems within systems of cities. *Papers in Regional Science*, 13(1), 147–163. <https://doi.org/10.1111/j.1435-5597.1964.tb01283.x>

Block, Per ; Stadtfeld, Christoph and Robins, Garry (2022): A statistical model for the analysis of mobility tables as weighted networks with an application to faculty hiring networks. *Social Networks*, 68, pp. 264–278.

doi:10.1016/j.socnet.2021.08.003

Bloze, Gintautas (2009): Interregional migration and housing structure in an East European transition country: A view of Lithuania 2001–2008. *Baltic Journal of Economics*, 9(2), pp. 47–66. doi:10.1080/1406099X.2009.10840461

Blumenstock, Joshua E ; Chi, Guanghua and Tan, Xu (2025): Migration and the Value of Social Networks. *The Review of Economic Studies*, 92(1), pp. 97–128.

doi:10.1093/restud/rdad113

Büchel, Konstantin ; Ehrlich, Maximilian V. ; Puga, Diego and Viladecans-Marsal, Elisabet (2020): Calling from the outside: The role of networks in residential mobility. *Journal of Urban Economics*, 119, pp. 103277.

doi:10.1016/j.jue.2020.103277

Burger, Martijn J. ; Knaap, Bert van der and Wall, Ronald S (2014): Polycentricity and the Multiplexity of Urban Networks. *European Planning Studies*, 22(4), pp. 816–840. doi:10.1080/09654313.2013.771619

Burger, Martijn and Meijers, Evert (2012): Form follows function? Linking morphological and functional polycentricity. *Urban Studies*, 49(5), 1127–1149.

<https://doi.org/10.1177/0042098011407095>

Camagni, Roberto P. and Salone, Carlo (1993): Network urban structures in Northern Italy: Elements for a theoretical framework. *Urban Studies*, 30(6), 1053–1064.

<https://doi.org/10.1080/00420989320080941>

Castells, Manuel (2020): “Space of Flows, Space of Places: Materials for a Theory of Urbanism in the Information Age”. In *The City Reader* 7th edn. Routledge.

Cavalleri, Maria Chiara ; Luu, Nhung and Causa, Orsetta (2021): Migration, housing and regional disparities: A gravity model of inter-regional migration with an application to selected OECD countries. *OECD Economics Department Working Papers*. doi:10.1787/421bf4aa-en

Chan, Sewin (2001): Spatial Lock-in: Do Falling House Prices Constrain Residential Mobility?. *Journal of Urban Economics*, 49(3), pp. 567–586.
doi:10.1006/juec.2000.2205

Charyyev, Batyr and Gunes, Mehmet Hadi (2019): Complex network of United States migration. *Computational Social Networks*, 6(1), pp. 1. doi:10.1186/s40649-019-0061-6

Christaller, Walter (1933): *Die zentralen Orte in Süddeutschland*. Jena: Gustav Fischer. English translation: *Central Places in Southern Germany* (C. W. Baskin, Trans., 1966). Englewood Cliffs, NJ: Prentice-Hall.

Clark, William A.V. and Lisowski, William (2019): Extending the human capital model of migration: The role of risk, place, and social capital in the migration decision. *Population, Space and Place*, 25(4), pp. e2225. doi:10.1002/psp.2225

Clemente, G. P. and Grassi, R (2018): Directed clustering in weighted networks: A new perspective. *Chaos, Solitons & Fractals*, 107, pp. 26–38.
doi:10.1016/j.chaos.2017.12.007

Coffey, William J (1998): Urban Systems Research: An Overview. *Canadian Journal of Regional Science*, 21(3), pp. 327–364.

Coulter, Rory; Ham, Maarten van and Findlay, Allan M (2016): Re-thinking residential mobility: Linking lives through time and space. *Progress in Human Geography*, 40(3), pp. 352–374. doi:10.1177/0309132515575417

Csizmady, Adrienne; Bagyura, Márton and Olt, Gergely (2022): From a Small Village to an Exclusive Gated Community: Unplanned Suburbanisation and Local Sovereignty in Post-Socialist Hungary. *Urban Planning*, 7(3), pp. 115–129.
doi:10.17645/up.v7i3.5275

Cushing, Brian and Poot, Jacques (2003): Crossing boundaries and borders: Regional science advances in migration modelling. *Papers in Regional Science*, 83(1), pp. 317–338. doi:10.1007/s10110-003-0188-5

David, Quentin ; Janiak, Alexandre and Wasmer, Etienne (2010): Local social capital and geographical mobility. *Journal of Urban Economics*, 68(2), pp. 191–204.
doi:10.1016/j.jue.2010.04.003

- Davoudi, Simin (2003): Polycentricity in European spatial planning: From an analytical tool to a normative agenda. *European Planning Studies*, 11(8), 979–999. doi:10.1080/0965431032000146169
- Dawkins, Casey J (2006): Are Social Networks the Ties that Bind Families to Neighborhoods? *Housing Studies*, 21(6), pp. 867–881. doi:10.1080/02673030600917776
- De Goei, Bastiaan ; Burger, Martijn J. ; Van Oort, Frank G. and Kitson, Michael (2010): Functional Polycentrism and Urban Network Development in the Greater South East, United Kingdom: Evidence from Commuting Patterns, 1981–2001. *Regional Studies*, 44(9), pp. 1149–1170. doi:10.1080/00343400903365102
- De Haas, Hein (2010): The Internal Dynamics of Migration Processes: A Theoretical Inquiry. *Journal of Ethnic and Migration Studies*, 36(10), pp. 1587–1617. doi:10.1080/1369183X.2010.489361
- De Haas, Hein (2021): A theory of migration: the aspirations-capabilities framework. *Comparative Migration Studies*, 9(1), pp. 8. doi:10.1186/s40878-020-00210-4
- De Vries, Jan (1984): *European Urbanization, 1500–1800*. Cambridge, MA: Harvard University Press. doi: <https://doi.org/10.4324/9780203716526>
- Desmarais, Bruce A. and Cranmer, Skyler J (2012): Statistical Inference for Valued-Edge Networks: The Generalized Exponential Random Graph Model. *PLOS ONE*, 7(1), pp. e30136. doi:10.1371/journal.pone.0030136
- Deutschmann, Emanuel (2022): *Mapping the Transnational World: How We Move and Communicate across Borders, and Why It Matters*. Princeton University Press, Princeton.
- DeWaard, Jack ; Kim, Keuntae and Raymer, James (2012): Migration Systems in Europe: Evidence From Harmonized Flow Data. *Demography*, 49(4), pp. 1307–1333. doi:10.1007/s13524-012-0117-9
- D’Ingiullo, Dario ; Odoardi, Iacopo and Quaglione, Davide (2023): Stay or emigrate? How social capital influences selective migration in Italy. *Regional Studies, Regional Science*, 10(1), pp. 529–548. doi:10.1080/21681376.2023.2205501

Dijkstra, Lewis, Poelman, Hugo and Veneri, Paolo (2019): The EU-OECD definition of a functional urban area. OECD Regional Development Working Papers, No. 2019/11. Paris: OECD Publishing.

Directorate-General for Economic and Financial Affairs (European Commission) ; Cousin, Guillaume ; Frayne, Christine ; Martins Dias, Vítor ; and Vašíček, Bořek (2025): Housing in the European Union: market developments, underlying drivers, and policies. Publications Office of the European Union.

Dong, Lei ; Li, Ruiqi ; Zhang, Jiang and Di, Zengru (2016): Population-weighted efficiency in transportation networks. *Scientific Reports*, 6(1), pp. 26377. doi:10.1038/srep26377

Dragan, Alexandru ; Crețan, Remus ; Jucu, Ioan Sebastian and Oancea, Oana Andreea (2024): Rural Landscapes as Cultural Heritage and Identity along a Romanian River. *Heritage*, 7(8), pp. 4354–4373. doi:10.3390/heritage7080205

Engelhardt, Gary V (2003): Nominal loss aversion, housing equity constraints, and household mobility: evidence from the United States. *Journal of Urban Economics*, 53(1), pp. 171–195. doi:10.1016/S0094-1190(02)00511-9

Ermisch, John and Mulder, Clara H (2019): Migration Versus Immobility, and Ties to Parents. *European Journal of Population*, 35(3), pp. 587–608. doi:10.1007/s10680-018-9494-0

Ermisch, John and Washbrook, Elizabeth (2012): Residential Mobility: Wealth, Demographic and Housing Market Effects. *Scottish Journal of Political Economy*, 59(5), pp. 483–499. doi:10.1111/j.1467-9485.2012.00591.x

Etzo, Ivan (2008): Internal migration: a review of the literature. MPRA Paper.

Faist, Thomas (2013): The mobility turn: a new paradigm for the social sciences?. *Ethnic and Racial Studies*, 36(11), pp. 1637–1646. doi:10.1080/01419870.2013.812229

Ferreira, Fernando ; Gyourko, Joseph and Tracy, Joseph (2010): Housing busts and household mobility. *Journal of Urban Economics*, 68(1), pp. 34–45. doi:10.1016/j.jue.2009.10.007

- Ferreira, Fernando ; Gyourko, Joseph and Tracy, Joseph (2011): Housing Busts and Household Mobility: An Update. Working Paper no. 17405; NBER Working Papers.
- Fussell, Elizabeth and Massey, Douglas S (2004): The limits to cumulative causation: International migration from Mexican Urban Areas. *Demography*, 41(1), pp. 151–171. doi:10.1353/dem.2004.0003
- Gagyí, Agnes and Vigvári, András (2025): Limits and openings for peri-urban gardening in the context of post-socialist extended urbanization: a case from Budapest. *Environmental Sociology*, 11(4), pp. 465–475.
doi:10.1080/23251042.2025.2465097
- Goetz, Christopher (2013): Falling House Prices and Labor Mobility: Evidence from Matched Employer-Employee Data (SSRN Scholarly Paper no. 2320572). Rochester, NY: Social Science Research Network. doi:10.2139/ssrn.2320572
- Granovetter, Mark S (1973): The Strength of Weak Ties. *American Journal of Sociology*, 78(6), pp. 1360–1380.
- Green, Nick (2007): Functional Polycentricity: A Formal Definition in Terms of Social Network Analysis. *Urban Studies*, 44(11), pp. 2077–2103.
doi:10.1080/00420980701518941
- Greenwood, Michael J (1975): Research on Internal Migration in the United States: A Survey. *Journal of Economic Literature*, 13(2), pp. 397–433.
- Greenwood, Michael J (1985): Human Migration: Theory, Models, and Empirical Studies. *Journal of Regional Science*, 25(4), pp. 521–544. doi:10.1111/j.1467-9787.1985.tb00321.x
- Greenwood, Michael J (1997a): Internal migration in developed countries. *Handbook of Population and Family Economics*, 1, Part B, pp. 647–720.
- Greenwood, Michael J (1997b): Chapter 13 Internal migration in developing countries. In *Handbook of Population and Family Economics Vol. 1*. Elsevier.
doi:10.1016/S1574-003X(97)80005-0
- Greenwood, Michael J. and Hunt, Gary L (2003): The Early History Of Migration Research. *International Regional Science Review*, 26(1), pp. 3–37.
doi:10.1177/0160017602238983

Giuliano, Genevieve and Small, Kenneth A. (1991): Subcenters in the Los Angeles region. *Regional Science and Urban Economics*, 21(2), 163–182. doi:10.1016/0166-0462(91)90032-I

Gyourko, Joseph ; Mayer, Christopher and Sinai, Todd (2013): Superstar Cities. *American Economic Journal: Economic Policy*, 5(4), pp. 167–199. doi:10.1257/pol.5.4.167

Hall, Peter and Pain, Kathy (2006): *The Polycentric Metropolis: Learning from Mega-City Regions in Europe*. London: Routledge. doi:10.4324/9781849773911

Handcock, Mark S. (2003): Assessing degeneracy in statistical models of social networks. Working Paper 39, Center for Statistics and the Social Sciences, University of Washington.

Harvey, David (1989): *The Condition of Postmodernity: An Enquiry into the Origins of Cultural Change*. Oxford: Blackwell.

Hegedüs, József and Tosics, Iván (1996): The disintegration of the East European housing model. In D. Clapham, J. Hegedüs, K. Kintrea, & I. Tosics (Eds.), *Housing Privatization in Eastern Europe*. Westport, CT: Greenwood Press.

Henley, Andrew (1998): Residential Mobility, Housing Equity and the Labour Market. *The Economic Journal*, 108(447), pp. 414–427. doi:10.1111/1468-0297.00295

Hill, Katherine ; Hirsch, Donald and Davis, Abigail (2021): The Role of Social Support Networks in Helping Low Income Families through Uncertain Times. *Social Policy and Society*, 20(1), pp. 17–32. doi:10.1017/S1474746420000184

Hirt, Sonja A. (2012): *Iron Curtains: Gates, Suburbs and Privatization of Space in the Post-socialist City*. Chichester: Wiley-Blackwell.

Hohenberg, Paul M. and Lees, Lynn Hollen (1985): *The Making of Urban Europe, 1000–1950*. Cambridge, MA: Harvard University Press

Holland, Paul W. and Leinhardt, Samuel (1981): An Exponential Family of Probability Distributions for Directed Graphs. *Journal of the American Statistical Association*, 76(373), pp. 33–50. doi:10.1080/01621459.1981.10477598

- Hsieh, Chang-Tai and Moretti, Enrico (2019): Housing Constraints and Spatial Misallocation. *American Economic Journal: Macroeconomics*, 11(2), pp. 1–39. doi:10.1257/mac.20170388
- Hu, Haoyu ; Shen, Jianfa ; Gu, Hengyu and Zhang, Junwei (2024): Identifying the spatio-temporal dynamics of mega city region range and hinterland: A perspective of inter-city flows. *Computers, Environment and Urban Systems*, 112, pp. 102146. doi:10.1016/j.compenvurbsys.2024.102146
- Hu, Xiaoqian ; Wang, Chao ; Wu, Junjie and Stanley, H. Eugene (2020): Understanding interurban networks from a multiplexity perspective. *Cities*, 99, pp. 102625. doi:10.1016/j.cities.2020.102625
- Huang, Peng and Butts, Carter T (2023): Rooted America: Immobility and Segregation of the Intercounty Migration Network. *American Sociological Review*, 88(6), pp. 1031–1065. doi:10.1177/00031224231212679
- Hummel, Ruth M., Hunter, David R., and Handcock, Mark S. (2012): Improving simulation-based algorithms for fitting ERGMs. *Journal of Computational and Graphical Statistics*, 21(4), 920–939. doi: 10.1080/10618600.2012.679224
- Hünteler, Bettina and Mulder, Clara H (2020): Geographic Proximity to Parents, Intergenerational Support Exchange, and Migration Within Germany. *European Journal of Population*, 36(5), pp. 895–918. doi:10.1007/s10680-020-09558-w
- Jackson, Matthew O (2020): A typology of social capital and associated network measures. *Social Choice and Welfare*, 54(2), pp. 311–336. doi:10.1007/s00355-019-01189-3
- Jia, Ning ; Molloy, Raven ; Smith, Christopher and Wozniak, Abigail (2023): The Economics of Internal Migration: Advances and Policy Questions. *Journal of Economic Literature*, 61(1), pp. 144–180. doi:10.1257/jel.20211623
- Jokela, Markus ; Soini, Eetu ; Laakasuo, Michael ; Parikka, Suvi ; Rotkirch, Anna and Hämäläinen, Hans (2025): Residential mobility and social capital: Regional analysis in Finland. *Population, Space and Place*, 31(1), pp. e2857. doi:10.1002/psp.2857

- Kan, Kamhon (2007): Residential mobility and social capital. *Journal of Urban Economics*, 61(3), pp. 436–457. doi:10.1016/j.jue.2006.07.005
- Keuschnigg, Marc (2019): Scaling trajectories of cities. *Proceedings of the National Academy of Sciences*, 116(28), pp. 13759–13761. doi:10.1073/pnas.1906258116
- Kheirkhahzadeh, Masoumeh ; Lancichinetti, Andrea and Rosvall, Martin (2016): Efficient community detection of network flows for varying Markov times and bipartite networks. *Physical Review E*, 93(3), pp. 032309. doi:10.1103/PhysRevE.93.032309
- Kloosterman, Robert C. and Musterd, Sako (2001): The polycentric urban region: Towards a research agenda. *Urban Studies*, 38(4), 623–633. doi.org:10.1080/00420980120035259
- Kluge, Lucas ; Levermann, Anders and Schewe, Jacob (2022): Radiation model for migration with directional preferences. *Physical Review E*, 106(6), pp. 064138. doi:10.1103/PhysRevE.106.064138
- Kocsis, János Balázs and Tomay, Kyra (2024): Suburban sustainability in Budapest agglomeration – The case of Törökbálint. *Sustainability*, 16(8), 3449. doi:10.3390/su16083449
- Kocsis, János Balázs and Varga, Virág (2025): Emerging subcenters in the inner Budapest metropolitan area. *Journal of Urban Affairs*. 1–23. doi:10.1080/07352166.2025.2573629
- Kok, Herman and Kovács, Zoltán (1999): The process of suburbanization in the agglomeration of Budapest. *Netherlands Journal of Housing and the Built Environment*, 14(2), pp. 119–141. doi:10.1007/BF02496818
- Kovács, Zoltán ; Farkas, Zsolt Jenő ; Egedy, Tamás ; Kondor, Attila Csaba ; Szabó, Balázs ; Lennert, József ; Baka, Dorián and Kohán, Balázs (2019): Urban sprawl and land conversion in post-socialist cities: The case of metropolitan Budapest. *Cities*, 92, pp. 71–81. doi:10.1016/j.cities.2019.03.018

Kovács, Zoltán and Tosics, Iván (2014): Urban sprawl on the Danube: The impacts of suburbanization in Budapest. In K. Stanilov & L. Sýkora (Eds.), *Confronting Suburbanization: Urban Decentralization in Postsocialist Central and Eastern Europe* (pp. 33–64). Chichester: Wiley-Blackwell.

Kováts, Bence and Kohl, Sebastian (2025): Varieties of East-Central European housing tenure structure: The long life of a north-south cleavage line. *Journal of Urban Affairs*, 47(10), pp. 3575–3596. doi:10.1080/07352166.2024.2351397

Központi Statisztikai Hivatal (2014): Magyarország településhálózata 1.: Agglomerációk, településegységek. Budapest: Központi Statisztikai Hivatal.

Krings, Gautier ; Calabrese, Francesco ; Ratti, Carlo and Blondel, Vincent D (2009): Urban gravity: a model for inter-city telecommunication flows. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(07), pp. L07003. doi:10.1088/1742-5468/2009/07/L07003

Krivitsky, Pavel N. (2012): Exponential-family random graph models for valued networks. *Electronic Journal of Statistics*, 6(none), pp. 1100–1128. doi:10.1214/12-EJS696

Krivitsky, Pavel N. and Handcock, Mark S (2014): A Separable Model for Dynamic Networks. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 76(1), pp. 29–46. doi:10.1111/rssb.12014

Lee, Everett S (1966): A theory of migration. *Demography*, 3(1), pp. 47–57. doi:10.2307/2060063

Leetmaa, Kadri and Tammaru, Tiit (2007): Suburbanization in countries in transition: Destinations of suburbanizers in the Tallinn metropolitan area. *Geografiska Annaler, Series B: Human Geography*, 89(2), 127–146. doi:10.1111/j.1468-0467.2007.00244.x

Leetmaa, Kadri ; Tammaru, Tiit and Anniste, Kristi (2009): From Priority-Led to Market-Led Suburbanisation in a Post-Communist Metropolis. *Tijdschrift Voor Economische En Sociale Geografie*, 100(4), pp. 436–453. doi:10.1111/j.1467-9663.2009.00551.x

- Lei, Weiqian ; Jiao, Limin ; Xu, Zhibang ; Zhou, Zhengzi and Xu, Gang (2021): Scaling of urban economic outputs: insights both from urban population size and population mobility. *Computers, Environment and Urban Systems*, 88, pp. 101657. doi:10.1016/j.compenvurbsys.2021.101657
- Leifeld, Philip ; Cranmer, Skyler J. and Desmarais, Bruce A (2018): Temporal Exponential Random Graph Models with btergm: Estimation and Bootstrap Confidence Intervals. *Journal of Statistical Software*, 83, pp. 1–36. doi:10.18637/jss.v083.i06
- Lengyel, Balázs ; Varga, Attila ; Ságvári, Bence ; Jakobi, Ákos and Kertész, János (2015): Geographies of an Online Social Network. *PLOS ONE*, 10(9), pp. e0137248. doi:10.1371/journal.pone.0137248
- Lerch, Mathias (2016): Internal and International Migration Across the Urban Hierarchy in Albania. *Population Research and Policy Review*, 35(6), pp. 851–876. doi:10.1007/s11113-016-9404-2
- Lesage, James P. and Ha, Christina L (2012): The Impact of Migration on Social Capital: Do Migrants Take Their Bowling Balls with Them? *Growth and Change*, 43(1), pp. 1–26. doi:10.1111/j.1468-2257.2011.00575.x
- Li, Fangzhou ; Feng, Zhiming ; Li, Peng and You, Zhen (2017): Measuring directional urban spatial interaction in China: A migration perspective. *PLOS ONE*, 12(1), pp. e0171107. doi:10.1371/journal.pone.0171107
- Li, Xiaomeng and Neal, Zachary P (2024): Are larger cities more central in urban networks: A meta-analysis. *Global Networks*, 24(2), pp. e12467. doi:10.1111/glob.12467
- Liang, Xiaofan ; Hidalgo, César A. ; Balland, Pierre-Alexandre ; Zheng, Siqi and Wang, Jianghao (2024): Intercity connectivity and urban innovation. *Computers, Environment and Urban Systems*, 109, pp. 102092. doi:10.1016/j.compenvurbsys.2024.102092
- Liu, Xingjian ; Derudder, Ben and Wu, Kang (2016): Measuring Polycentric Urban Development in China: An Intercity Transportation Network Perspective. *Regional Studies*, 50(8), pp. 1302–1315. doi:10.1080/00343404.2015.1004535

Lösch, August (1940): *Die räumliche Ordnung der Wirtschaft*. Jena: Gustav Fischer.
English translation: *The Economics of Location* (W. H. Woglom, Trans., 1954). New Haven, CT: Yale University Press.

Lőrincz, László and Németh, Brigitta (2022): How Social Capital is Related to Migration Between Communities? *European Journal of Population*, 38(5), pp. 1119–1143. doi:10.1007/s10680-022-09642-3

Luca, Massimiliano ; Lepri, Bruno ; Frias-Martinez, Enrique and Lutu, Andra (2022): Modeling international mobility using roaming cell phone traces during COVID-19 pandemic. *EPJ Data Science*, 11(1), pp. 22. doi:10.1140/epjds/s13688-022-00335-9

Lucas, Robert E. B (1997): Internal migration in developing countries. *Handbook of Population and Family Economics*, 1, Part B, pp. 721–798.

Lux, Martin and Sunega, Petr (2012): Labour Mobility and Housing: The Impact of Housing Tenure and Housing Affordability on Labour Migration in the Czech Republic. *Urban Studies*, 49(3), pp. 489–504. doi:10.1177/0042098011405693

Mabogunje, Akin L (1970): Systems Approach to a Theory of Rural-Urban Migration. *Geographical Analysis*, 2(1), pp. 1–18. doi:10.1111/j.1538-4632.1970.tb00140.x

MacDonald, John S. and MacDonald, Leatrice D. (1964): Chain migration, ethnic neighborhood formation and social networks. *The Milbank Memorial Fund Quarterly*, 42(1), 82–97. doi: <https://doi.org/10.2307/3348581>

Macedo, Mariana ; Lotero, Laura ; Cardillo, Alessio ; Barbosa, Hugo and Menezes, Ronaldo (2020): Gender Patterns of Human Mobility in Colombia: Reexamining Ravenstein's Laws of Migration. In Hugo Barbosa ; Jesus Gomez-Gardenes ; Bruno Gonçalves ; Giuseppe Mangioni ; Ronaldo Menezes and Marcos Oliveira (Eds). *Complex Networks XI*. Springer International Publishing. doi:10.1007/978-3-030-40943-2_23

Massey, D. S (1990): Social structure, household strategies, and the cumulative causation of migration. *Population Index*, 56(1), pp. 3–26.

- Massey, Douglas S. ; Arango, Joaquín ; Hugo, Graeme ; Kouaouci, Ali ; Pellegrino, Adela and Taylor, J. Edward (1993): Theories of International Migration: A Review and Appraisal. *Population and Development Review*, 19(3), pp. 431–466.
doi:10.2307/2938462
- McNeill, Graham ; Bright, Jonathan and Hale, Scott A (2017): Estimating local commuting patterns from geolocated Twitter data. *EPJ Data Science*, 6(1), pp. 24.
doi:10.1140/epjds/s13688-017-0120-x
- Meeteren, Michiel van ; Poorthuis, Ate ; Derudder, Ben and Witlox, Frank (2016): Pacifying Babel’s Tower: A scientometric analysis of polycentricity in urban research. *Urban Studies*, 53(6), pp. 1278–1298. doi:10.1177/0042098015573455
- Meijers, Evert J. (2007): From Central Place to Network Model: Theory and Evidence of a Paradigm Change. *Tijdschrift Voor Economische En Sociale Geografie*, 98(2), pp. 245–259. doi:10.1111/j.1467-9663.2007.00394.x
- Meijers, Evert J. and Burger, Martijn J. (2017): Stretching the concept of 'borrowed size'. *Urban Studies*, 54(1), 269–291. doi:10.1177/0042098015597642
- Menyhért, Márton ; Bokányi, Eszter ; Corten, Rense ; Heemskerk, Eelke M. ; Kazmina, Yuliia and Takes, Frank W. (2025). Connectivity and community structure of online and register-based social networks. *EPJ Data Science*, 14(1), pp. 8.
doi:10.1140/epjds/s13688-025-00522-4
- Michaelides, Marios (2011): The effect of local ties, wages, and housing costs on migration decisions. *The Journal of Socio-Economics*, 40(2), pp. 132–140.
doi:10.1016/j.socec.2011.01.010
- Milne, William J (1993): Macroeconomic Influences on Migration. *Regional Studies*, 27(4), pp. 365–373. doi:10.1080/00343409312331347625
- Moch, Leslie Page (1992): *Moving Europeans: Migration in Western Europe since 1650*. Bloomington: Indiana University Press.
- Molloy, Raven ; Smith, Christopher L. and Wozniak, Abigail (2011): Internal Migration in the United States. *Journal of Economic Perspectives*, 25(3), pp. 173–196. doi:10.1257/jep.25.3.173

Mónus, Gergely and Lőrincz, László (2025): Rural-urban flows determine internal migration structure across scales. *Cities*, 163, pp. 105992.

doi:10.1016/j.cities.2025.105992

Mu, Xiaoyan ; Yeh, Anthony Gar-On ; Zhang, Xiaohu ; Wang, Jiejing and Lin, Jian (2022): Moving down the urban hierarchy: Turning point of China's internal migration caused by age structure and hukou system. *Urban Studies*, 59(7), pp. 1389–1405. doi:10.1177/00420980211007796

Mulder, Clara (2006): Population and housing: A two-sided relationship.

Demographic Research, 15, pp. 401–412. doi:10.4054/DemRes.2006.15.13

Myrdal, Gunnar (1957): *Economic Theory and Under-developed Regions*. G. Duckworth.

Najarsadeghi, Mahsa and Dorostkar, Ehsan (2022): An assessment of the role of small towns in the development of periphery population centers using network analysis. *Cities*, 131, pp. 104026. doi:10.1016/j.cities.2022.104026

Neal, Zachary P (2011): From Central Places to Network Bases: A Transition in the U.S. Urban Hierarchy, 1900–2000. *City & Community*, 10(1), pp. 49–75.

doi:10.1111/j.1540-6040.2010.01340.x

Niedomysl, Thomas (2011): How Migration Motives Change over Migration Distance: Evidence on Variation across Socio-economic and Demographic Groups. *Regional Studies*, 45(6), pp. 843–855. doi:10.1080/00343401003614266

O'Brien, Thomas ; Crețan, Remus ; Jucu, Ioan Sebastian and Covaci, Raluca Narcisa (2023): Internal migration and stigmatization in the rural Banat region of Romania.

Identities, 30(5), pp. 704–724. doi:10.1080/1070289X.2022.2109276

OECD (2020): *Cities in the World: A New Perspective on Urbanisation*. OECD Publishing, Paris.

Olney, William W. and Thompson, Owen (2024): The Determinants of Declining Internal Migration (Working Paper no. 32123). National Bureau of Economic Research. doi:10.3386/w32123

Palloni, Alberto ; Massey, Douglas S. ; Ceballos, Miguel ; Espinosa, Kristin and Spittel, Michael (2001): Social Capital and International Migration: A Test Using Information on Family Networks. *American Journal of Sociology*, 106(5), pp. 1262–1298. doi:10.1086/320817

Park, Robert E., Burgess, Ernest W. and McKenzie, Roderick Duncan (1925): *The City*. Chicago: University of Chicago Press.

Parr, John (2004): The polycentric urban region: A closer inspection. *Regional Studies*, 38(3), 231–240. doi:10.1080/003434042000211114

Partridge, Mark D. ; Rickman, Dan S. ; Ali, Kamar and Olfert, M. Rose (2009): Agglomeration spillovers and wage and housing cost gradients across the urban hierarchy. *Journal of International Economics*, 78(1), pp. 126–140. doi:10.1016/j.jinteco.2009.02.004

Peng, Chien-Wen and Tsai, I-Chun (2019): The long- and short-run influences of housing prices on migration. *Cities*, 93, pp. 253–262. doi:10.1016/j.cities.2019.05.011

Pénzes, János (2024): A magyarországi belföldi migráció néhány területi sajátossága a rendszerváltás után. *Területi Statisztika*, 64(2), pp. 141–169. doi:10.15196/TS640201

Pitoski, Dino ; Lampoltshammer, Thomas J. and Parycek, Peter (2021a): Human migration as a complex network: appropriate abstraction, and the feasibility of Network Science tools. In Peter Haber ; Thomas Lampoltshammer ; Manfred Mayr and Kathrin Plankensteiner (Eds). *Data Science – Analytics and Applications*. Springer Fachmedien. doi:10.1007/978-3-658-32182-6_17

Pitoski, Dino ; Lampoltshammer, Thomas J. and Parycek, Peter (2021b): Network Analysis of Internal Migration in Austria. *Digit. Gov.: Res. Pract.*, 2(3), pp. 25:1–25:24. doi:10.1145/3447539

Pitoski, Dino ; Lampoltshammer, Thomas J. and Parycek, Peter (2021c): Network analysis of internal migration in Croatia. *Computational Social Networks*, 8(1), pp. 10. doi:10.1186/s40649-021-00093-0

- Pitoski, Dino ; Meštrović, Ana and Schmeets, Hans (2024): The complex network patterns of human migration at different geographical scales: network science meets regression analysis. *Applied Network Science*, 9(1), pp. 35. doi:10.1007/s41109-024-00635-1
- Plane, D. A. ; Henrie, C. J. and Perry, M. J (2005): Migration up and down the urban hierarchy and across the life course. *Proceedings of the National Academy of Sciences*, 102(43), pp. 15313–15318. doi:10.1073/pnas.0507312102
- Plantinga, Andrew J. ; Détang-Dessendre, Cécile ; Hunt, Gary L. and Piguet, Virginie (2013): Housing prices and inter-urban migration. *Regional Science and Urban Economics*, 43(2), pp. 296–306. doi:10.1016/j.regsciurbeco.2012.07.009
- Poot, Jacques ; Alimi, Omoniyi ; Cameron, Michael P. and Maré, David C (2016): The Gravity Model of Migration: The Successful Comeback of an Ageing Superstar in Regional Science (SSRN Scholarly Paper no. 2864830). Rochester, NY: Social Science Research Network.
- Pred, Allan (1977): *City-Systems in Advanced Economies: Past Growth, Present Processes and Future Development Options*. London: Routledge.
doi:10.4324/9781315103136
- Prieto-Curiel, Rafael ; Schumann, Abel ; Heo, Inhoi and Heinrigs, Philipp (2022): Detecting cities with high intermediacy in the African urban network. *Computers, Environment and Urban Systems*, 98, pp. 101869.
doi:10.1016/j.compenvurbsys.2022.101869
- Quigley, John M (2002): Homeowner Mobility and Mortgage Interest Rates: New Evidence from the 1990s. *Real Estate Economics*, 30(3)
- Ravenstein, E. G (1885): The Laws of Migration. *Journal of the Statistical Society of London*, 48(2), pp. 167–227. doi:10.2307/2979181.
- Ravenstein, E. G (1889): The Laws of Migration. *Journal of the Royal Statistical Society*, 52(2), pp. 241–305. doi:10.2307/2979333
- Redford, Arthur (1926): *Labour Migration in England, 1800–1850*. Manchester: Manchester University Press.

Rees, Philip and Lomax, Nik (2019): Ravenstein Revisited: The Analysis of Migration, Then and Now. *Comparative Population Studies*, 44. doi:10.12765/CPoS-2020-10

Roseman, Curtis C (1971): Migration as a Spatial and Temporal Process. *Annals of the Association of American Geographers*, 61(3), pp. 589–598. doi:10.1111/j.1467-8306.1971.tb00809.x

Rosvall, M. ; Axelsson, D. and Bergstrom, C. T (2009): The map equation. *The European Physical Journal Special Topics*, 178(1), pp. 13–23. doi:10.1140/epjst/e2010-01179-1

Rosvall, Martin and Bergstrom, Carl T (2008): Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4), pp. 1118–1123. doi:10.1073/pnas.0706851105

Rowe, Francisco ; Bell, Martin ; Bernard, Aude ; Charles-Edwards, Elin and Ueffing, Philipp (2019): Impact of Internal Migration on Population Redistribution in Europe: Urbanisation, Counterurbanisation or Spatial Equilibrium? *Comparative Population Studies*, 44. doi:10.12765/CPoS-2019-18

Rowe, Francisco and Patias, Nikos (2020): Mapping the spatial patterns of internal migration in Europe. *Regional Studies, Regional Science*, 7(1), pp. 390–393. doi:10.1080/21681376.2020.1811139

Sardadvar, Sascha and Vakulenko, Elena (2020): Estimating and interpreting internal migration flows in Russia by accounting for network effects. *Socio-Economic Planning Sciences*, 69, pp. 100685. doi:10.1016/j.seps.2019.01.007

Sassen, Saskia (1991): *The Global City: New York, London, Tokyo*. Princeton, NJ: Princeton University Press.

Schaub, Michael T. ; Delvenne, Jean-Charles ; Yaliraki, Sophia N. and Barahona, Mauricio (2012): Markov Dynamics as a Zooming Lens for Multiscale Community Detection: Non Clique-Like Communities and the Field-of-View Limit. *PLOS ONE*, 7(2), pp. e32210. doi:10.1371/journal.pone.0032210

Schubert, Gregor (2021): House Price Contagion and U.S. City Migration Networks (SSRN Scholarly Paper no. 4984580). Rochester, NY: Social Science Research Network. doi:10.2139/ssrn.4984580

Selod, Harris and Shilpi, Forhad (2021): Rural-urban migration in developing countries: Lessons from the literature. *Regional Science and Urban Economics*, 91, pp. 103713. doi:10.1016/j.regsciurbeco.2021.103713

Simini, Filippo ; González, Marta C. ; Maritan, Amos and Barabási, Albert-László (2012): A universal model for mobility and migration patterns. *Nature*, 484(7392), pp. 96–100. doi:10.1038/nature10856

Siriban, Charles ; Bernard, Aude ; Pojani, Dorina and Wilson, Tom (2025): Internal migration responses to housing dynamics before and after COVID-19 in Australia. *Applied Geography*, 178, pp. 103548. doi:10.1016/j.apgeog.2025.103548

Smiljanić, Jelena ; Blöcker, Christopher ; Holmgren, Anton ; Edler, Daniel ; Neuman, Magnus and Rosvall, Martin (2026): Community Detection with the Map Equation and Infomap: Theory and Applications. *ACM Computing Surveys*, 58(7), pp. 1–34. doi:10.1145/3779648

Snijders, Tom A. B. ; Pattison, Philippa E. ; Robins, Garry L. and Handcock, Mark S (2006): New Specifications for Exponential Random Graph Models. *Sociological Methodology*, 36(1), pp. 99–153. doi:10.1111/j.1467-9531.2006.00176.x

Spilimbergo, Antonio and Ubeda, Luis (2004): Family attachment and the decision to move by race. *Journal of Urban Economics*, 55(3), pp. 478–497. doi:10.1016/j.jue.2003.07.004

Stanilov, Kiril (2007): Housing trends in Central and Eastern European cities during and after the period of transition. In Kiril Stanilov (Ed.). *The Post-Socialist City: Urban Form and Space Transformations in Central and Eastern Europe after Socialism*. Dordrecht: Springer Netherlands. doi:10.1007/978-1-4020-6053-3_9

Stanilov, Kiril and Sýkora, Luděk (2014): Postsocialist Suburbanization Patterns and Dynamics. In *Confronting Suburbanization*. John Wiley & Sons, Ltd. doi:10.1002/9781118295861.ch9

- Stark, Oded and Taylor, J. Edward (1989): Relative deprivation and international migration. *Demography*, 26(1), pp. 1–14. doi:10.2307/2061490
- Stawarz, Nico ; Sander, Nikola and Sulak, Harun (2021): Internal migration and housing costs – A panel analysis for Germany. *Population, Space and Place*, 27(4), pp. e2412. doi:10.1002/psp.2412
- Stein, Jeremy C (1995): Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects. *The Quarterly Journal of Economics*, 110(2), pp. 379–406.
- Stephens, Mark, Lux, Martin and Sunega, Petr (2015): Post-socialist housing systems in Europe: Housing welfare regimes by default? *Housing Studies*, 30(8), 1210–1234. doi:10.1080/02673037.2015.1013090
- Stewart, John Q (1948): Demographic Gravitation: Evidence and Applications. *Sociometry*, 11(1/2), pp. 31–58. doi:10.2307/2785468
- Sun, YiFan and Pan, Kunfeng (2014): Prediction of the intercity migration of Chinese graduates. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(12), pp. P12022. doi:10.1088/1742-5468/2014/12/P12022
- Sýkora, Luděk and Bouzarovski, Stefan (2012): Multiple Transformations: Conceptualising the Post-communist Urban Transition. *Urban Studies*, 49(1), pp. 43–60. doi:10.1177/0042098010397402
- Sýkora, Luděk and Stanilov, Kiril (2014): The Challenge of Postsocialist Suburbanization. In *Confronting Suburbanization*. John Wiley & Sons, Ltd. doi:10.1002/9781118295861.ch1
- Szabó, Tünde ; Szabó, Balázs and Kovács, Zoltán (2014): Polycentric urban development in post-socialist context: the case of the Budapest Metropolitan Region. *Hungarian Geographical Bulletin*, 63(3), pp. 287–301. doi:10.15201/hungeobull.63.3.4
- Tacoli, Cecilia (1998): Rural-urban interactions: a guide to the literature. *Environment & Urbanization*, 10(1), pp. 147–166. doi:10.1177/095624789801000105

Thomas, William I. and Znaniecki, Florian (1918): *The Polish Peasant in Europe and America* (Vol. 1). Boston: Richard G. Badger.

Tilly, Charles (1990): Transplanted networks. In V. Yans-McLaughlin (Ed.), *Immigration Reconsidered: History, Sociology, and Politics* (pp. 79–95). New York: Oxford University Press.

Timár, Judit and Váradi, Mónika Mária (2001): The uneven development of suburbanisation during transition in Hungary. *European Urban and Regional Studies*, 8(4), 349–360. doi.org:10.1177/096977640100800407

Topa, Giorgio and Zenou, Yves (2015): Neighborhood and Network Effects. In *Handbook of Regional and Urban Economics* Vol. 5. Elsevier. doi:10.1016/B978-0-444-59517-1.00009-X

Tóth, Gergő ; Wachs, Johannes ; Di Clemente, Riccardo ; Jakobi, Ákos ; Ságvári, Bence ; Kertész, János and Lengyel, Balázs (2021): Inequality is rising where social network segregation interacts with urban topology. *Nature Communications*, 12(1), pp. 1143. doi:10.1038/s41467-021-21465-0

Tóth, József (1994): Urbanization and spatial structure in Hungary. *GeoJournal*, 32(4), pp. 343–350. doi:10.1007/BF00807353

Vasanen, Antti (2012): Functional Polycentricity: Examining Metropolitan Spatial Structure through the Connectivity of Urban Sub-centres. *Urban Studies*, 49(16), pp. 3627–3644. doi:10.1177/0042098012447000

Vries, Jacob J. de ; Nijkamp, Peter and Rietveld, Piet (2001): Alonso’s Theory of Movements: Developments in Spatial Interaction Modeling. *Journal of Geographical Systems*, 3(3), pp. 233–256. doi:10.1007/PL00011478

Wachs, Johannes ; Yasseri, Taha ; Lengyel, Balázs and Kertész, János (2019): Social capital predicts corruption risk in towns. *Royal Society Open Science*, 6(4), pp. 182103. doi:10.1098/rsos.182103

Wang, Yuxia ; Wang, Fahui ; Zhang, Yi and Liu, Yu (2019): Delineating urbanization “source-sink” regions in China: Evidence from mobile app data. *Cities*, 86, pp. 167–177. doi:10.1016/j.cities.2018.09.016

- White, Michael J. and Lindstrom, David P (2005): Internal Migration. In Dudley L. Poston and Michael Micklin (Eds). *Handbook of Population*. Boston, MA: Springer US. doi:10.1007/0-387-23106-4_12
- Windzio, Michael (2018): The network of global migration 1990–2013: Using ERGMs to test theories of migration between countries. *Social Networks*, 53, pp. 20–29. doi:10.1016/j.socnet.2017.08.006
- Windzio, Michael ; Teney, Céline and Lenkewitz, Sven (2021): A network analysis of intra-EU migration flows: how regulatory policies, economic inequalities and the network-topology shape the intra-EU migration space. *Journal of Ethnic and Migration Studies*, 47(5), pp. 951–969. doi:10.1080/1369183X.2019.1643229
- Xu, Anqi (2024): Mind the gap: Exploring urban–rural differences in US inter-county migration decisions. *Demographic Research*, 50, pp. 115–130. doi:10.4054/DemRes.2024.50.4
- Xu, Zengwang (2018): The structure and dynamics of population migration among economic areas in the United States from 1990 to 2011. *Papers in Regional Science*, 97(3), pp. 785–801. doi:10.1111/pirs.12282
- Zelinsky, Wilbur (1971): The Hypothesis of the Mobility Transition. *Geographical Review*, 61(2), pp. 219–249. doi:10.2307/213996
- Zhang, Weili ; Chong, Zhaohui ; Li, Xiaojian and Nie, Guibo (2020): Spatial patterns and determinant factors of population flow networks in China: Analysis on Tencent Location Big Data. *Cities*, 99, pp. 102640. doi:10.1016/j.cities.2020.102640
- Zhao, Liquiu and Yao, Xianguo (2017): Does local social capital deter labour migration? Evidence from rural China. *Applied Economics*, 49(43), pp. 4363–4377. doi:10.1080/00036846.2017.1282146
- Zhao, Miaoxi ; Derudder, Ben and Huang, Junhao (2017): Examining the transition processes in the Pearl River Delta polycentric mega-city region through the lens of corporate networks. *Cities*, 60, pp. 147–155. doi:10.1016/j.cities.2016.08.015

Zheng, Wensheng ; Kuang, Aiping ; Wang, Xiaofang and Chen, Jing (2020): Measuring Network Configuration of the Yangtze River Middle Reaches Urban Agglomeration: Based on Modified Radiation Model. *Chinese Geographical Science*, 30(4), pp. 677–694. doi:10.1007/s11769-020-1131-2

Zhong, Chen ; Arisona, Stefan Müller ; Huang, Xianfeng ; Batty, Michael and Schmitt, Gerhard (2014): Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science*, 28(11), pp. 2178–2199. doi:10.1080/13658816.2014.914521

Zhu, Ruoxin ; Wang, Yujing ; Lin, Diao ; Jendryke, Michael ; Xie, Mingxia ; Guo, Jianzhong and Meng, Liqiu (2021): Exploring the rich-club characteristic in internal migration: Evidence from Chinese Chunyun migration. *Cities*, 114, pp. 103198. doi:10.1016/j.cities.2021.103198

Zipf, George Kingsley (1946): The P1 P2/D Hypothesis: On the Intercity Movement of Persons. *American Sociological Review*, 11(6), pp. 677–686. doi:10.2307/2087063

AUTHOR'S OWN PUBLICATIONS RELATED TO THE TOPIC

Juhász, Sándor ; Pintér, Gergő ; Kovács, Ádám J. ; Borza, Endre ; Mónus, Gergely ; Lőrincz, László and Lengyel, Balázs. (2023). Amenity complexity and urban locations of socio-economic mixing. *EPJ Data Science*, 12(1), pp. 34.
doi:10.1140/epjds/s13688-023-00413-6

Mónus, Gergely and Lőrincz, László. (2025). Rural-urban flows determine internal migration structure across scales. *Cities*, 163, pp. 105992.
doi:10.1016/j.cities.2025.105992

Rottenkolber, Alexandra ; Ali, Ola ; Mónus, Gergely ; Li, Jiaxuan ; Kim, Jisu ; Perrotta, Daniela and Akbaritabar, Aliakbar. (2026). The role of professional networks and institutional prestige in shaping the first career moves of scholars. *PNAS Nexus*, 5(6), pp. pgag168. doi:10.1093/pnasnexus/pgag168