

RESEARCH

Open Access



The role of higher education in spatial mobility

Zsolt Tibor Koszttyán^{1,2,3*} , Vivien Valéria Csányi^{1,2}, Zsuzsanna Banász^{1,2}, Ákos Jakobi^{3,4,5},
Ildikó Neumanné-Virág^{1,3} and András Telcs^{1,2,6,7}

*Correspondence:

kzst@gtk.uni-pannon.hu

¹ Department of Quantitative Methods, University of Pannonia, Egyetem str. 10, Veszprém 8200, Hungary
Full list of author information is available at the end of the article

Abstract

The role of higher education in social and spatial mobility has attracted considerable attention. However, there are very few countrywide databases that follow the career paths of graduates from their place of birth, through their enrollment in university, and ultimately to their workplace. However, in Hungary, there is an excellent source maintained by the government's Education Authority containing information on career tracks, which allows one to follow all students from their place of birth, through their choice of higher education institution, to their workplace. With the combination of gravity-like economic models and the proposed mobility network, this paper examines the mediating and retaining role of institutions. This paper also proposes how to calculate the added value of location and institution in salaries and how to use these values to explain mobility between locations. The paper also shows how economic inequities influence revealed application preferences through the asymmetry of the mobility network.

Keywords: Mobility, Graduate tracking, Temporal and spatial networks, Career tracking system

Introduction

Having knowledge of the migratory behavior of students has been considered extremely important for the relevant institutions and to those who attempt to promote and direct higher education. Certain migratory behavior is determined by socioeconomic and cultural features of regions (Lourenço and Sá 2019). Psychologically, those districts where the habitation of certain age group is higher have larger outgoing (Dotti et al. 2013; Telcs et al. 2015) and incoming streams (Beine et al. 2014), as there is a positive correlation between the amount of tutorial services and population. A larger total number of students (Dotti et al. 2013) tend to be attracted to and held by regions where the unemployment is lower and economic opportunities are greater, as the educated workforce often expects to remain in the region (Lourenço and Sá 2019). Kazakis (2019) reports that individuals favor places with lower inequality, housing prices, and taxes. Utility levels and experience factors also play a substantial role in the decision making of possible student settlers (Faggian and Franklin 2014; Franklin and Faggian 2014). On the other hand, living in such densely populated regions is more costly. Beine et al. (2014)

finds that language, quality of higher education and university rankings have positive impacts while distance and migration costs have negative impacts on international student mobility to destination areas. The features and reputations of institutions also play a substantial role. Their quality is often measured by proxies (Lourenço and Sá 2019): for example, the employability of graduates (Sá et al. 2012; Lourenço and Sá 2019), the teacher-student ratio (Sá et al. 2004), student quality (Dotti et al. 2013), research effectiveness (Adkisson and Peach 2008), and the results of international rankings (Ciriaci 2014).

The reported effect of quality differs across studies (Lourenço and Sá 2019). One study reports that there is not enough difference to identify an impact of institutional quality Sá et al. (2012). In other cases, quality seems to matter (Ciriaci 2014; Cooke and Boyle 2011). A recent paper investigates whether research quality is associated with a university's ability to attract students from other provinces in Italy. The estimates suggest that research performance is a significant predictor of student enrollment. Cross-country differences in the quality of higher education institutions (HEIs) may also play a substantial role in international and internal student mobility (Bratti and Verzillo 2019). Mobility decisions depend on the attractiveness of the origin and destination and one's age and may be justified by emotional or family reasons (Lourenço et al. 2020). However, the advantages of migration are explored from various perspectives: more interesting teaching services, distinct sociocultural skills, and the opportunity to leave one's home (Holdsworth 2009; Lourenço et al. 2020). Further research shows a negative effect of distance on students' mobility decisions (Sá et al. 2004). The direct cost of attending an HEI has received considerable attention in empirical work. Financial support packages, which at least partly cover the expenses of a college education, are available in most higher education systems. The amount of financial aid, in the form of grants and scholarships, is expected to have a positive effect on the probability of enrollment (Fuller et al. 1982; Catsiapis 1987). Nevertheless, these financial aid packages rarely cover all out-of-pocket expenses, such that students are often dependent on their families' financial resources. Household income is among the most commonly evoked factors when discussing the decision to continue studying after the secondary level, with most studies finding that the higher the income of the household is, the higher the demand for post-secondary education and the propensity to be in school after the secondary level (see, for instance, Savoca 1990; Duchesne and Nonneman 1998; Checchi 2000; Hartog and Diaz-Serrano 2007). Parental educational level and occupational status are sometimes used either to proxy for this income effect or examined in their own right, and they exert a positive influence on young people's decisions to attend higher education (e.g., Checchi 2000; Nguyen et al. 2003; Hartog and Diaz-Serrano 2007).

Application to universities (so-called application mobility) is the first step in the spatial mobility of young people. The birth-to-school and school-to-work (so-called occupation mobility) transitions can be examined using both economic models (see, for instance, Tuckman 1970; Orsuwan and Heck 2009; Niu 2015), such as gravity models (e.g., Agasisti and Dal Bianco 2007; Alm and Winters 2009; Abbott and Silles 2016; Cullinan and Duggan 2016), and social network analysis (see, e.g., González Canché 2018; Bilecen et al. 2018; Kondakci et al. 2018).

This study demonstrates advantages of combining economic and data science methods in the analysis of application and occupation mobility. Gadar and Abonyi (2018) analyzed the school-to-work transition in the framework of a bipartite network model using the integrated database of the National Tax Administration, the National Health Insurance Fund, and the Education Authority. Based on the Education Authority database, application mobility is investigated using gravity models and logistic regressions (Telcs et al. 2015); however, to the best of our knowledge, no prior work has combined economic and social science methods to analyze the dynamic structure of the mobility network. Combining time-series gravity models and dynamic network science methods allows us to understand changes in the structure of the mobility network. These models can cross-validate one another, and in this way, they can be used for model triangulation (Modell 2015).

Spatial mobility

Spatial or geographical mobility is the movement across different locations. It concerns the physical motion from one space to another (Powell and Finger 2013). It is usually distinguished from social mobility, which is the ability to move up or down in social class (typically defined in terms of wealth). However, sociologists argue that the social mobility of individuals with diverse backgrounds is adversely affected by educational experiences (Powell and Finger 2013; Haveman and Smeeding 2006; Brown 2013). Therefore, the spatial mobility of applicants (application mobility) and of graduated early-career people (occupation mobility) are usually motivated by the promise of social advancement, such as a better salary, better existence, and better social esteem (Montmarquette et al. 2002; Hilmer and Hilmer 2012). Kazakis (2019) finds a positive and significant relationship between the migration flows of skilled individuals and innovation (patents as proxy), productivity (using, e.g., total factor productivity and labor productivity as proxies), higher population density, and higher investments in R&D. He reveals that technological development is highly correlated with education. Highly innovative regions are more attractive to professionals. Grogger and Hanson (2015) analyze how economic and political conditions influence foreign students' decisions to live in the United States after receiving PhDs from US universities. The authors find that students who receive merit-based fellowships or scholarships during their studies and have more educated parents are more likely want to stay in the United States. The authors contend that a stronger US economy makes it more attractive for the graduates to remain in the United States instead of returning to a home country with a weaker economy. Lucas (2001) reveals the factors affecting the international migration of highly skilled people and some benefits of the technology transfers, international trade and capital flows induced by "brain drain". Faggian and Franklin (2014) contend that the migration of highly educated persons is fundamental for policy makers. Higher human capital can improve a region's relative position. They use a negative binomial regression to estimate a gravity-type model of the interstate migration of "college-bound" high school students in the US. Chetty (2020) measure children's outcomes in adulthood to advance the research on human capital development and find that neighborhoods have causal effects on children's long-term outcomes.

Despite the divergent results on spatial mobility in the recent social science literature, there are studies examining geographical mobility trends within countries over time (Kulu et al. 2018; Chetty 2020), especially in the case of application and occupation mobility.

Methods for exploring spatial mobility

The diverse methodological approaches for measuring spatial mobility include, on the one hand, direct techniques to quantify the volume of people's movement in space. Examples such as traffic counts (Cascetta 1984), population census data (Ette et al. 2008), or commuting surveys (Rüger et al. 2011) rely to varying degrees on concrete measures of the flows between origin and destination; however, the availability of data for such studies is often limited. Another widely used methodological solution is, therefore, the application of indirect tools for exploring geographical mobility. Researchers estimate spatial mobility numbers, for instance, from GPS tracking data (Zheng et al. 2008; Sifa-Nowicka et al. 2016; Zignani and Gaito 2010) or cell phone information (Mohall 2015; Candia et al. 2008; Gonzalez et al. 2008); such methodologies could provide reasonable assumptions on geographical mobility numbers.

Other indirect tools apply proxies, such as distance-based probabilities, to measure the volume of spatial mobility between regions. Rogerson (1990) applied geometric probability methods to estimate migration distances based on the spatial distribution of population and region shape data. It is also common to approximate spatial interactions or mobility volume between regions by taking the size of regions and spatial proximities into account, namely, by applying the gravity model approach from social physics in a geographical context. Moreover, Poot et al. (2016) explicitly highlights the use of gravity modeling in spatial mobility research.

Methodologies related to occupation (or labor) mobility account, for example, for the interaction between the returns to geographic mobility and to the level of education by applying distance functions. Based on a large dataset, Lemistre and Moreau (2009) calculated the distance between the place of education and the location of the first employment of graduated students. Their results suggest decreasing returns to spatial mobility in the distance covered and increasing returns to mobility with higher levels of education. Similarly, Magrini and Lemistre (2013) examine an income-distance trade-off model and found that the most highly skilled young people do not receive a positive wage return from migration but that less-skilled young workers do. Early-career spatial mobility was also analyzed by Venhorst et al. (2015), who applied an instrumental variable approach in their model and found positive wage returns related to spatial mobility; however, when controlling for self-selection, a strong reduction was observed in the effect of spatial mobility on job match quality. Similarly, Javakhishvili Larsen and Mitze (2015) applied treatment variables in panel econometric models of individual-based longitudinal data to determine the interconnectedness of spatial mobility and early career effects.

The gravity model

The so-called “gravity” equations are widely used in empirical analysis of foreign trade, migration and even capital flows due to mass-based spatial movements (Anderson

1979). Gravity models have become popular due to their flexibility, simplicity and high explanatory power (excellent fit) Anderson and Van Wincoop (2003). To analyze the spatial interaction between two or more locations using this mathematical model, one applies Newton's gravitational law, as with gravity in physics (Paas 2003). Countries and municipalities with high economic power exert attraction on smaller ones around them (Nemes Nagy and Tagai 2011). Attractive areas are geographical points (e.g., cities, small regions) where the attractiveness of the place is stronger than that of any other geographical point. Based on the physical analogy, there are two basic areas of application: the examination of spatial flow (the intensity of the flow) and the delimitation and demarcation of attractive areas (Nemes Nagy and Tagai 2011).

Henri-Guillaume Desart developed a version of the gravity model for analyzing passenger travel and applied it to railway planning, and the American economist Henry Carey presented a statement that resembled the notion of a gravity model in 1858 (Odlyzko 2015). According to a survey by Fotheringham et al. (2000), Carey (1858) and Ravenstein (1885) observed that there is a parallel between the movement of individuals between cities and the law of universal attraction, namely, there is a more intense flow between larger cities than between smaller towns (Fotheringham et al. 2000; Ravenstein 1889).

International trade has its own gravity model, developed by Jan Tinbergen (1962). It is a multivariate linear regression model for modeling bilateral and regional trade that is employed for analyzing cross-sectional and panel data (Tinbergen 1962; Anderson 1979). The gravity model of foreign trade, like other gravity models in social science, predicts bilateral trade flows based on the size and distance of the partner economies (Anderson and Van Wincoop 2003). It states that trade between two countries is directly linked to the "gravitational" pull of their national incomes (GDP) and inversely proportional to the distance between them (Paas 2003). The model predicts bilateral trade flows based on economic size (usually measured in GDP) and distance (Anderson 1979; Bergstrand 1985, 1989; Anderson and Van Wincoop 2003). The gravity model has been widely used to estimate the impact of a variety of policy issues, including regional trading groups, currency unions, political blocks, various trade distortions and agreements, border region activities and historical linkages (Paas 2003; Westerlund and Wilhelmsson 2011).

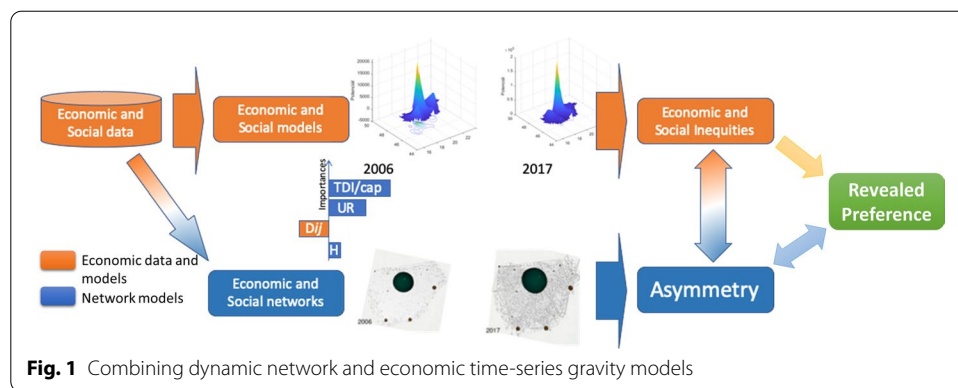
The gravity model also underlies migration studies. Most studies using the gravity model approach have sought to rationalize labor force mobility across locations, especially internal migration waves. There has been highly accurate empirical work on this topic for countries such as the United States (Ashby 2007), China (Shen 1999; Poston and Zhang 2008), Germany (Bierens and Kontuly 2008), Hungary (Cseres-Gergely 2012) and Spain (Devillanova and García-Fontes 1998). Based on the objectives of the present study, the most relevant articles are those examining international migration streams. In addition to other works, we can also mention articles on international migration to the European Union (Breitenfellner et al. 2008; Warin and Svaton 2008) and to the United States (Karemera et al. 2000) as the two main regions that draw foreign immigration. These articles tend to include standard economic variables such as per capita gross domestic product (GDP) or population, sometimes using different changes in variables, as in Warin and Svaton (2008), which is based on calculations related to GDP,

while in Karemera et al. (2000), political variables are also included, which has a negative impact on migration flows. The leader countries where most international students are recruited from include China, India and other parts of Asia. Another study finds that “China is becoming an important destination for students due to the distinctiveness of the language, the rise of its universities in global rankings and the country’s economic growth” (Ahmad and Shah 2018). According to the gravity model, the number of migrants between two regions is directly proportional to the population in each region and indirectly proportional to the squared distance between the location they leave and the region they enter.

The gravity model approach is also common in higher-education-related spatial mobility studies. Bernela et al. (2018) tested the significance of the impact of the scientific size of regions and spatial proximity on PhD mobility in France by applying the Heckman (1979)’s two-step gravity model with a selection equation to evaluate the existence of potential spatial and nonspatial proximity effects. In addition, there are examples of measuring higher-education-related application mobility with large survey-type datasets. Extensive research on the spatial mobility of graduates was performed by Venhorst et al. (2011), who applied multinomial logit models to investigate the relationships between migration and both regional economic circumstances and individual characteristics. They found that the presence of a large labor market is the most important structural economic determinant of higher retention rates in regions.

The network model

Mobility can be modeled by networks. In this case, nodes represent locations, while directed arcs represent the mobility from one location to another. The weights of arcs represent the number of domestic migrants (spatial mobility) between given locations. In network science, null model creation is a common and useful tool. The null model assumes that the network is random (Newman and Girvan 2004) and thus that the weights of the arcs are independent of one another. Liu and Murata (2010)’s null model assumes that the probability of weights on an arc depends on the distance between locations, while Gadár et al. (2018) assumes that the number of links (weights of arcs) fit well to the values predicted by a gravity model. Null models are also essential for modularity-based community detection. Modularity-based community analysis is performed in two separate phases: first, the detection of a meaningful community structure from a network and, second, the evaluation of the appropriateness of the detected community structure. Systematic deviations from a random configuration or from other null models without a characteristic modularity structure allow us to define a quantity called modularity, which is a measure of the quality of partitions. Newman and Girvan consider only the degree of nodes as a null model, which is equivalent to rewiring the network while preserving the degree sequence (Newman and Girvan 2004). This random model overlooks the economic nature of the network and thus modules. However, economic-based null models can connect these aspects: modularity-based community detection to find and explain communities where mobility exceeds the expected value. Economic null models predict the weight of arcs and, in this way, establish a baseline and explain network properties such as density, asymmetry or clustering. Therefore, it is worth connecting gravity and link prediction models. Based on link prediction, other properties of



the network, such as asymmetry, can be evaluated. In this study, we show how to match the network asymmetry and the revealed preference matrix; therefore, the application preference order can also be modeled using economic models.

Most studies on student mobility focus on international movement (see, e.g., Beine et al. 2014; Shields 2013), and very few studies have investigated mobility within a county (see an excerpt of Bacci and Bertaccini 2020). The reason for the low number of papers in this field is that it is very difficult to access a reliable database containing data on students, employees, and institutions, while several databases on international mobility networks are freely available (Gadár et al. 2020). However, an increasing number of countries, including Italy, Estonia, and Hungary, are registering applicants, and this allows for the investigation of the mobility network. To the best of our knowledge, Hungary is the first in the field of data integration because it has registered applications in a central database since 2001, and since 2011, this database has been integrated into the early career database, which was already an integrated database. In addition, this anonymous database is freely available to researchers.¹

Contribution to the literature

To the best of our knowledge, very few studies have attempted to combine gravity-like and network models (see several exceptions in Gadár et al. 2018; Bacci and Bertaccini 2020). However, they are not combined to explain mobility network formation. Although very few studies still consider economic model-based link prediction, network properties, such as asymmetry, are not modeled by economic inequities, and they are not used to estimate the revealed application preferences. In this study, aspects of describing young people's mobility are combined (see Fig. 1). The main result of this study is to apply economic gravity models for link prediction, which provides an explanation for why mobility is stronger between certain locations. We demonstrate the connection between revealed preferences and network asymmetries, and we show how the revealed application preferences can be explained by economics or rooted in spatial economic inequalities reflected by the gravity-based network asymmetries.

¹ <http://dimplomantul.hu>.

Methods

In section “[Data sources](#)”, the common data sources are introduced as indicated in Fig. 1. Then, in section “[Applied null models in mobility networks](#)”, the fundamental network properties are introduced, which will be evaluated for the network based on the economic gravity model, which is introduced in section “[Applied gravity models](#)”. Finally, in section “[Methods](#)”, we present the estimates of the inequities in the application preferences through network asymmetry; see section “[Modeling application preferences via asymmetries in the application mobility network](#)”.

Data sources

Several specific data sources are involved in the study. One of the main databases is the Hungarian central system for tracking graduates’ careers (HCSTGC). Similar tracking systems were recently developed by Estonia and Italy (Bacci and Bertaccini 2020; Kovacs and Kasza 2018). At present, there are few such systems, but we believe that a career path tracking system is the key data source for analyzing student mobility and the impact of higher education on society and the economy. Therefore, it would be worthwhile for decision makers to consider introducing such a system at least at the European Union level.

HCSTGC includes anonymized information on the location of residence (NUTS4 subregion), the city (or subregion) of the HEI, the county and subregion of the workplace, and starting salary of all graduated employees who earned their absolution or degree between 09/01/2014 and 01/31/2015. Among these individuals, this study focuses on those who were employed as of May 2016, representing 47,165 graduate students. The occupation and economic activity codes are also available for the work and workplace. The occupation coding uses the International Standard Classification of Occupations (ISCO) codes. The economic activity coding uses the International Standard Industrial Classification of all Economic Activities (ISIC) codes. Following Gadar and Abonyi (2018)’s work and databases, the first two letters of the occupation codes are matched to thirteen scientific fields of graduation; we called these fields *occupation categories*.

The next applied database is the student application database (2006–2017) which contains anonymized data from applicants and HEIs. We used the following data for the analysis: the location (NUTS4 subregion) of the applicant, the location of the HEI, the applied for BA/BSc/MA/MSc program and the scientific field of the program. To maintain consistency between the student application database and HCSTGC, henceforward, only the thirteen matches for occupation-scientific field category (occupation category) are considered: (1) agriculture, (2) human studies, (3) social sciences, (4) information technology, (5) law & public administration, (6) military, (7) business & economics, (8) engineering, (9) health & medical sciences, (10) pedagogy, (11) sport sciences, (12) natural sciences, and (13) arts. The aim of the matching was to ensure consistent nomenclature. In this study, only categories of the ISCO are considered, but to ensure consistency, nomenclatures of thirteen scientific fields are used as occupation categories ($O_k, k = 1, 2, \dots, 13$).

The last included data sources are already found in all countries in the European Union (such as Eurostat) and most countries in the world. The per-capita gross domestic

income (GDI/cap) of a location (i.e., subregion) between 2006 and 2017 comes from the Hungarian Central Statistical Office. From the Hungarian National Employment Service, we obtained the mean of the (gross) salary² for all 19 counties and the capital city (Budapest) for 2015 and 2016. These national (not only for recent graduates) salary statistics are available via ISIC/ISCO codes at the county level.

Applied null models in mobility networks

The mobility network can be described as a *directed graph* and is an ordered pair $G = (V, E)$ where V is a set of vertices (also called nodes, i.e., locations); $E \subseteq \{(x, y) | (x, y) \in V^2 \wedge x \neq y\}$ is a set of edges (also called arcs) that are ordered pairs of distinct vertices (i.e., an edge is associated with two distinct locations in a mobility graph). The number of movements between locations is associated with the edges. E is the adjacency matrix of graph G , where the elements of the matrix indicate whether pairs of vertices are adjacent in the graph.

Denote e_{ij} as the matrix element of adjacency matrix E of mobility graph G . The first null model that will be considered is the random configuration model that calculates the arc probabilities p_{ij}^{NG} , assuming a random graph conditioned to preserve the degree sequence of the original network:

$$p_{ij}^{\text{NG}} = \frac{id_i od_j}{L} \quad (1)$$

where id represents the in-degree and od the out-degree: $id_j = \sum_i e_{ij}$, $od_i = \sum_j e_{ij}$, L is the number of arcs (links) between nodes.

The distance-dependent (Liu and Murata 2010) version can also be used for null models.

$$p_{i,j}^{\alpha,\beta} = \frac{od_i^\alpha id_j^\beta}{f(d_{ij})} \quad (2)$$

where $p^{\alpha,\beta}$ is the distant-dependent null model. $f(d_{ij})$ is a monotone function of distance decay. The α, β parameters are called importance values estimated by regression analysis.

For the economic model we use the notation $q_{ij}^\Gamma = p_{ij}^{\alpha,\beta,\delta}$ with multiindex $\Gamma = \{\alpha, \beta, \delta\}$

$$q_{ij}^\Gamma = \gamma d_{ij}^\delta m_i^\alpha m_j^\beta \quad (3)$$

where m_i is an economic value, such as GDP, GDI or another economic quantity of the location of node i . d_{ij} is the distance between location i and location j . α, β, δ are importance values of locations and the distance between locations. They were estimated through a regression analysis. In this study, we propose a generalized gravity model for link prediction (null model):

² There is a flat tax of 33.5% in Hungary.

$$q_{ij}^{\Gamma} = \gamma d_{ij}^{\delta} \prod_{k=1}^N m_{i_k}^{\alpha_k} m_{j_k}^{\beta_k} \quad (4)$$

$$\log(q_{ij}^{\Gamma}) = \log(\gamma) + \delta \log(d_{ij}) + \sum_{k=1}^N (\alpha_k \log(m_{i_k}) + \beta_k \log(m_{j_k})) \quad (5)$$

where N is the number of economic parameters while $\alpha_k, \beta_k, \gamma, \delta$ are regression parameters.

The time-series version of the null model predicts the arcs of the dynamic network.

$$\log(q_{t,ij}^{\Gamma}) = \log(\gamma) + \delta \log(d_{ij}) + \sum_{k=1}^N (\alpha_k \log(m_{t,i_k}) + \beta_k \log(m_{t,j_k})) \quad (6)$$

Null models are mainly used in modularity-based community analysis. Originally, the method was specified for edges.

$$f(C) = (\text{fraction of arcs within communities}) - (\text{null model based expected fraction of such arcs}). \quad (7)$$

In the case of a directed network, this difference can be formulated as follows:

$$f(C) = \frac{1}{L} \sum_{i,j} (e_{ij} - p_{ij}) \delta(C_i, C_j) \quad (8)$$

p_{ij} represents the number of estimated arcs/weights proceeding from the i -th to the j -th location, and $\delta(C_i, C_j)$ is the Kronecker delta function that is equal to one if the i -th and j -th locations are assigned to the same community.

The goal of modularity-based community analysis is to separate the network into groups of nodes that have fewer connections between them than inside communities (Newman and Girvan 2004). The modularity of partition C can be calculated as the sum of the modularities of the $C_c, c = 1, \dots, n_c$ communities:

$$M_c = \frac{1}{L} \sum_{(i,j) \in C_c} (e_{ij} - p_{ij}). \quad (9)$$

Originally, these null models were specified for estimating arcs in a binary graph; however, they have been extended to handle weighted graphs, but null models can also be used to evaluate asymmetry.

The value of modularity M_c of a cluster C_c can be positive, negative or zero. Should it be equal to zero, then the community has as many links as the null model predicts. When modularity is positive, the C_c subgraph tends to be a community that exhibits a stronger degree of internal cohesion than the model predicts.

To obtain real communities, null models (p in Eq. (10)) should approximate a given property (λ), such as the probability of arcs, weights, reciprocity or edge asymmetry as much as possible. Therefore, when seeking null models, the following equation should be minimized:

$$\sum_{ij} |\lambda_{ij} - p_{ij}| = \epsilon \rightarrow \min \quad (10)$$

where λ is the modeled parameter, such as the arc/weights between node i and node j or the asymmetry of arc ij .

Measuring asymmetry: We consider a directed weighted network specified by the (nonnegative) weight matrix \mathbf{E} , where e_{ij} indicates the weight of the directed from node i to node j . In the case of no connection from i to j , $e_{ij} = 0$. \mathbf{E} can be specified as the sum of a symmetric (\mathbf{P}) and skewed symmetric matrix \mathbf{Q} , where

$$\mathbf{E} = (\mathbf{P} + \mathbf{Q}) \quad (11)$$

$$\mathbf{P} = \frac{1}{2} (\mathbf{E} + \mathbf{E}^T) \quad (12)$$

$$\mathbf{Q} = \frac{1}{2} (\mathbf{E} - \mathbf{E}^T) \quad (13)$$

The edge asymmetry matrix (\mathbf{A}) is the rate of skew-symmetric (\mathbf{Q}) and symmetric (\mathbf{P}) components of the weight matrix (\mathbf{W}).

$$\mathbf{A} = \frac{\mathbf{Q}}{\mathbf{P}}, \quad a_{ij} = \frac{e_{ij} - e_{ji}}{e_{ij} + e_{ji}} \in \mathbf{A} \quad (14)$$

where $a_{ij} = \frac{e_{ij} - e_{ji}}{e_{ij} + e_{ji}} \in \mathbf{A}$ is also a skewed symmetric matrix. For $\forall i \neq j$, where $e_{i,j} + e_{j,i} \neq 0$, edge asymmetry can only be interpreted for the interconnected i and j nodes. It satisfies that $-1 \leq a_{ij} = -a_{ji} \leq 1$ — in binary networks $a_{ij} \in \{-1, 0, 1\}$.

The economic null model for asymmetry (denoted as λ) can be expressed as follows:

$$\lambda_{ij}^{\Gamma} = \frac{q_{ij}^{\Gamma} - q_{ji}^{\Gamma}}{q_{ij}^{\Gamma} + q_{ji}^{\Gamma}} = \frac{\gamma d_{ij}^{\delta} m_i^{\alpha} m_j^{\beta} - \gamma d_{ij}^{\delta} m_j^{\alpha} m_i^{\beta}}{\gamma d_{ij}^{\delta} m_i^{\alpha} m_j^{\beta} + \gamma d_{ij}^{\delta} m_j^{\alpha} m_i^{\beta}} = \frac{m_i^{\alpha} m_j^{\beta} - m_j^{\alpha} m_i^{\beta}}{m_i^{\alpha} m_j^{\beta} + m_j^{\alpha} m_i^{\beta}} \quad (15)$$

Equation (15) also shows that while the economic version of the law of gravity is distance dependent (see Eq. (3)), by substituting it into Eq. (14), we obtain a distance-independent model. In addition, the economic null model of asymmetry only depends on the economic parameters (denoted as m) of locations and their importance factors (α, β). The indicator λ_{ij}^{Γ} has economic content itself. If $q_{i,j}^{\Gamma}$ estimates domestic immigration and $q_{j,i}^{\Gamma}$ estimates domestic emigration between location i and location j , then $q_{i,j}^{\Gamma} - q_{j,i}^{\Gamma}$ represents the net mobility exchange while $q_{i,j}^{\Gamma} + q_{j,i}^{\Gamma}$ represents the gross mobility exchange between location i and location j . If the arcs represent import/export goods, then the absolute value of λ_{ij}^{Γ} is a well-known intraindustry trade (IIT), while in this study, this value is called the intramobility rate (IMR), which characterizes the asymmetry of the mobility between locations.

Applied gravity models

According to the introduced data sources (see "The gravity model" section) we can identify and calculate the following parameters and coefficients (see Table 1):

Table 1 Notations and explanations

| Notation | Explanation |
|----------------------|--|
| L_j | Location (subregion at the NUTS level 4) L_j indexed by j |
| $C(L_j)$ | County of location L_j . |
| O_k | Occupation O_k (coded by ISCO) indexed by k . |
| Y_{ij} | Number of movements from location L_i to location L_j . |
| $Y_{ij k}$ | Number of movements from location L_i to location L_j , restricted only to the occupation O_k . |
| HEI_m | Higher education institution (HEI) m , indexed by m . |
| $HEI(L_j)$ | 1, if there is an HEI at location L_j . 0, otherwise. |
| $HEI_j k$ | 1, if there is an HEI that has a program for the expectation of occupation O_k . |
| L_m | Location (subregion at the NUTS level 4) of HEI_m |
| $RANK_{m,t}$ | Faculty excellence rank of HEI_m in year t (based on Eduline domestic HEI rank). |
| $A_{ij,t}$ | Number of applications from location L_i to HEI_j in year t |
| $A_{m,m',t}$ | The number applications where HEI_m is ahead of $HEI_{m'}$ on the preference list in year t |
| $I_j = GDI/cap(L_j)$ | Per-capita gross domestic income of location L_j , and I_m will be used for GDI/cap at the location of HEI_m . |
| $UR_j = UR(L_j)$ | Unemployment rate of location L_j . |
| E_n | Employee E_n , indexed by n |
| $S_{k,j,m}^g(n)$ | Starting salary of a graduated employee E_n working in occupation O_k at location L_j and graduated in HEI_m . |
| \bar{S}_k^n | Mean (starting) salary of undergraduate or not graduated employees in occupation O_k . |
| \bar{S}_k^g | Mean (starting) salary of graduated employees in occupation O_k . |
| $\bar{S}_{k,j}^g$ | Mean (starting) salary of graduated employees in occupation O_k at location L_j |
| $\bar{S}_{k,j,m}^g$ | Mean (starting) salary of graduated employees in occupation O_k at location L_j , graduated from HEI_m . |

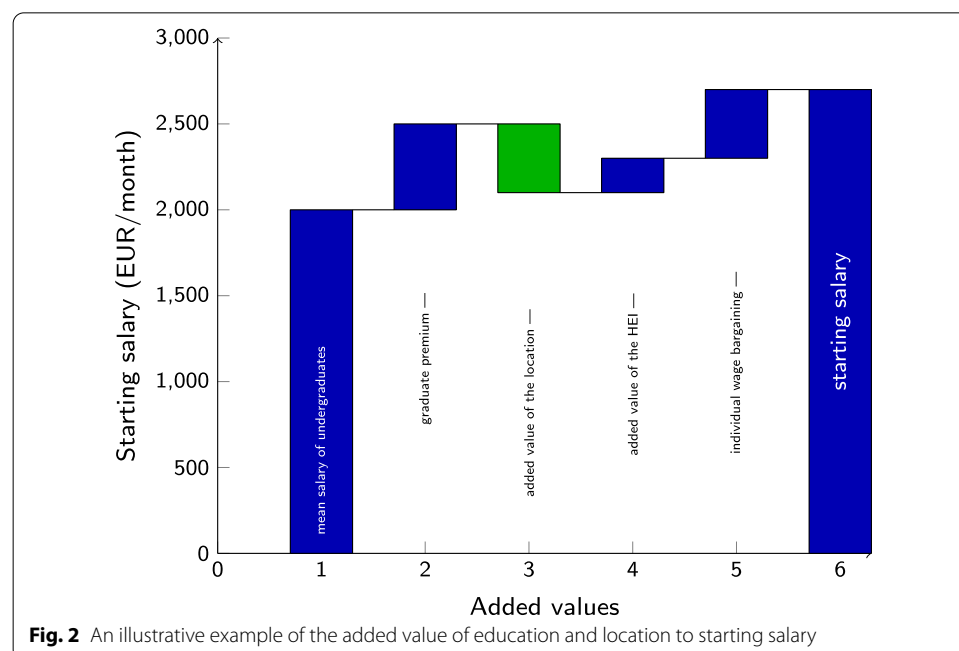
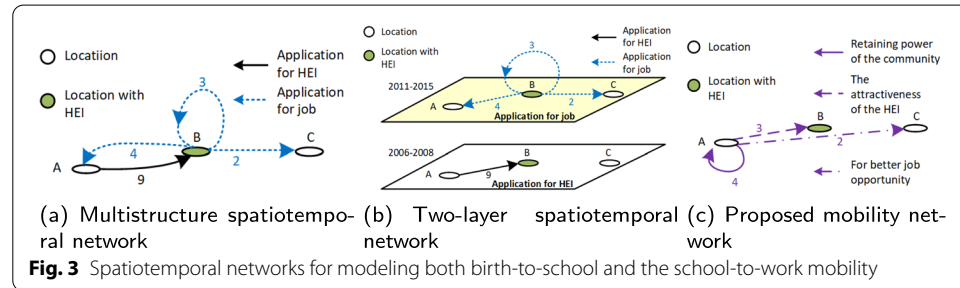


Table 2 Income differences and their meaning

| Differences | Meaning |
|--|--|
| $\Delta \bar{S}_k^g = \bar{S}_k^g - \bar{S}_k^n$ | Graduate premium |
| $\Delta \bar{S}_{k,j}^g = \bar{S}_{k,j}^g - \bar{S}_k^g$ | Added value of the location |
| $\Delta \bar{S}_{k,j,m}^g = \bar{S}_{k,j,m}^g - \bar{S}_{k,j}^g$ | The added value of HEI (HEI_m) at location (L_j) for occupation (O_k). |



The added value measured by the salary premium, which is the difference between the starting and mean salary of not graduated employees, can be specified by a sum of a graduate premium, the added value of the location, added value of the HEI and individual wage bargaining (see the waterfall plot in Fig. 2).

Figure 2 shows that the added value of a location can be negative; however, except for the mean salary of undergraduates, every factor can either be positive or negative (Table 2).

Factors of starting salary: After matching education-occupation pairs and the calculation of average salaries, the following differences can be specified:

Formally, the starting salary for employee n can be explained as follows (see Eq. (16)):

$$S_{k,j,m}^g(n) - \bar{S}_{k,j,m} = \Delta \bar{S}_k^g + \Delta \bar{S}_{k,j}^g + \Delta \bar{S}_{k,j,m}^g + \epsilon_n \quad (16)$$

where ϵ_n is the result of individual wage bargaining. Figure 2 explains how the salary premium ($S_{k,j,m}^g(n) - \bar{S}_{k,j,m}$) for employee n can be divided into four parts, namely, the graduate premium ($\Delta \bar{S}_k^g$), the added value of the location ($\Delta \bar{S}_{k,j}^g$), the added value of the HEI ($\Delta \bar{S}_{k,j,m}^g$), and the individual bargaining (ϵ_n) as described in Eq. (16).

These values serving as proxies for the added value to salaries will be measured, and they will be applied as dependent variables in the proposed spatiotemporal model.

Two kinds of movement can be specified based on the Hungarian career tracking database: (1) application to an HEI (see the solid arrows in Fig. 3a) and (2) application to a job (see the dotted arrows in Fig. 3a). These movements can be separated (see Fig. 3b) into two distinct layers. Two kinds of mobility, occupation mobility (Y_{L_i,L_j}) and application mobility ($A_{L_i,HEI_j,t}$), can be explored. In addition, based on the application databases, application mobility can be modeled over the period 2006–2017 with a fixed effects gravity model. Occupation mobility is also separated in time (see Fig. 3b). One of the problems is that these movements differ over time. Therefore, it

is a so-called two-layer spatiotemporal network. Nevertheless, if we only consider the movement between the location of residence before the application to an HEI and the location of a workplace, we can propose a so-called **mobility network** (see Fig. 3c). In this case, the edges between nodes connect different locations not only in space but also in time.

There are two ways to understand the proposed mobility network. We can regard the mobility network as a spatiotemporal mobility network where the communities represent the “attractiveness” of a location. The sizes of communities can differ, regardless of whether an HEIs is present. Counties that typically represent sources and sinks for different occupations can be specified. The role of an HEI in becoming a source or sink county can also be analyzed.

The other understanding is to regard mobility as a flow (of people) between locations. Gravity models make it possible to model the “attractiveness” of a location, such as the strength of the regional economy (GDP/cap, GDI/cap), salary opportunities, the distance dependencies between locations for every single occupation, and the role of institutions.

The logarithmic form of the proposed spatiotemporal gravity model can be specified as follows for occupation mobility:

$$\begin{aligned}\log Y_{j,j'} = & \beta_0 + \beta_1 \log I_{j,t_j} + \beta_2 \log I_{j',t_{j'}} + \\ & + \beta_3 \log UR_{j,t_j} + \beta_4 \log UR_{j',t_{j'}} + \beta_5 \log d_{j,j'} + \\ & + \beta_6 HEI(L_j) + \beta_7 HEI(L_{j'}) + u_{j,j'}\end{aligned}\quad (17)$$

where $u_{j,j'}$ is the residue of the regression model. β s are the regression parameters to be estimated. Note that the years of the data values $I_{j,t_j} = GDI/cap(L_j, t_j)$ and $I_{j',t_{j'}} = GDI/cap(L_{j'}, t_{j'})$ are usually different. When considering the source location (L_j), the $GDI/cap(L_j, t_j)$ should be from the year of application, while when considering the target ($L_{j'}$) (workplace), the $GDI/cap(L_{j'}, t_{j'})$ should be from the year of hiring at the workplace.

Using dummy variables, the role of HEIs is included in Eq. (17). GDI/cap values are proxies for the expected salaries; therefore, if the salary surplus (as an added value) of the location for an occupation O_k is considered, then a detailed model can be specified as follows:

$$\begin{aligned}\log Y_{j,j'}|_k = & \beta_0 + \beta_1 \log \bar{S}_{k,j}^g + \beta_2 \log \bar{S}_{k,j'}^g + \beta_3 \log UR_{j,t_j} + \\ & + \beta_4 \log UR_{j',t_{j'}} + \beta_5 \log d_{j,j'} + \beta_6 HEI_j|_k + \beta_7 HEI_{j'}|_k + \\ & + u_{j,j'}|_k\end{aligned}\quad (18)$$

Application mobility can be explored with a time-series (fixed effects gravity) model. In this way, similar indicators can be specified.

$$\begin{aligned}\log A_{j,m,t} = & \beta_0 + \beta_1 \log I_{j,t} + \beta_2 \log I_{m,t} + \\ & + \beta_3 \log UR_{m,t} + \beta_4 \log UR_{j,t} + \beta_5 \log d_{j,m} + \\ & + \beta_6 RANK_{m,t} + u_{j,m,t}\end{aligned}\quad (19)$$

Table 3 Connections between null models and economic gravity models

| Network | Static or dynamic | Null model | Economic model |
|------------------------------|-------------------|---------------|----------------------|
| Occupation mobility network | Static | Equation (5) | Equations. (17)–(18) |
| Application mobility network | Dynamic | Equation (6) | Equation (19) |
| Preference graph | Dynamic | Equation (15) | Equation (20) |

Table 4 The unordered aggregated preference and asymmetry matrices

| V | l_1 | l_2 | l_3 | l_4 |
|-----------------------------------|-------|-------|-------|-------|
| (a) Preference matrix | | | | |
| l_1 | 0,0 | 1,0 | 2,0 | 3,0 |
| l_2 | 2,0 | 0,0 | 2,0 | 3,0 |
| l_3 | 1,0 | 1,0 | 0,0 | 3,0 |
| l_4 | 0,0 | 0,0 | 0,0 | 0,0 |
| \sum | 3,0 | 2,0 | 4,0 | 9,0 |
| N^o | 2 | 1 | 3 | 4 |
| Opposite preferences=2+1+1=4 | | | | |
| A | l_1 | l_2 | l_3 | l_4 |
| (b) Asymmetry matrix | | | | |
| l_1 | 0,0 | −1/3 | 1/3 | 1,0 |
| l_2 | 1/3 | 0,0 | 1/3 | 1,0 |
| l_3 | −1/3 | −1/3 | 0,0 | 1,0 |
| l_4 | −1,0 | −1,0 | −1,0 | 0,0 |
| \sum | −1,0 | −5/3 | 1/3 | 3,0 |
| N^o | 2 | 1 | 3 | 4 |
| Sum of the lower triangular=−10/3 | | | | |

We linked the gravity models (Eqs. (17)–(19)) and the linked prediction models (see Eqs. (5)–(6)).

Owing to the time-series data, the application mobility can be explored using time-series (i.e., fixed effects gravity) models. Therefore, the dynamic null model predicts the dynamic mobility graph.

Table 3 reports which null models are estimated with economic models.

Modeling application preferences via asymmetries in the application mobility network

The Hungarian Education Authority has made all application data available to researchers. We considered the interval between 2006 and 2017. According to (Telcs et al. 2016), individual application preferences can be aggregated, and in this way, professional, regional, and institutional aggregated preference matrices can be calculated. Here, an (i, j) cell from the aggregated preference matrix shows how many times the i -th institution preceded the j -th institution in student applications. Telcs et al. (2016) offered several heuristic methods to estimate aggregated preference orders from an aggregated preference matrix. They showed that these methods approximate the optimal preference order very well, where the optimal preference order is a preference order and the number of opposite application preferences (i.e., the sum of values in the lower triangle of the

Table 5 The ordered aggregated preference and asymmetry matrices

| V | l_2 | l_1 | l_3 | l_4 |
|----------------------------------|-------|-------|-------|-------|
| (a) Preference matrix | | | | |
| l_2 | 0,0 | 2,0 | 2,0 | 3,0 |
| l_1 | 1,0 | 0,0 | 2,0 | 3,0 |
| l_3 | 1,0 | 1,0 | 0,0 | 3,0 |
| l_4 | 0,0 | 0,0 | 0,0 | 0,0 |
| \sum | 2,0 | 3,0 | 4,0 | 9,0 |
| N^o | 1 | 2 | 3 | 4 |
| Opposite preferences (E)=1+1+1=3 | | | | |
| A | l_2 | l_1 | l_3 | l_4 |
| (b) Asymmetry matrix | | | | |
| l_2 | 0,0 | 1/3 | 1/3 | 1,0 |
| l_1 | -1/3 | 0,0 | 1/3 | 1,0 |
| l_3 | -1/3 | -1/3 | 0,0 | 1,0 |
| l_4 | -1,0 | -1,0 | -1,0 | 0,0 |
| \sum | -5/3 | -1,0 | 1/3 | 3,0 |
| N^o | 1 | 2 | 3 | 4 |
| Sum of the lower triangular=-4 | | | | |

preference matrix) is minimal (Telcs et al. 2016). One possible choice is the column sum method.

The aggregated revealed preference is the order of the column sums (see Table 4a). Since Eq. (14) is a monotonous transformation, the order of the column sum is not modified (see Table 4b), the asymmetry is not perturbed, and therefore, we gain a model for the preference order (see Eq. (20)).

Table 4 provides an example of an (unordered) aggregated preference matrix.

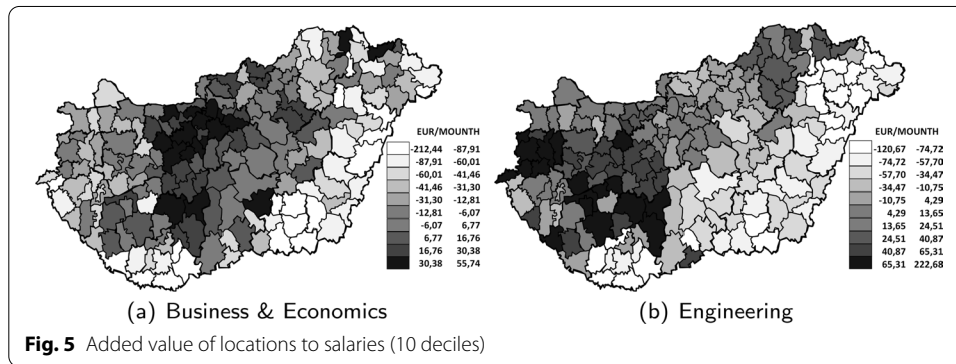
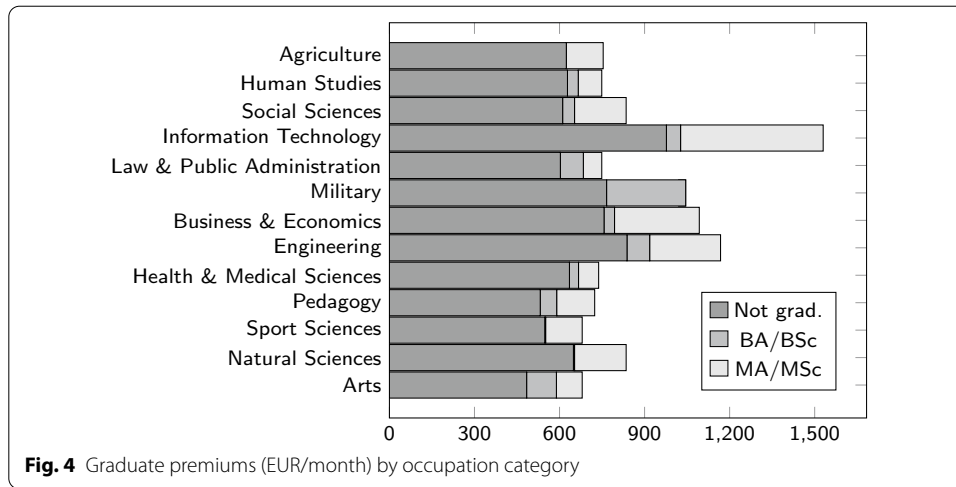
Telcs et al. (2016) showed that if the matrix is (re)ordered by the preference orders, then the opposite preferences (the sum of the values in the lower triangle) can be decreased. The algorithm will stop if the (re)orderings of the preference matrix cannot reduce the sum of opposite preferences (see Table 5).

Equation (20) shows an example of the power of combining economic models and network science. Equation (20) explains the preference value between institutions HEI_m and $HEI_{m'}$ in year t . If these models fit well, then the institutional preference order can also be modeled and explained.

Equation 20 defines the asymmetry between HEIs m and m' . A generalized gravity model is used to express the attractiveness of the HEIs to express the asymmetry of the nodes in the network representing the HEIs.

$$A_{m,m',t} = \frac{I_m^{\alpha_1} UR_m^{\alpha_2} RANK_j^{\alpha_3} - I_{m'}^{\beta_1} UR_{m'}^{\beta_2} RANK_{m'}^{\beta_3}}{I_m^{\alpha_1} UR_m^{\alpha_2} RANK_m^{\alpha_3} + I_{m'}^{\beta_1} UR_{m'}^{\beta_2} RANK_{m'}^{\beta_3}} \quad (20)$$

For shorter notation here, we used I_m instead of $GDI/cap(L_{HEI_m})$ with a little abuse of the original notation.



Results

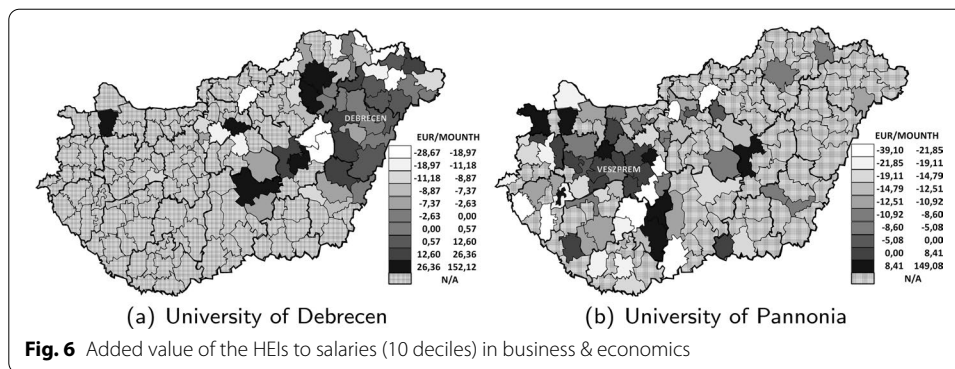
In this section, we show how the location and the location's economic state influence the students mobility, career path and how it can be inferred by our models.

Added value of the locations and education

Following Eq. (16), the salary premium can be decomposed into four components. The first is the graduate premium $\Delta \bar{S}_k^g$, which is the difference between the mean starting salary of graduated employees (\bar{S}_k^g) and the mean starting salary of not graduated employees

(\bar{S}_k^h) in occupation category k . Figure 4 shows the graduate premiums by occupation category k . The highest mean salaries are in information technology, engineering and business & economics. The highest added value of a master's diploma is also in these three categories.

The second factor is the added value of the location. Figure 5 shows the added value of the location for business & economics (see Fig. 5a) and engineering (see Fig. 5b). The added value, which can be positive or negative, is ordered into ten deciles. In the case of business and economics, only the seventh decile contains positive values, and these deciles are concentrated in the center of Hungary, while in the case of engineering,



added value is concentrated in the western part of the country. These results are highly correlated with the spatial distributions of the companies. While the center of economics is concentrated in the capital city, the companies requiring engineers are concentrated in the more industrialized western part of Hungary.

The third factor is the added value of HEIs. The role of HEIs in spatial mobility has already been shown in Fig. 6. If an employee graduated in the eastern part of Hungary (i.e., the University of Debrecen, see Fig. 6a) then he/she usually does not go to work in to another part of the country, and vice versa. For example if employees graduated in the western part of the country, for example from the University of Pannonia (see Fig. 6b), which has several campuses in the western part of Hungary, then they usually do not work in the eastern part of the country. The added value of rural universities is positive or neutral in the locations and neighboring locations of the campuses; however, it is usually negative in the capital (Budapest) and varies substantially in more distant subregions.

The remaining factor is the individual bargaining on salary, which is the unexplained (residual) part of the economic model in Eq. (16). This value follows the normal distribution.

Analyzing spatial mobility to reflect the role of HEIs

In this section, application and occupation mobility are investigated. First, traditional gravity models are applied to determine the roles of economic and institutional indicators of mobility. Then, a mobility network is built and the main network properties are examined. In the null models, parameters, such as link weights, density and asymmetry, are predicted by the unified economic-network model. Then, the preference orders of institutes and subregions are explained by the proposed economic-network model.

Analyzing application mobility: the first step

Applying gravity models: With Eq. (19), the primary preferences can be analyzed within the period 2011–2017. We used a fixed effects gravity model. Table 6 shows the results.

$I = GDI/cap$ is only a proxy for the cost of living and expected salaries. It is not surprising that the most important value is the GDI/cap in the subregion of the HEI. Nevertheless, the second most important value is the GDI/cap in the subregion of residence

Table 6 Results of the gravity model for all first-place applications to higher education institutions (2011–2017)

| Variables | Coefficients | 2011 | 2013 | 2014 | 2015 | 2016 | 2017 | Panel model (2011–2017) | Relative importance |
|--------------|--------------|--------|--------|--------|--------|--------|--------|----------------------------|------------------------|
| $l_{j,t}$ | β_1 | 2.038 | 1.671 | 2.798 | 1.081 | 1.367 | 0.716 | 1.462 | 21.27% |
| $l_{m,t}$ | β_2 | 2.938 | 2.719 | 2.882 | 2.974 | 2.442 | 2.828 | 2.814 | 39.82% |
| $UR_{j,t}$ | β_3 | 0.099 | 0.677 | 0.509 | 1.126 | 0.750 | 1.141 | 0.808 | 10.15% |
| $UR_{m,t}$ | β_4 | −0.448 | −0.192 | −0.430 | −0.580 | −0.678 | −1.814 | −1.257 | 16.34% |
| $d_{j,m}$ | β_5 | −0.457 | −0.596 | −0.619 | −0.688 | −0.701 | −0.724 | −0.667 | 9.20% |
| $RANK_{m,t}$ | β_6 | −0.428 | −0.386 | −0.278 | −0.138 | −0.107 | −0.095 | −.241 | 3.22% |
| | R^2 | 0.872 | 0.847 | 0.729 | 0.682 | 0.539 | 0.730 | 0.663 | 100.00% |
| | \bar{R}^2 | 0.871 | 0.846 | 0.728 | 0.682 | 0.539 | 0.730 | 0.663 | 100.00% |

of the applicants, which indicates that the economic properties of the sending subregion also play an important role in being accepted by an HEI.

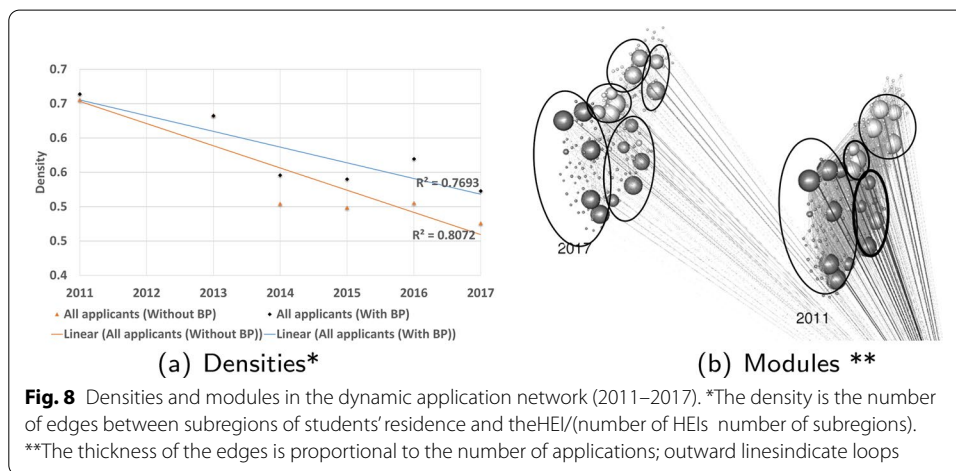
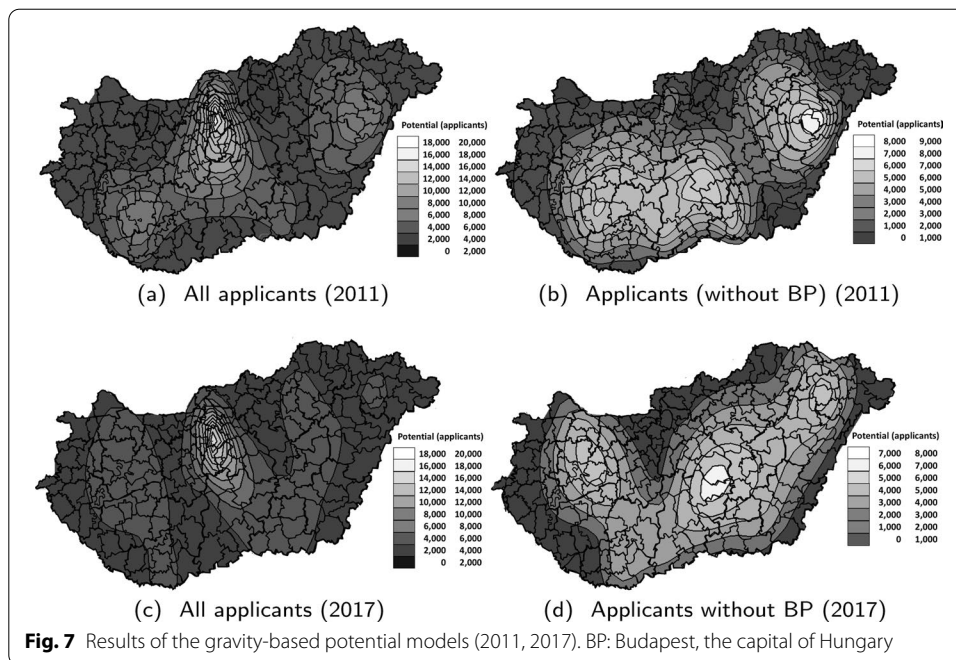
The unemployment rate (*UR*) of the subregion of the place of residence has a positive effect while the unemployment rate of the subregion of the university has a negative effect on applications. This means that there are more applicants from subregions with higher unemployment rates but that fewer apply to institutions in subregions where this indicator is also high. In the national ranking of Hungarian HEIs, the better HEIs are linked to fewer applicants; therefore, having a negative value indicates that more reputable institutions enroll students.

Note that the role of institutional reputation is decreasing each year. The negative coefficient of distance is also in line with the gravity models. The larger the distance is, the higher the travel costs, resulting in a greater financial burden for parents and students. In addition, between 2011 and 2016, this coefficient linked to physical distance increased, which indicates a decrease in mobility since the geographical distances did not change.

Telcs et al. (2015) suggested that gravity models should be used without HEIs in capital cities because the high centralization of institutes may distort the results.

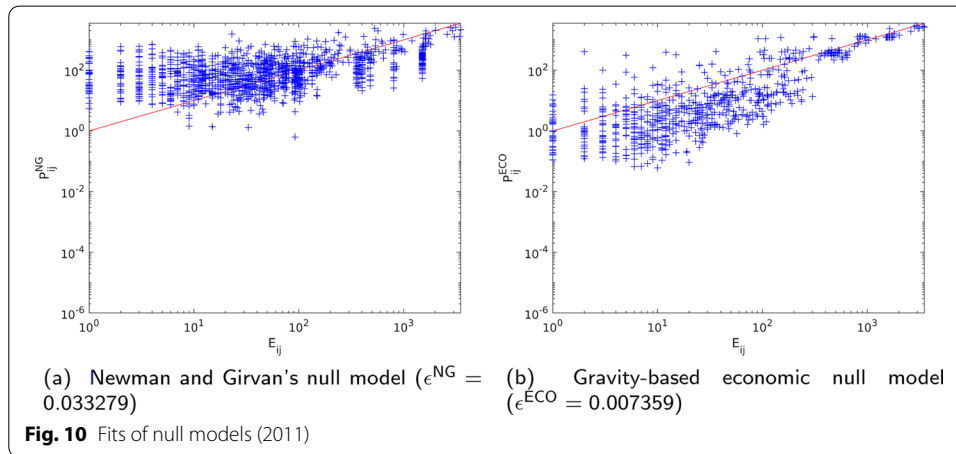
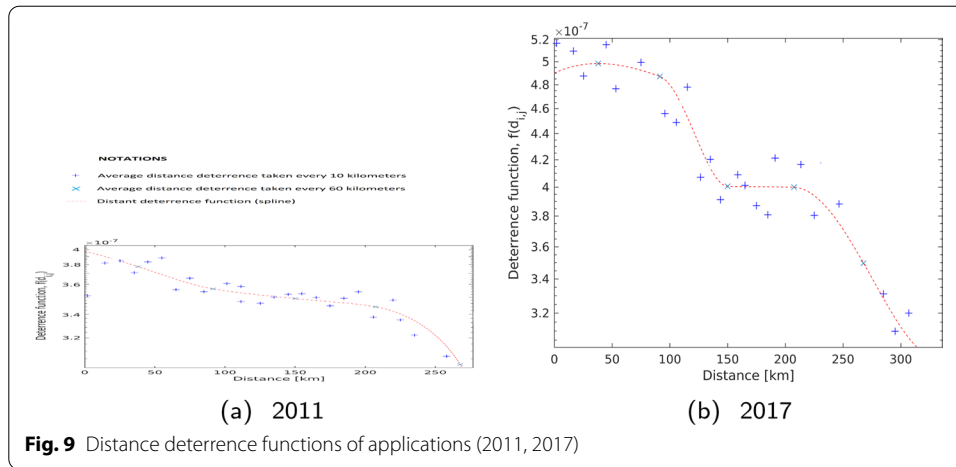
The gravity-based potential model shows that the role of Budapest (BP) is increasing. The strength of Budapest-centric HEIs is reflected in the greater difference in potential values between Budapest and rural HEIs (see Fig. 7).

Applying network science models: Network analysis can not only confirm the results of the gravity model but also offers additional insights. One of the fundamental dynamic network indicators is the change in network density over time (see Fig. 8a) and the change in the structure of modules (see Fig. 8b). Figure 8a shows that between 2011 and 2013, the density of the mobility graphs was almost equal regardless of whether one includes or excludes the HEIs in Budapest. The linear trends show that the density is increasing, which indicates that fewer locations are connected. Furthermore, not only has the number of applicants decreased, but students have applied to HEIs from fewer locations. The differences in the slope of the linear trend indicate that the decrease in mobility is greater if HEIs in Budapest are excluded. In 2017, there were 27% fewer applicants to HEIs than in 2011. This causes the lines to be thinner in Fig. 8b. The module-based community analysis identified similar catchment areas as the potential analysis.



However, it indicates that the Eastern part is more fragmented. In addition, the decreasing density (with the exception of 2016) indicates that fewer students applied to HEIs and that they applied from fewer places.

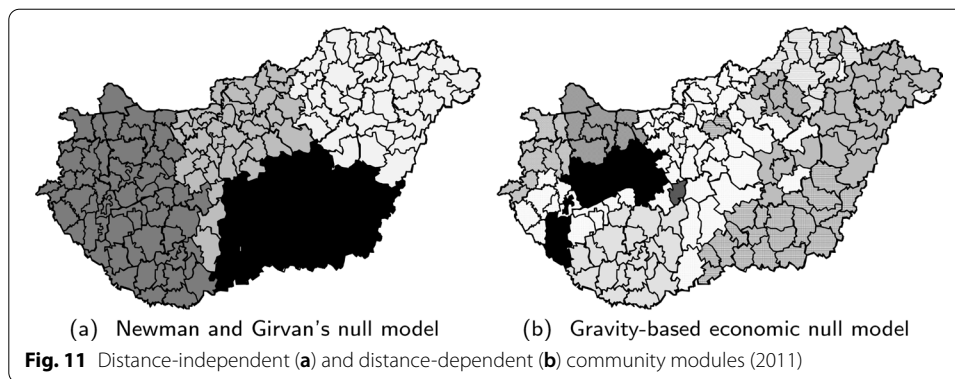
The results of the fixed panel model show the increasing roles of the distance between the subregion of the student's residence and the subregion of the HEI (see Table 6). The higher values of the distance deterrence function at low distances in 2017 also reflect these results; nevertheless, the distance deterrence function shows the nature of the distance distribution between the student's residences and the locations of HEIs. The maximum of all deterrence functions is in the low distance between locations, which indicates the retention role of HEIs. In both curves, a fraction can be seen showing that students are welcome to apply to institutions in the center of the



country (Budapest) but are far less welcome at either end of the country. This fraction is stronger in 2017. Figure 9 shows the distance deterrence function calculated by Eq. (2). These spline curves indicate different shapes of distance deterrence comparing 2011 and 2017. The higher values at low distances in 2017 show that the role of distance is increasing. Students apply to closer HEIs. Since Budapest is in the middle of Hungary, the distance deterrence function becomes fractioned.

The combination of gravity and network science models generates a lower error value for the null models. These null models can be used to specify distance-based communities (see Fig. 11), which better reflect students' preferences and the catchment areas of HEIs. Therefore, the asymmetry of the application network can also be modeled (see Eq. (15)). The application preferences can be modeled using see Eq. (19), the asymmetry prediction.

Figure 10 shows the fits of the null models. $e_{j,m} \in \mathbf{E}$ represents the applications from location j to location of HEI_m , while matrix \mathbf{P} contains the estimated values (see Eq. (1) based on Newman and Girvan's null model and Eqs. (4)–(5) for the economic null

**Table 7** Estimation of application preference via network asymmetry (2011–2017)

| Terms | Coefficients | 2011 | 2013 | 2014 | 2015 | 2016 | 2017 |
|--------------------------|-------------------------------|--------|--------|--------|--------|--------|--------|
| $l_{m,t}/l_{m',t}$ | $\tau_1 = \alpha_1 - \beta_1$ | 0.056 | 0.048 | 0.042 | 0.016 | 0.027 | 0.105 |
| $UR_{m,t}/UR_{m',t}$ | $\tau_2 = \alpha_2 - \beta_2$ | −0.043 | −0.035 | −0.031 | −0.005 | −0.019 | −0.026 |
| $RANK_{m,t}/RANK_{m',t}$ | $\tau_3 = \alpha_3 - \beta_3$ | −0.106 | −0.081 | −0.071 | −0.062 | −0.042 | −0.047 |
| | R^2 | 0.606 | 0.528 | 0.478 | 0.476 | 0.438 | 0.485 |
| | \bar{R}^2 | 0.606 | 0.528 | 0.478 | 0.475 | 0.437 | 0.485 |

model). In this calculation, the economic null model (Eq. (5)) uses the coefficients from the application mobility gravity model (see Eq. (19)).

In this way, the gravity model and the economic null model are matched. First, the matched model is used for community detection (see Eq. (9)). Figure 11a shows that without considering distances, four modules can be defined.³ The economic-gravity-model-based communities (see Fig. 11b) better reflect the catchment areas of HEIs. Note that Newman and Girvan's null model is a distance-independent null model; therefore, if subregions are geographically connected, such as in Fig. 11a, then this shows that distance is an important factor. It does not neglect from the null model (Gádár et al. 2018). The economic null model is already a distance-dependent null model where the modules are more fragmented; see Fig. 11b. Nevertheless, we can see, for example, that an economic-based null model that is simultaneously a gravity model explains mobility better (see Table 6) than the module that provides catchment areas of HEIs (compare Figs. 7 and 11b).

The other approach is to combine gravity-based economic models and null models from network science, and the revealed preferences are used to model asymmetry with economic null models (see Eq. (15)). The first advantage of this model is that it is distant independent (see Eq. (15)); therefore, only differences between the importance values are estimated and analyzed.

Table 7 shows the parameter estimation of Eq. (20). The results of the gravity models in Table 6 demonstrate that a more preferred HEI that has a higher number of applicants has a lower unemployment rate, higher per-capita GDI in its region and better

³ Compare Fig. 8b as a graph representation and Fig. 11a as a spatial representation of distance-independent community based modules.

Table 8 Estimation of application preference (2011)

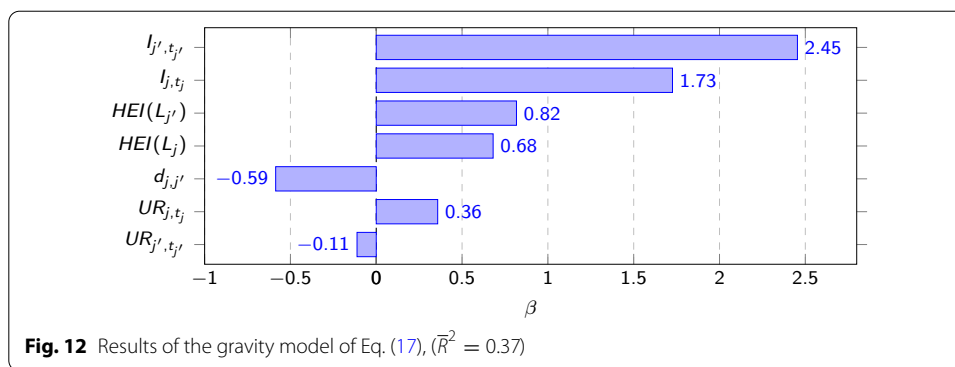
| Appl. pref. order | Estimated order | m=HEI | $I_m = GDI/cap_m$ (EUR) | UR_m (%) | $RANK_m$ | Number of 1 st order appl. | Estimation of 1 st ord. appl. | All applications |
|-------------------|-----------------|-------|-------------------------|------------|----------|---------------------------------------|--|------------------|
| 1 | 1 | ELTE | 5513 | 4.12 | 1 | 15,339 | 13,461 | 44,213 |
| 2 | 3 | DE | 4407 | 9.10 | 3 | 12,552 | 10,437 | 39,965 |
| 3 | 2 | SZTE | 4366 | 6.18 | 2 | 11,953 | 10,653 | 37,603 |
| 4 | 6 | PTE | 4443 | 8.03 | 6 | 10,803 | 9,212 | 34,045 |
| 5 | 7 | BCE | 5513 | 4.12 | 8 | 9605 | 8,263 | 28,383 |
| 6 | 18 | BGF | 5513 | 4.12 | 41 | 7388 | 5,778 | 28,274 |
| 7 | 10 | SZIE | 4848 | 4.67 | 13 | 8056 | 7,583 | 26,884 |
| 8 | 9 | BME | 5513 | 4.12 | 11 | 8580 | 6,858 | 23,165 |
| 9 | 5 | NYME | 3980 | 2.07 | 3 | 6582 | 6,232 | 19,524 |
| 10 | 14 | ME | 4241 | 11.94 | 9 | 5308 | 5,005 | 17,189 |

rank position (smaller rank order value) on faculty excellence. The results show that an HEI is more preferred if the prospects of living there are more favorable (higher per-capita GDI, lower unemployment rate) and there are better HEIs, which is in line with the results of the gravity models (see Table 7). Nevertheless, this model also shows that the differences in the importance of the faculty excellence of HEIs are increasingly being reduced.

By examining the coefficients over time, we can see that the determination value (R^2) and the adjusted determination value (\bar{R}^2), with the exception of those of 2017, decrease between 2011 and 2016. This means that the importance of other, unmodeled parameters is increasing. The values of all coefficients are the same as in the gravity model. Nevertheless, in the gravity models, the decline in coefficients indicated that they were playing increasingly less of a role in the top preference of applicants. Here, however, a decrease in values indicates that the importance of coefficient values for institutions located in different places in diminishing order of preference is being equalized.

Where a good fit is found for asymmetry, it is worth comparing the real and modeled values. In this case, we can answer the question of how well the preference ordering formed on the basis of the application and the ordering modeled on the basis of economic, unemployment, and faculty excellence data correlate with one another (see Table 8). As in section "[Modeling application preferences via asymmetries in the application mobility network](#)", where we showed how we can restore the aggregate preference matrix from the asymmetry matrix by knowing the total number of students, we can estimate the preference order from the estimation of the asymmetry matrix; therefore, all applications and, more important for institutions, all first-place applications can be predicted. The policymakers of a given HEI can gain particular insight into their comparative advantages (disadvantages) by examining the factors that determine the asymmetry between them and their primary competitors.

Table 8 shows the preference order of the first 10 institutions (see Telcs et al. 2016), which follows most of *all applications*. The Spearman rank correlation of the modeled and real preference order is 0.61. To understand why the proposed method ranked certain institutions lower or higher, we report the per-capita GDI and the unemployment



rate in the subregions of the HEIs and display the rank positions of the faculty excellence values.

Budapest Business School: BGF (now BGE) is ranked lower in the model because its faculty excellence ($RANK_m$) is significantly lower than that of other listed institutions. Based on the results presented above, although faculty excellence is an important factor, there are institutions that are popular, despite that this value is lower for them. The University of Miskolc (ME) would not be among the top ten institutions according to the model. The reason for this is that in the subregion of Miskolc, unemployment was very high in 2011, at above 10%. The University of West Hungary (NYME), which no longer exists, would be ranked higher by the model, despite the lower per-capita GDI (I_m); however, there is a low unemployment rate (UR_m), and it is third place in faculty excellence ($RANK_m$). We can see that the model underestimates the number of applications for the first 10 institutions and overestimates them at the end of the ranking.

Analyzing occupation mobility: the second step

First, Eq. (17) is applied to measure the role of HEIs in occupational mobility. Figure 12 shows the results of the gravity model where all parameters are significant, and the parameters are arranged in descending ordered by the absolute β values.

Not surprisingly, the highest coefficient is the $I = GDI/cap$ at the host location (j') which is a proxy for expected salaries. The coefficient of GDI/cap of the source location j is also positive, which resonates with the results of the applied gravity models. A high GDI/cap in the source location creates a chance for the student to attend to university and later to remain at that location or go to another place to secure employment.

It is a very interesting and important result that just behind the income data, the importance of two dummy variables (whether there is an HEI in host j' and the source j location) appears ahead of the unemployment rates and distance. Mobility is facilitated both by having an HEI close to the workplace and even more so by having such an institution close to the source location. It is no coincidence that multinational companies prefer to settle around university cities, as these institutions play a significant role in both attracting students and keeping them there after graduation. Due to their importance in innovation, HEIs can have a positive impact on the social, economic and cultural development of a given region. This is because due to the continuous decrease in public expenditures, some HEIs are turning to the local public and business sector

Table 9 Variable importance (in %) of occupation mobility model

| Occupation categories | \bar{S}_{kj}^g | $\bar{S}_{kj'}^g$ | $UR_{j,tj}$ | $UR_{j',t,j'}$ | $d_{jj'}$ | $HEI(L_j)$ | $HEI(L_{j'})$ | \bar{R}^2 |
|-----------------------------|------------------|-------------------|-------------|----------------|-----------|------------|---------------|-------------|
| Agriculture | 3 | 1 | 1 | 3 | 16 | 13 | 63 | 0.459 |
| Human Studies | 2 | 4 | 1 | 4 | 22 | 3 | 64 | 0.429 |
| Social Sciences | 8 | 1 | | 3 | 14 | 6 | 68 | 0.468 |
| Information Technology | 9 | | | | 15 | 10 | 66 | 0.462 |
| Law & Public Administration | 14 | | | 3 | 15 | 20 | 48 | 0.421 |
| Military | 9 | 2 | | | 14 | 1 | 74 | 0.629 |
| Business & Economics | 5 | 4 | | 3 | 16 | 6 | 66 | 0.487 |
| Engineering | 4 | 2 | | 1 | 15 | 4 | 74 | 0.503 |
| Health & Medical Sciences | 7 | 1 | | 6 | 26 | 10 | 50 | 0.465 |
| Pedagogy | 4 | 3 | | 5 | 11 | 6 | 71 | 0.486 |
| Sport Sciences | 7 | 2 | 4 | | 33 | 22 | 32 | 0.520 |
| Natural Sciences | 4 | | | 2 | 26 | 7 | 61 | 0.554 |
| Arts | 12 | | 2 | | 27 | 2 | 57 | 0.509 |
| All | 1 | 3 | 1 | 4 | 7 | 5 | 79 | 0.445 |

and attempting to recruit more students from the immediate environment, as well as to increase revenues by providing professional services. The retention of students and the regional involvement of HEIs, which provide space and audiences for cultural events, are indisputable, in addition to professional relations with the business sector. Mobility, albeit to a lesser extent, is positively affected if the unemployment is high in the source location; however, the attractiveness of the host location is slightly decreased if the unemployment rate in the host location is high. The coefficient of the unemployment rate is lower than the coefficient of distance, which is, not surprisingly, a negative value.

If instead of GDI/cap, the mean of local starting salaries in an occupation is considered, then the adjusted determination coefficient (\bar{R}^2) can be increased (see Table 9). Table 9 shows the variable importance (measured by the contribution to the determination coefficient) instead of betas. Since the *HEI* variables are dummy variables, they better reflect the real role of HEIs.

The results show that except for sport sciences, the most important value for occupational mobility is whether there are any HEIs close to the workplace. These results match well with the results on the added value of the HEIs, where Figure 6 shows that students attempt to find jobs close to the university where they graduate. Furthermore, larger companies prefer to settle around cities where HEIs operate.

By combining network science methods with gravity models, the impact of HEIs can be further investigated. The asymmetry in occupation mobility can be calculated for the subregions, and they can be ranked. Table 10 shows the top 10 most attractive subregions.

In line with the former results, it is not surprising that, first, ten subregions have (at least one) HEIs. This accords with the former results and highlights the role of HEIs. Only the first two subregions have a positive balance of graduated employees. The first is the capital city of Hungary, and the second is also located next to Budapest. The other subregion already has a negative balance. If only all people's mobility is considered is

Table.10 Top 10 most attractive subregions for freshly graduated employees (2015). (Subregions are ranked by their asymmetry values)

| Rank | Subregion | Is there any HEI? | Incoming Graduated* | Outgoing | Incoming All* | Outgoing |
|------|-----------------|-------------------|---------------------|----------|---------------|----------|
| | | | Employees (2015) | | People (2015) | |
| 1 | Budapest | Yes | 17,656 | 4163 | 68,672 | 66,263 |
| 2 | Gödöllői | Yes | 2148 | 300 | 3624 | 3305 |
| 3 | Győri | Yes | 742 | 778 | 6348 | 5045 |
| 4 | Székesfehérvári | Yes | 619 | 887 | 4850 | 4389 |
| 5 | Veszprémi | Yes | 407 | 579 | 3727 | 3679 |
| 6 | Kecskeméti | Yes | 439 | 642 | 4369 | 4051 |
| 7 | Tatabányai | Yes | 348 | 350 | 2806 | 2679 |
| 8 | Szegedi | Yes | 708 | 1140 | 5976 | 5204 |
| 9 | Pécsi | Yes | 549 | 1096 | 5510 | 5150 |
| 10 | Egri | Yes | 407 | 723 | 3107 | 2620 |

* without loops, it means that the source and the host location must be different

this balance positive in the top subregions. Nevertheless, the lack of freshly graduated employees can indicate that the structure of society may change over time.

Discussion

Data-driven career tracking offers new insights for scholars to analyze application and occupation mobility. This database represents the whole population of graduated employees and applicants; therefore, more reliable models can be proposed. This database is also important for potential applicants when choosing an HEI. A pilot web page is already accessible at <https://www.diplomantul.hu>, where potential applicants can see the potential salaries by occupation categories and the added value of HEIs (see Figs. 5 and 6). Nevertheless, future research should examine how the availability of this information reinforces the asymmetry that already exists in mobility networks. Modeling asymmetry is one of the key issues because the asymmetry makes it possible to use revealed preferences and the preference order to model economic inequities (see Table 7). In addition, asymmetries in occupational mobility may indicate changes in demographic and social characteristics and structure as well as subjective preferences, components that determine popularity of HEIs. It is unquestionable that HEIs play a key role in changing social and demographic processes (see Table 10).

The combination of network and economic models, on the one hand, can cross-validate the results and offers a technique for model triangulation. On the other hand, the economic models can explain the formation and properties, such as asymmetries, of mobility networks (see Tables 7 and 8). In addition, the combination offers better community detection (see Figs. 10 and 11 and a better explanation and prediction of revealed application preferences (see Table 8). Explaining network properties with economic models can open new horizons because it is possible to move from descriptive statistics on network properties to explanatory models, and thus, we can better understand the mechanism underlying the formation of networks.

Summary and conclusions

This study showed how which to measure the role of HEIs in application and occupation mobility by combining economic and network models. These results are based on application and career tracking databases that include all applicants and all freshly graduated people within the explored time interval. The paper demonstrates how to calculate the graduate premium, the added value of the locations and the HEIs. One can see that the added value of locations (on starting salaries) varies across occupation categories and correlates with the distribution of the companies (see Fig. 5). The added value of HEIs is usually limited to the vicinity of the HEI in question. The impact of a HEI at a larger distance is questionable (see Fig. 6).

The results highlighted the role of HEIs, especially in occupational mobility (see the high importance values in Table 9). Nevertheless, in the case of application mobility, the coefficient of faculty excellence decreases over time (see Table 6). In addition, the declining R^2 values over the years make it probable that factors other than economic factors, unemployment and faculty excellence are increasingly influencing student mobility (see Table 7). Finally, the appropriate model for network asymmetry can explain the precedence orders and predict the number of applications of HEIs (see Table 8).

Although the authors only analyzed spatial (such as application and occupation) mobility in Hungary, the presented methods can also be generalized to data from other countries and other spatial and economic data. The proposed methods can cross-validate each other and create new insight to explain the formation of mobility networks. By using these methods, actual players in these processes may better understand their competitive position and design their strategies.

Abbreviations

BA:: Bachelor of Arts; BP:: Budapest, the capital of Hungary; BSc:: Bachelor of Science; GDI:: Gross Domestic Income; GDP:: Gross Domestic Product; HCSO:: Hungarian Central Statistical Office; HCSTGC:: Hungarian Central System for Tracking Graduates' Careers; HEI:: Higher Education Institution; ISCO:: International Standard Classification of Occupations; ISIC:: International Standard Industrial Classification; MA:: Master of Arts; MSc:: Master of Science; NUTS:: Nomenclature of Territorial Units; UR:: Unemployment Rate.

Acknowledgements

This work was supported by the TKP2020-NKA-10 project financed under the 2020-4.1.1-TKP2020 Thematic Excellence Programme by the National Research, Development and Innovation Fund of Hungary and by the Research Centre at Faculty of Business and Economics (No. PE-GTK-GSKK A095000000-1) of University of Pannonia (Veszprém, Hungary). This work was partially supported by the BME-Artificial Intelligence FIKP grant of EMMI (BME FIKP-MI/SC) and by the Ministry of Innovation and Technology and the National Research, Development and Innovation Office within the Artificial Intelligence National Laboratory of Hungary.

Authors' contributions

ZTK and AT designed the study. ZB and VVC prepared and cleaned the data, and ÁJ and IN-V conducted the literature review, but all authors were involved in writing the paper. All authors read and approved the final manuscript.

Funding

There were no funding sources used for the work in this study.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Competing interests

Authors declare that they have no competing interests.

Author details

¹Department of Quantitative Methods, University of Pannonia, Egyetem str. 10, Veszprém 8200, Hungary. ²MTA-PE Budapest Ranking Research Group, Budapest, Hungary. ³Institute of Advanced Studies (iASK), Kőszeg, Hungary. ⁴Eötvös Loránd University, Budapest, Hungary. ⁵Big Geospatial Data Lab, Research Centre for Astronomy and Earth Sciences, Budapest, Hungary. ⁶Wigner RCP Department of Computational Sciences, Budapest, Hungary. ⁷Department of Computer Science and Information Theory, Budapest University of Technology and Economics, Budapest, Hungary.

Received: 21 May 2021 Accepted: 18 October 2021

Published online: 08 November 2021

References

- Abbott A, Silles M (2016) Determinants of international student migration. *World Econ* 39(5):621–635
- Adkisson RV, Peach JT (2008) Non-resident enrollment and non-resident tuition at land grant colleges and universities. *Educ Econ* 16(1):75–88
- Agasisti T, Dal Bianco A (2007) Determinants of college student migration in Italy: Empirical evidence from a gravity approach. SSRN
- Ahmad AB, Shah M (2018) International students' choice to study in China: an exploratory study. *Tertiary Educ Manag* 24(4):325–337
- Alm J, Winters JV (2009) Distance and intrastate college student migration. *Econ Educ Rev* 28(6):728–738
- Anderson JE (1979) A theoretical foundation for the gravity equation. *Am Econ Rev* 69(1):106–116
- Anderson JE, Van Wincoop E (2003) Gravity with gravitas: a solution to the border puzzle. *Am Econ Rev* 93(1):170–192
- Ashby NJ (2007) Economic freedom and migration flows between US states. *Southern Econ J* 677–697
- Bacci S, Bertaccini B (2020) Assessment of the university reputation through the analysis of the student mobility. *Soc Indic Res*
- Beine M, Noël R, Ragot L (2014) Determinants of the international mobility of students. *Econ Educ Rev* 41:40–54
- Bergstrand JH (1985) The gravity equation in international trade: some microeconomic foundations and empirical evidence. *Rev Econ Stat* 474–481
- Bergstrand JH (1989) The generalized gravity equation, monopolistic competition, and the factor-proportions theory in international trade. *Rev Econ Stat* 143–153
- Bernela B, Bouba-Olga O, Ferru M (2018) Spatial patterns of PhDs' internal migration in France, 1970–2000. *J Innov Econ Manag* 1(25):33–56
- Bierens HJ, Kontuly T (2008) Testing the regional restructuring hypothesis in Western Germany. *Environ Plan A* 40(7):1713–1727
- Bilecen B, Gamper M, Lubbers MJ (2018) The missing link: Social network analysis in migration and transnationalism. *Social Networks, The missing link: Social network analysis in migration and transnationalism*. 53:1–3
- Bratti M, Verzillo S (2019) The 'gravity' of quality: research quality and the attractiveness of universities in Italy. *Region Stud* 53(10):1385–1396
- Breitenfellner A, Cuaresma JC, Mooslechner P, Ritzberger-Grünwald D et al (2008) The impact of EU enlargement in 2004 and 2007 on FDI and migration flows: gravity analysis of factor mobility. *Monet Policy Econ* 2(8):101–120
- Brown P (2013) Education, opportunity and the prospects for social mobility. *Br J Sociol Econ* 34(5–6):678–700
- Candia J, González MC, Wang P, Schoenharl T, Madey G, Barabási A-L (2008) Uncovering individual and collective human dynamics from mobile phone records. *J Phys A: Math Theor* 41(22):224015
- Carey HC (1858) *Principles of Social Science*. J.B. Lippincott, Philadelphia
- Cascetta E (1984) Estimation of trip matrices from traffic counts and survey data: a generalized least squares estimator. *Transp Res B: Methodol* 18(4–5):289–299
- Catsiapis G (1987) A model of educational investment decisions. *Rev Econ Stat* 33–41
- Checchi D (2000) University education in Italy. *Int J Manpower*
- Chetty R, Friedman JN, Hendren N, Jones MR, Porter SR (2020) The opportunity atlas: Mapping the childhood roots of social mobility. NBER Working Paper Series
- Ciriaci D (2014) Does university quality influence the interregional mobility of students and graduates? the case of Italy. *Reg Stud* 48(10):1592–1608
- Cooke TJ, Boyle P (2011) The migration of high school graduates to college. *Educ Eval Policy Anal* 33(2):202–213
- Cseres-Gergely Z (2012) Can the modernisation of a public employment service be an effective labour market intervention? The Hungarian experience, 2004–2008. *Eur J Govern Econ* 1(2):145–162
- Cullinan J, Duggan J (2016) A school-level gravity model of student migration flows to higher education institutions. *Spat Econ Anal* 11(3):294–314
- Devillanova C, García-Fontes W (1998) Migration across Spanish provinces: evidence from the social security records (1978–1992). *Universitat Pompeu Fabra Economics WP* 1(318)
- Dotti NF, Fratesi U, Lenzi C, Percoco M (2013) Local labour markets and the interregional mobility of Italian university students. *Spat Econ Anal* 8(4):443–468
- Duchesne I, Nonneman W (1998) The demand for higher education in Belgium. *Econ Educ Rev* 17(2):211–218
- Ette A, Unger R, Graze P, Sauer L (2008) Measuring spatial mobility with the German microcensus: The case of German return migrants. *Z Bevölkerungswiss* 33(3–4):409–431
- Faggian A, Franklin RS (2014) Human capital redistribution in the USA: the migration of the college-bound. *Spat Econ Anal* 9(4):376–395
- Fotheringham AS, Brunsdon C, Charlton M (2000) *Quantitative geography: perspectives on spatial data analysis*. Sage, London

- Franklin RS, Faggian A (2014) College student migration in new england: Who comes, who goes, and why we might care. *Northeastern Geographer*, 6
- Fuller WC, Manski CF, Wise DA (1982) New evidence on the economic determinants of postsecondary schooling choices. *J Human Resour* 477–498
- Gadar L, Abonyi J (2018) Graph configuration model based evaluation of the education-occupation match. *PLoS ONE* 13(3):1–19
- Gadár L, Koszttyán ZT, Abonyi J (2018) The settlement structure is reflected in personal investments: distance-dependent network modularity-based measurement of regional attractiveness. *Complexity*, 2018
- Gadár L, Koszttyán ZT, Telcs A, Abonyi J (2020) A multilayer and spatial description of the Erasmus mobility network. *Scientific Data* 7(1):1–11
- González Canché M (2018) Geographical network analysis and spatial econometrics as tools to enhance our understanding of student migration patterns and benefits in the u.s. higher education network. *Rev High Educ* 41(2):169–216
- Gonzalez MC, Hidalgo CA, Barabasi A-L (2008) Understanding individual human mobility patterns. *Nature* 453(7196):779–782
- Grogger J, Hanson GH (2015) Attracting talent: Location choices of foreign-born phds in the united states. *J Law Econ* 33(S1 Part2):S5–S38
- Hartog J, Diaz-Serrano L (2007) Earnings risk and demand for higher education: a cross-section test for Spain. *J Appl Econ* 10(1):1–28
- Haveman R, Smeeding T (2006) The role of higher education in social mobility. *Future Child* 16(2):125–150
- Heckman JJ (1979) Sample selection as a specification error. *Econometrica* 47:153–161
- Hilmer MJ, Hilmer CE (2012) On the relationship between student tastes and motivations, higher education decisions, and annual earnings. *Econ Educ Rev* 31(1):66–75
- Holdsworth C (2009) 'going away to uni': mobility, modernity, and independence of English higher education students. *Environ Plan A* 41(8):1849–1864
- Javakhishvili Larsen N, Mitze T (2015) Spatial mobility and early career effects in the Danish Labour market
- Karemera D, Oguledo VI, Davis B (2000) A gravity model analysis of international migration to North America. *Appl Econ* 32(13):1745–1755
- Kazakis P (2019) On the nexus between innovation, productivity and migration of us university graduates. *Spat Econ Anal* 14(4):465–485
- Kondakci Y, Bedenlier S, Zawacki-Richter O (2018) Social network analysis of international student mobility: uncovering the rise of regional hubs. *High Educ* 75(3):517–535
- Kovacs L, Kasza G (2018) Learning to integrate domestic and international students: the Hungarian experience. *Int Res Rev* 8(1):26–43
- Kulu H, Lundholm E, Malmberg G (2018) Is spatial mobility on the rise or in decline? an order-specific analysis of the migration of young adults in Sweden. *Popul Stud* 72(3):323–337 (PMID: 29663847)
- Lemistre P, Moreau N (2009) Spatial mobility and returns to education: some evidence from a sample of French youth. *J Reg Sci* 49(1):149–176
- Liu X, Murata T (2010) An efficient algorithm for optimizing bipartite modularity in bipartite networks. *J Adv Comput Intell Inform* 14(4):408–415
- Lourenço D, Sá C (2019) Spatial competition for students: What does (not) matter? *Ann Reg Sci* 63(1):147–162
- Lourenço D, Sá C, Tavares O, Cardoso S (2020) Enrolling in higher education: the impact of regional mobility and public-private substitution effects. *J Econ Issues* 54(1):183–197
- Lucas REB (2001) Diaspora and development: Highly skilled migrants from east Asia. Report prepared for the World Bank
- Magrini M-B, Lemistre P (2013) Distance-income migration trade-off of young French workers: an analysis per education level. *Reg Stud* 47(2):282–295
- Modell S (2015) Theoretical triangulation and pluralism in accounting research: a critical realist critique. *Account Audit Account J*
- Mohall M (2015) Measuring spatial mobility-towards new perspectives on accessibility, nr. 957., kulturgeografiska institutionen, uppsala universitet, Sweden
- Montmarquette C, Cannings K, Mahseredjian S (2002) How do young people choose college majors? *Econ Educ Rev* 21(6):543–556
- Nemes Nagy J, Tagai G (2011) Regional inequalities and the determination of spatial structure. *Region Stat: J Hungarian Central Statistical Office* 14(51):15–28
- Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. *Phys Rev E* 69(2):026113
- Nguyen A, Taylor J, Bradley S (2003) Relative pay and job satisfaction: some new evidence
- Niu S (2015) Leaving home state for college: differences by race/ethnicity and parental education. *Res High Educ* 56(4):325–359
- Odlyzko A (2015) The forgotten discovery of gravity models and the inefficiency of early railway networks. *CEconomia. History, Methodology, Philosophy*, 5(2)
- Orsuwan M, Heck RH (2009) Merit-based student aid and freshman interstate college migration: testing a dynamic model of policy change. *Res High Educ* 50(1):24–51
- Paas T (2003) Regional integration and international trade in the context of eu eastward enlargement. *HWWA Discussion Paper* 1(208):1–34
- Poot J, Alimi O, Cameron MP, Maré DC (2016) The gravity model of migration: the successful comeback of an ageing superstar in regional science. *J Region Res* 36:63–86
- Poston DL, Zhang L (2008) Ecological analyses of permanent and temporary migration streams in China in the 1990s. *Popul Res Policy Rev* 27(6):689
- Powell JJ, Finger C (2013) The bologna process's model of mobility in Europe: the relationship of its spatial and social dimensions. *Eur Educ Res J* 12(2):270–285
- Ravenstein EG (1885) The laws of migration. *J Stat Soc London* 48(2):167–235
- Ravenstein EG (1889) The laws of migration. *J Roy Stat Soc* 52(2):241–305

- Rogerson PA (1990) Buffon's needle and the estimation of migration distances. *Math Popul Stud* 2(3):229–238
- Rüger H, Feldhaus M, Becker KS, Schlegel M (2011) Circular job-related spatial mobility in Germany: comparative analyses of two representative surveys on the forms, prevalence and relevance in the context of partnership and family development. *Comp Popul Stud* 36(1):221–248
- Sá C, Florax RJ, Rietveld P (2004) Determinants of the regional demand for higher education in the Netherlands: a gravity model approach. *Reg Stud* 38(4):375–392
- Sá C, Florax RJ, Rietveld P (2012) Living arrangement and university choice of Dutch prospective students. *Reg Stud* 46(5):651–667
- Savoca E (1990) Another look at the demand for higher education: measuring the price sensitivity of the decision to apply to college. *Econ Educ Rev* 9(2):123–134
- Shen G (1999) Estimating nodal attractions with exogenous spatial interaction and impedance data using the gravity model. *Pap Reg Sci* 78(2):213–220
- Shields R (2013) Globalization and international student mobility: a network analysis. *Comp Educ Rev* 57(4):609–636
- Sila-Nowicka K, Vandrol J, Oshan T, Long JA, Demšar U, Fotheringham AS (2016) Analysis of human mobility patterns from gps trajectories and contextual information. *Int J Geogr Inf Sci* 30(5):881–906
- Telcs A, Kosztyán ZT, Neumann-Virág I, Katona A, Török A (2015) Analysis of Hungarian students' college choices. *Procedia-Soc Behav Sci* 191:255–263
- Telcs A, Kosztyán ZT, Török A (2016) Unbiased one-dimensional university ranking - application-based preference ordering. *J Appl Stat* 43(1):212–228
- Tinbergen JJ (1962) *Shaping the world economy; suggestions for an international economic policy*. Twentieth Century Fund, New York
- Tuckman HP (1970) Determinants of college student migration. *South Econ J* 37(2):184–189
- Venhorst V, Cörvers F, et al (2015) Entry into working life: Spatial mobility and the job match quality of higher-educated graduates. Maastricht University, Graduate School of Business and Economics (GSBE). Research Centre for Education and the Labour Market. ROA Research Memoranda, No. 003
- Venhorst V, Van Dijk J, Van Wissen L (2011) An analysis of trends in spatial mobility of Dutch graduates. *Spat Econ Anal* 6(1):57–82
- Warin T, Svaton P (2008) European migration: Welfare migration or economic migration? *Glob Econ J* 8(3):1850140
- Westerlund J, Wilhelmsson F (2011) Estimating the gravity model without gravity using panel data. *Appl Econ* 43(6):641–649
- Zheng Y, Li Q, Chen Y, Xie X, Ma W-Y (2008) Understanding mobility based on gps data. In: *Proceedings of the 10th international conference on Ubiquitous computing*, pp 312–321
- Zignani M, Gaito S (2010) Extracting human mobility patterns from gps-based traces. In: *2010 IFIP Wireless Days*, pp 1–5. IEEE

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
