STUDIES

The cognitive and geographical structure of knowledge links and how they influence firms' innovation performance

Tom Broekel
Department of Economic Geography,
Faculty of Geosciences,
Utrecht University,
The Netherlands
E-mail:
t.broekel@uu.nl

Ron Boschma
Department of Economic Geography,
Faculty of Geosciences,
Utrecht University,
The Netherlands
E-mail:
r.boschma@geo.uu.nl

Firms' embeddedness in knowledge networks has received much attention in literature. However, little is known about the structure of firms' knowledge exchange with respect to different types of proximities. Based on survey data of 295 firms in 8 European regions, we show that firms' knowledge exchange systematically differs in their geographical and cognitive dimensions. We find that firms' innovation performance is enhanced if the firm primarily links to technologically related as well as technologically similar organizations. Connecting with organizations at different geographical levels yields positive effects as well.

Keywords:
geographical proximity,
knowledge networks,
technological relatedness,
innovation performance

Introduction

Access to external knowledge is crucial for firms' research and development activities as it allows complementing internal resources (Powell et al. 1996). This implies that firms' embeddedness in knowledge networks is crucial for their economic success (cf. Uzzi 1996, Cantner–Graf 2004). This raises the question about the type of factors that influence the structure of knowledge networks. Boschma (2005) proposed five types of proximity that influence the probability of organizations to establish a knowledge link. Researchers have analysed the role these proximities play for the development of knowledge networks (cf. Breschi–Lissoni 2003, Balland, 2011) and firms' performance (cf. Gluckler 2007, Broekel–Boschma 2012).

To assess the relevance of proximities for firms' innovative success, studies usually relate the innovative output of a firm to the *average* proximity (based on one of the proximity dimensions) with its knowledge exchange partners. This means that these dimensions are often condensed into a single measure. However, little is known about the *structure* of the links of firms for knowledge exchange, concerning the types of proximity. This also applies to the impact of such structures on their innovative success. In addition, studies tend to focus only on one type of proximity, and when they do consider multiple proximity types, these are treated as being independent of each other. However, proximity in one dimension can substitute for proximity in another dimension. For instance, when two partners are cognitively distant, being geographically close allows for better communication because of the better possibilities to interact face-to-face. The proximity types may also work together when it comes to their effects on innovation. In this respect, cognitive proximity has an exceptional position among the proximity types (Boschma 2005). Similar to the other proximities, it not only facilitates knowledge exchange but also defines the learning potential for the creation of novel ideas and solutions. Therefore, we argue that the effects of the other types of proximities on innovative success have to be identified and empirically investigated in relation to that of cognitive proximity.

This paper analyses the cognitive and geographical structures of 295 firms' links for knowledge exchange in 8 European regions. In particular, we focus on the relationship between the geographical and cognitive structure of links and analyze whether firms interact more intensively with technologically similar, related, or dissimilar organizations at various geographical levels (regional and non-regional). By means of a cluster analysis technique, we establish an interaction between the cognitive with the geographical proximity dimension and systematically identify groups of firms with particular *combinations* of links with specific geographical and cognitive characteristics. We use that information to test whether particular geographical or cognitive structures of links or combinations of these are more conducive for innovative performance than other structures.

The paper is structured as follows. Section 2 briefly introduces the proximity framework and presents arguments on how the different proximities matter for innovative success. Section 3 unveils the employed data, and Section 4 sets out the empirical approach. Sections 5 and 6 present the results. Section 7 concludes the study.

Proximity, networks, and innovation

Firms' embeddedness into knowledge networks is increasingly being recognized as a factor influencing their economic and innovative success (Powell et al. 1996). A rich body of literature has emerged that analyses factors driving the formation of knowledge networks. Inspired by the French school of proximity (cf. Torre–Rallet 2005), Boschma (2005) proposes five types of proximity (i.e. cognitive, social, organizational, institutional, and geographical) that may affect firms' likelihood to engage in knowledge exchange with other organizations. A flourishing body of empirical literature has taken up these ideas and has investigated the driving forces behind network formation (cf. Breschi–Lissoni 2003, Cantner–Graf 2004, Ter Wal–Boschma 2009, Balland 2011, Broekel–Boschma 2012).

In addition, the structure of firms' links has received attention, with studies focusing on geographically proximate and distant partners (cf. Meyer-Krahmer 1985, Arndt–Sternberg 2000, Bathelt et al. 2004) as well as cognitively proximate and distant partners (cf. Fritsch 2003, Nooteboom et al. 2007, Balland et al. 2013). Especially, the economic relevance of the type of partner (in terms of the proximity dimensions) has been the focus of researchers. This focus is more relevant as high degrees of proximity may ease communication and make the establishment of links between firms more likely, but high levels of proximity may not lead to higher innovation performance *per se*, and may even possibly reduce it. This has been referred to as the 'proximity paradox' (Boschma–Frenken 2010). Accordingly, the benefits a firm can gain from particular links may be related to optimal levels of proximity in various dimensions. For instance, Sternberg & Arndt (2001) showed that balancing intra- and inter-regional links is most conducive for firms' innovation performance. Phelps (2010) found that access to greater variety improves innovation activities. Nooteboom et al. (2007) and Broekel & Boschma (2012) provided evidence that, in terms of innovative success, an optimal level of cognitive proximity might exist. That is, both higher cognitive proximity (greater technological similarity) and lower cognitive proximity (lower technological similarity) tend to reduce the innovation performance of firms.

In this context, Boschma & Frenken (2010) made a distinction between social, organizational, institutional, and geographical proximities on the one hand and cognitive proximity on the other hand. For the first four types of proximity, an 'optimal level' refers to a *combination* of links with low and high proximity. For instance, with respect to geographical proximity (and similarly for the other three proximity types), they state that an optimal level of geographical proximity does not equal an optimal geographical distance between two agents. Instead, agents should focus on balancing or combining local and non-local linkages. Contrarily, to be beneficial, they propose that an optimal level of cognitive proximity needs to exist for each link. This is not only relevant from a theoretical perspective but it also matters empirically. For instance, many studies that test for the effect of cognitive proximity estimate the *average* technological similarity (cognitive distance) of a focal firm to its direct knowledge exchange partners (Nooteboom et al. 2007, Broekel–Boschma 2012). However, the estimated *average* can be identical for very different distributions. In a very simplified example, two firms might have two partners each. For firm **A**, one link has a maximum technological similarity index of 1 and the other one is at a minimum with value 0. The average technological similarity of this firm to its partners is 0.5, which is the same value obtained for firm **B,** which has two partners with a similarity of 0.5 each (mean similarity). Accordingly, both firms obtain the average similarity index value of 0.5. However, in the proximity framework, firm **A**'s links are expected to lower the innovation performance of the firm as a great cognitive overlap (similarity of 1) implies a lack of variety for innovation creation and a low cognitive overlap (similarity of 0) goes hand in hand with a lack of effective communication. Firm **B** is expected to have a high innovation performance, since effective communication and sufficient levels of new variety characterise all the links of firm **B**. Analyses using the averaged similarity index treat both firms **A** and **B** identically, and therefore fail to discriminate between the two very different cases.

Boschma (2005) points out that proximity types are related to each other in two ways. First, proximities can be substitutes rather than complements when it comes to the establishment of links and ensure their success. It can be sufficient to be proximate in just one dimension to link with an organization: being proximate in another dimension does not yield further effects (Boschma–Frenken 2010). This argument is backed by some empirical studies. For instance, Singh (2005) shows that geographical proximity fosters the establishment of links among researchers who are distant in the cognitive dimension. Ponds et al. (2007) find that geographical proximity helps to overcome institutional distance. As revealed in their study, industry-university relationships are more likely to be established if firms and universities are geographically proximate, while university-university relationships span greater geographical distances.

Second, proximities can be related in terms of their effects on innovation. It is still empirically unclear if proximity in one dimension can substitute for proximity in another dimension in terms of firms' innovation performance. For instance, it seems plausible that geographical proximity allows effective communication because of the possibilities of face-to-face interaction, even when two partners might be cognitively distant. However, in many cases geographical proximity might just capture the effects of other types of proximity (social, institutional, organizational), questioning its direct and independent impact.

However, cognitive proximity has an exceptional position among the other proximity types. Similar to the other types of proximity, cognitive proximity not only facilitates knowledge exchange, but also defines the learning potential, i.e. the potential for the creation of novel ideas and solutions. In this context, some cognitive distance is necessary for facilitating interactive learning and innovation (Nooteboom 2000). The other proximity types (i.e. geographical, social, institutional, and organizational) are not directly related to the learning potential and primarily impact communication efficiency and the success of knowledge exchange (Boschma 2005). Therefore, we propose that the effects of these proximities on innovative success must be seen in relation to cognitive proximity. Our study takes up this issue by analyzing the relationship between geographical proximity (as an enabling factor of knowledge exchange) and cognitive proximity for innovation performance of firms. To be precise, we perform tests to determine if particular geographical or cognitive structures (or a combination of both) of knowledge exchange links are more conducive for firms' innovation performance.

Data

Our empirical study is based on a database collected as part of a European project on '*Constructing regional advantage: Towards state-of-the-art regional innovation policies in Europe?*' funded by the European Science Foundation. Research teams in a number of European countries (the Czech Republic, Germany, the Netherlands, Sweden, and Turkey) interviewed firms in different European regions and industries with a harmonized questionnaire. The questions were concerned with the R&D activities of firms, their innovative success, and their engagement in knowledge exchange activities. The firms provided detailed information on all kinds of organizations (other firms, associations, universities, etc.) that they were interacting with during the last three years and with whom they exchanged technological knowledge relevant for their innovation activities. The interviews were conducted by the word of mouth, and the interviewer adapted the questions according to the context of corresponding industry. The information facilitates an analysis on the sources of technological knowledge of the firms and the manner in which it affects their innovation performance.

Table 1

Overview of the cases

Industry	Region	Country	Firms
Automotive	Southwest Saxony	Germany	58
Aerospace	Netherlands	Netherlands	71
Biotechnology	North Rhine-Westphalia	Germany	23
Biotechnology	Prague	Czech Republic	16
Biotechnology	Scania	Sweden	30
ICT	Prague	Czech Republic	29
Video Game	Hamburg	Germany	20
Moving Media	Scania	Sweden	37
Textiles	Denizli	Turkey	31
Food	Scania	Sweden	28
Electronics	South Moravia	Czech Republic	29
Total			*372*

As shown in Table 1, data was collected on 372 firms in eight European regions and nine industries. A total of 295 firms reported at least one link with which they exchanged technological knowledge. On average, 5.31 links were mentioned (see Table A1). The distribution of the number of links is strongly skewed, with 50% of the firms reporting four or more links. Only 8% of the 295 firms mentioned just one organization with which they exchanged technological knowledge. The maximum number of links was 36.

In this study, we grouped firms that have similar links concerning their cognitive and geographical proximity. Our data restricted us to these two proximity types. In a similar manner as Cassiman et al. (2005), the technological similarity between an interviewed firm and its knowledge exchange partner is reported on a scale of 1 to 5 for each link. Accordingly, we measured technological similarity in a relative manner allowing for comparison across industries. For all 2,110 links, on average, a technological similarity of 2.83 was reported. For all firms and contacts, the geographical location (i.e. the latitude and longitude) was recorded, allowing the estimation of the geographical distance between the firm and the contact organization. The mean geographical distance is large with 886 km, which is however caused by a number of long distance links (>10,000). The median distance is just about 177 km.

Empirical Approach

Defining the structure of links

As pointed out earlier, studies analysing the effects of proximities on innovation often rely on the average geographical and cognitive distance between a firm and its direct knowledge exchange partners. However, the average distance gives only limited insights since the mean can be identical for very different empirical distributions.

To obtain a picture of the cognitive structure of firms' links, we constructed three variables: (1) the share of links that show low technological similarity; (2) the share of links with medium technological similarity; and (3) the share of links with organizations whose knowledge base is similar to that of the focal firm. These shares are derived from the 1-5 similarity variable and are as follows: values equal to 1 are defined as low technological similarity (SHARE-VARIETY), values from 2 to 3 are considered as medium similarity relationships (SHARE-RELATED), and values of 4 and 5 refer to high similarity (SHARE-SIMILAR). This classification is based on own experiences during interviews and results in a uniform distribution of observations across these classes, as shown in Table 2. We then grouped the firms according to the similarity in their profiles by performing a cluster analysis on the three shares.

Table 2

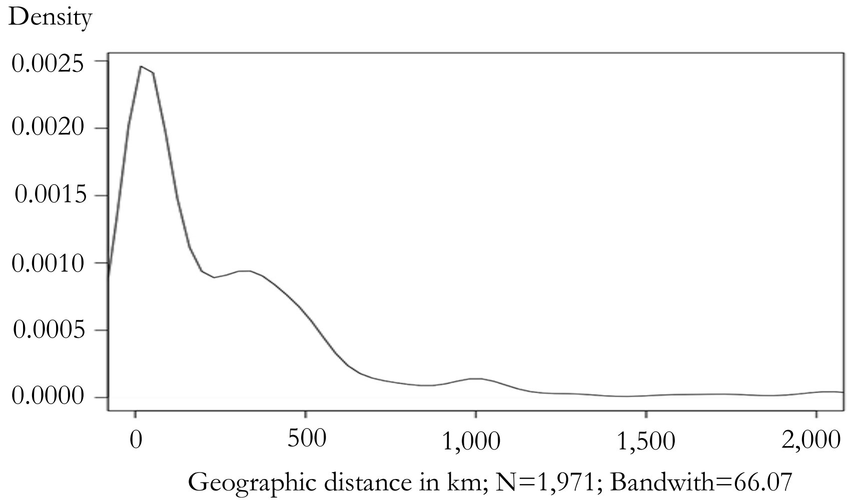
Mean values basic variables

Variables	Mean
SHARE-VARIETY	0.21
SHARE-RELATED	0.47
SHARE-SIMILAR	0.32
SHARE-REGIONAL	0.36
SHARE-MEDIUM-DIST	0.45
SHARE-LARGE-DIST	0.19
SHARE-REGIONAL&VARIETY	0.07
SHARE-REGIONAL&RELATED	0.17
SHARE-REGIONAL&SIMILAR	0.12
SHARE-NO-REGIONAL&VARIETY	0.13
SHARE-NO-REGIONAL&RELATED	0.30
SHARE-NO-REGIONAL&SIMILAR	0.20

We took an approach similar to that taken for cognitive dimension to analyse geographical dimension. We are particularly interested in the embeddedness of firms in regional knowledge networks, which refer to links with organizations at low geographical distances. It implied the need to differentiate between regional and non-regional linkages. The data covers observations from different countries and regions. For this reason, 'regional' may refer to quite different distances in each case. For example, distances between Dutch cities rarely exceed 200 km, while this is different in Sweden. Accordingly, the size of regions is likely to vary substantially, which makes the definition of a universal threshold distance problematic. Figure 1 shows the distribution of the geographical distances of all links (excluding those larger than 2,000 km), smoothed with a standard Gaussian kernel density. A clear peak of link distances at around 40 km and a flattening of the curve from 600 km onwards are visible. Based on this, three variables are created: (1) the share of links with distances equal or less than 40 km (SHARE-REGIONAL), (2) the share of linkages with distances between 40 and 600 km (SHARE-MEDIUM-DIST), and (3) the share of knowledge relations exceeding a distance of 600 km (SHARE-LARGE-DIST). Clusters for these three variables are identified with a cluster analysis.

Figure 1

Distribution of links' geographic distance



As shown above, we treated the cognitive and geographical structure of firms' partners independently of each other. However, this does not take into account the relationship between geographical and cognitive proximity, which refers to the idea that the two might have a complementary or substitutive relationship. To consider this empirically, the two dimensions are made to interacted. In order to keep the complexity at an acceptable level, we dropped the differentiation between the medium-range and large distance linkages, as most of the literature regarding the effects of geographical proximity emphasize the difference between regional and non-regional knowledge links. Based on this, the following variables were created: (1) the share of regional links with low technological similarity (SHARE-REGIONAL&VARIETY), (2) the share of regional links with medium technological similarity (SHARE-REGIONAL&RELATED), and (3) the share of regional links with high technological similarity (SHARE-REGIONAL&SIMILAR). The same variables are defined for non-regional links: SHARE-NO-REGIONAL&VARIETY, SHARE-NO-REGIONAL&RELATED, and SHARE-NO-REGIONAL&SIMILAR. As stated before, a cluster analysis is performed on the six variables to group firms with similar geographical and cognitive links.

The cognitive and geographical structure of firms' knowledge links

While there are several approaches to performing cluster analysis (Kaufman–Rousseeuw 1990), we used the *pdfCluster* method (Bin–Risso 2011). It employs nonparametric model-based clustering techniques as proposed by Azzalini & Torelli (2007). The advantages of this method are that the 'correct' number of clusters is derived from the data and that it works well in situations of high dimensionality, i.e. situations in which few observations but many variables are available (for more details see Bin–Risso 2011).

Cognitive structure of firms' knowledge links

The application of the *pdfCluster* method to the three shares representing the cognitive dimension (SHARE-VARIETY, SHARE-RELATED, and SHARE-SIMILAR) results in the identification of six clusters, whose characteristics are presented in Table 3. Each of the six clusters summarizes at least 30 observations. Cluster 1, 2, and 4 clearly represent firms with primary links to technologically similar (2: SIMILAR), technologically related (1: RELATED), and technologically unrelated (4: VARIETY) organizations. The other three clusters summarize firms with many related and some highly similar links (3: RELATED&SIMILAR), firms with many highly similar and a few related links (5: SIMILAR&RELATED), and firms with equal shares of links with low and medium similarity (6: VARIETY&RELATED).

Table 3

Cluster: Cognitive dimension

Cluster	1	2	3	4	5	6
Size	70	36	71	36	50	32
SHARE-VARIETY	0.02	0.01	0.07	0.93	0.11	0.46
SHARE-RELATED	0.95	0.01	0.61	0.03	0.27	0.47
SHARE-SIMILAR	0.02	0.99	0.33	0.04	0.62	0.07
Cluster Names:	**RELATED**	**SIMILAR**	**RELATED& SIMILAR**	**VARIETY**	**SIMILAR&R ELATED**	**VARIETY& RELATED**
Share of AUTO	0.09	0.12	0.30	0.04	0.32	0.14
Share of BIO	0.36	0.14	0.22	0.09	0.14	0.05
Share of AERO	0.34	0.06	0.25	0.09	0.06	0.21
Share of TEXT	0.12	0.24	0.12	0.32	0.08	0.12
Share of MEDIA	0.23	0.13	0.19	0.23	0.13	0.10
Share of ICT	0.00	0.26	0.26	0.16	0.32	0.00
Share of VIDEO	0.10	0.05	0.35	0.10	0.35	0.05
Share of FOOD	0.46	0.04	0.23	0.12	0.04	0.12
Mean LINKS_TEC	5.24	3.56	8.18	5.14	10.10	6.50
Mean AGE	37.39	21.00	30.42	26.53	20.86	30.47
Mean EMPL	170.27	82.69	706.17	78.06	84.40	118.63
Mean SKILL	49.13	43.22	41.82	46.31	36.34	39.69
Mean RD_INT	0.27	0.28	0.30	0.37	0.36	0.31
Mean SIM	2.52	4.45	3.02	1.07	3.59	1.91

In our example, average similarity (Mean SIM) seems to be a reliable approximation of the distribution as it clearly allows differentiating between firms that are dominated by links of low technological similarity (VARIETY), those dominated by links of medium similarity (RELATED), and firms that primarily link with very similar organizations (SIMILAR) links. In case of the two clusters (i.e. RELATED&SIMILAR and SIMILAR&RELATED), the cluster analysis offers richer information on the cognitive structure of firms' links than the simple mean value of technological similarity (SIM).

Some more insights into the structure are derived when relating the cluster classification to a number of firm characteristics collected during the interviews. The first characteristic is the absorptive capacity of firms. It is approximated by four variables: the share of R&D employees in total employment (RD_INT), the share of employees with at least a bachelor's degree (SKILL), firm's age (AGE), and the number of employees (EMPL). Last, it might be worthwhile to look at the absolute number of links for exchanging technological knowledge (LINKS_TEC). This variable represents a firms' general propensity to engage in technological knowledge exchange. Moreover, based on the eight industry dummies (AUTO, BIO, AERO, TEXT, MEDIA, ICT, VIDEO, and FOOD), we estimate the share of firms in a cluster with respect to each of these industries.

Table 3 highlights that there are no ICT-producing firms in Cluster RELATED and firms based on the textile and video gaming industry are barely present in clusters SIMILAR and VARIETY. It is important to note that no cluster is dominated by a single industry and firms of each industry are included in almost all the clusters. Therefore, the clusters do not represent industry characteristics. Cluster RELATED&SIMILAR tends to contain larger organizations (Mean EMPL) with more links (Mean LINK_TEC). Firms mainly connected with very similar organizations (SIMILAR) tend to have relatively small number of links.

Geographical structure of firms' links

The results of the cluster analysis for the geographical dimension suggest the existence of five clearly defined groups of firms, as shown in Table 4. The smallest cluster summarizes 18 firms and the largest summarizes 88 firms. Similar to the cognitive dimension, clusters 1, 2 and 4 represent the 'specialized' cases that summarize: firms which primarily link at medium distances (1: MEDIUM-DIST); firms that mainly have regional linkages (2: REGIONAL); and firms that extensively engage in long distance knowledge exchange (4: LARGE-DIST). Cluster 5 (5: MIX-DIST) includes firms that maintain an equal share of linkages to organizations at all three spatial levels. Contrarily, Cluster 3 (3: REGIONAL&MEDIUM-DIST) represents firms that lack large-distance linkages.

Table 4

Clusters: geographical dimension

Clusters	1	2	3	4	5
Size	69	53	88	18	67
SHARE_REGIONAL	0.04	0.94	0.44	0.03	0.24
SHARE_MED-DIST	0.91	0.04	0.52	0.04	0.32
SHARE_LARGE-DIST	0.05	0.02	0.05	0.94	0.44
Cluster Names:	**MEDIUM-DIST**	**REGIONAL**	**REGIONAL&MEDIUM-DIST**	**LARGE-DIST**	**MIX-DIST**
Share of AUTO	0.35	0.02	0.53	0.02	0.09
Share of BIO	0.17	0.14	0.28	0.05	0.36
Share of AERO	0.49	0.09	0.23	0.02	0.17
Share of TEXT	0.12	0.44	0.04	0.12	0.28
Share of MEDIA	0.10	0.39	0.19	0.16	0.16
Share of ICT	0.00	0.16	0.47	0.11	0.26
Share of VIDEO	0.05	0.30	0.30	0.10	0.25
Share of FOOD	0.19	0.23	0.23	0.04	0.31
Mean LINKS_TEC	7.52	4.06	8.05	4.11	6.84
Mean AGE	37.09	27.15	27.96	17.72	25.81
Mean EMPL	151.58	817.17	91.57	235.00	146.31
Mean SKILL	27.57	47.019	45.36	44.50	52.70
Mean RD_INT	0.25	0.36	0.32	0.31	0.32
Mean DIST	437	107	322	4,447	1,887

The average geographical distance measure (Mean DIST) differentiates well between the 'specialized' cases: REGIONAL, MEDIUM-DIST, and LARGE-DIST. However, the measure works less well for differentiating between these three specialized clusters and those cluster summarizing firms whose links fall in multiple distance categories (clusters 3: REGIONAL&MEDIUM-DIST and cluster 5: MIX-DIST). In particular, firms with equal shares of links at all three geographical levels (MIX-DIST) will be classified as firms with rather long-distance relations when using the average geographical distance of their links, which obtains the second highest value of 1,887 km. However, the majority of these firms' links (56%) is still shorter than 600 km.

As in the case of the clusters based on the technological characteristics of links, we relate the cluster classification to firm characteristics. Table 4 reveals that all the clusters in this analysis contain firms of different industries. Therefore, the identified geographical patterns do not 'simply' reflect industry specifics. Moreover, firms in the clusters REGIONAL and LARGE-DIST have, on average, fewer linkages for exchanging technological knowledge. On average, the cluster REGIONAL has larger firms (EMPL), while the cluster LARGE has relatively younger firms (AGE). Cluster MED consists of smaller (EMPL) and older firms (AGE) and firms with lower shares of highly qualified employees (SKILL).

Table 5

Clusters: cognitive and geographical dimension

Clusters	1	2	3	4	5	6	7
Size	97	22	46	15	58	40	17
SHARE-REGIONAL&SHARE_VARIETY	0.03	0.00	0.10	0.01	0.02	0.03	0.70
SHARE-REGIONAL&SHARE-RELATED	0.04	0.02	0.02	0.96	0.42	0.10	0.15
SHARE-REGIONAL&SHARE_SIMILAR	0.03	0.02	0.22	0.00	0.09	0.39	0.06
SHARE-NO-REGIONAL&SHARE-VARIETY	0.06	0.02	0.55	0.01	0.06	0.08	0.06
SHARE-NO-REGIONAL&SHARE-RELATED	0.66	0.05	0.07	0.02	0.28	0.09	0.01
SHARE-NO-REGIONAL&SHARE-SIMILAR	0.18	0.89	0.03	0.00	0.14	0.32	0.03
Cluster names:	NO-REGIONAL	NOREGIONAL&SIMILAR	**REGIONAL&SIMILAR/ NO-REGIONAL&VARIETY**	**REGIONAL&RELATED**	**RELATED**	**SIMILAR**	**REGIONAL&VARIETY**
Share of AUTO	0.42	0.12	0.07	0.00	0.23	0.16	0.00
Share of BIO	0.36	0.06	0.13	0.09	0.16	0.17	0.03
Share of AERO	0.51	0.04	0.15	0.06	0.15	0.08	0.02
Share of TEXT	0.08	0.00	0.48	0.04	0.16	0.08	0.16
Share of MEDIA	0.16	0.13	0.19	0.07	0.23	0.07	0.16
Share of ICT	0.11	0.21	0.16	0.00	0.16	0.32	0.05
Share of VIDEO	0.20	0.05	0.10	0.05	0.25	0.25	0.10
Share of FOOD	0.39	0.00	0.12	0.08	0.31	0.04	0.08
Mean LINKS_TEC	8.51	5.14	4.85	2.40	7.33	6.58	5.24
Mean AGE	34.04	13.82	25.33	53.33	31.17	20.00	19.24
Mean EMPL	159.20	102.46	92.65	65.87	863.95	57.00	30.88
Mean SKILL	38.61	45.50	39.70	53.80	46.95	44.28	49.71
Mean RD_INT	0.29	0.33	0.29	0.22	0.32	0.31	0.51

Combining cognitive and geographical dimensions of knowledge links

In this section, we explore whether firms interact more intensively with technologically similar or dissimilar organizations at particular geographical distances. For the sake of simplicity, we reduced the geographical dimension to two variables: share of regional (SHARE-REGIONAL) and non-regional links (SHARE-NO-REGIONAL). These variables are interacted with the three variables on technological similarity (SHARE-VARIETY, SHARE-RELATED, SHARE-SIMILAR). This results in six variables that serve as inputs for the cluster analysis.

As shown in Table 5, the analysis identifies seven distinct clusters, which vary in size between 15 and 97 firms. Cluster 5 (RELATED) and cluster 6 (SIMILAR) resemble mainly the cognitive structure of links. Two clusters represent firms with primarily links to organizations outside their region that are either technologically related (Cluster 1, NO-REGIONAL&RELATED) or highly similar (cluster 2, NO-REGIONAL&SIMILAR). Two other clusters consist of firms interacting primarily with regional organizations that are either technologically dissimilar (cluster 6, REGIONAL&VARIETY) or technologically related (cluster 4, REGIONAL& RELATED). Last, there is one 'mixed' cluster (cluster 3, REGIONAL& SIMILAR/NO-REGIONAL&VARIETY) that summarizes firms with complex partner structures: firms have large shares of regional relationships with technologically similar organizations and large shares of non-regional links with dissimilar organizations.

Substitution versus complementarity

The identification of these clusters is interesting insofar as they allow inference to either a substitutive or a complementary relationship between the cognitive and geographical dimension. According to the concept of a substitutive relationship, geographical proximity allows for easier communication and overcoming cognitive distance between firms. The existence of the REGIONAL&VARIETY cluster corresponds precisely to this concept as it summarizes firms interacting with technologically dissimilar knowledge sources at close geographical distance. Cluster NO-REGIONAL&SIMILAR describes the opposite case in which large geographical distances are overcome when the knowledge bases of two organizations are technologically similar (low cognitive distance). However, we did not find a particular geographical pattern for the exchange of related knowledge: organizations interact with organizations having related knowledge at close geographical distance (REGIONAL&RELATED), over larger distances (NO-REGIONAL&RELATED), or at a combination of geographical distances (RELATED).

However, two clusters confute this idea of substitution. They rather suggest that geographical proximity enhances the interaction with technologically similar organizations. The first cluster, SIMILAR, refers to firms linking with technologically similar organizations at the regional as well as at the non-regional level. The cluster is orthogonal to the existence of a substitutive relationship between the two dimensions. To an even larger extent, the same applies to the cluster REGIONAL&SIMILAR/ NO-REGIONAL&VARIETY. Here, firms interact with technologically similar organizations at close geographical distance and simultaneously with dissimilar organizations at high geographical distance. Therefore, these two clusters rather hint at a complementary relationship between the two proximity dimensions.

Out of the total number of firms assigned to each cluster, it is observed that… 39 firms fully support the substitution hypothesis (part of the NO-REGIONAL&SIMILAR and REGIONAL&VARIETY clusters), while 46 firms (part of the REGIONAL&SIMILAR/NO-REGIONAL&VARIETY clusters) suggest a complementary relationship. Therefore, the cluster analysis indicates that both complementary and substitutive relationships exist between the geographical and cognitive proximity types. These relationships between the two proximity dimensions was explored further by identifying the characteristics of the firms. This was achieved through a multinomial regression analysis. In this analysis, we selected all the clusters that are did not indicate either of the relationships as reference groups, i.e. all the clusters that represented relationships between organizations with medium levels of technological similarity (NO-REGIONAL&RELALTED, REGIONAL& RELATED, RELATED). We also aggregated the two clusters that represented the substitution hypothesis (REGIONAL&VARIETY, NO-REGIONAL&SIMILAR), owing to the presence of a small number of firms in the REGIONAL&VARIETY cluster. The two clusters that support the complementarity hypothesis (REGIONAL&SIMILAR/NO-REGIONAL&VARIETY and SIMILAR) are kept separate, as the SIMILAR cluster does not clearly indicate complementarity.

The variables that were introduced as firm characteristics include RD_INT and SKILL, AGE, EMPL, LINKS_TEC, and eight industry dummies. However, to obtain estimates that are more precise we included additional control variables. We control for the technological orientation of firms by considering the share of engineers (ENG) in total employment. Second, organizations may also differ with respect to their openness towards external knowledge. The variable (OPEN) describes if a firm perceives external knowledge as being highly relevant for its innovation activities.[1] Third, the total number of links regarding the exchange of market knowledge (LINKS_MA) is included to control for the differences in firms' general collaboration behaviour and the potential innovation stimulating effects resulting from superior market information. As we have data on biotech firms in three

[1] This information was collected by the following question we posted during the interviews: 'Please indicate in terms of percentage the relative importance of: a) knowledge acquired inside the company; b) knowledge acquired outside the company (adding up to 100%)'.

different countries, we took into account two sector-country dummies (SWED_BIO, CZECH_BIO) to control for potential country effects in this sector.

The results of the analysis are presented in Table 6. The first interesting finding is concerned with the coefficients of the sector dummies. While we have already pointed out that all the clusters contain firms that belong to different industries, the regression analysis adds further evidence that none of the detected geographical and cognitive structures is restricted to just one industry. For instance, the cluster NO-REGIONAL&SIMILAR/REGIONAL&VARIETY seems to correspond primarily to the characteristics of firms in the ICT sector. However, firms of the same sector are also significant members in the cluster SIMILAR (positive significant coefficient of ICT). Similar findings are observed for other sector dummies and clusters.

Table 6

Multinomial regression clusters

Clusters	NO-REGIONAL&SIMILAR/REGIONAL&VARIETY	REGIONAL&SIMILAR/NO_REGIONAL&VARIETY	SIMILAR	NO-REGIONAL&RELATED REGIONAL&RELATED RELATED
Obs:	39	46	40	170
CONST	0.903	0.143	–0.178	Reference group
Log(AGE)	–0.569*	0.101	–0.103	
Log(EMPL)	–0.215	–0.561**	–0.412	
SKILL	–0.002	0.001	–0.009	
ENG	–0.005	0.006	–0.007	
RD_INT	0.295	–1.140	–1.256	
AUTO	1.478	0.185	2.336*	
BIO	–0.231	1.656*	2.968**	
AERO	–0.103	0.176	1.243	
TEXT	1.553	3.143***	1.947	
MEDIA	0.960	0.939	1.347	
VIDEO	0.498	0.221	2.822*	
ICT	2.273*	1.285	4.016***	
FOOD	(omitted)	(omitted)	(omitted)	
CZECH_BIO	0.429	–2.041	–1.154	
SWED_BIO	–0.290	–2.119*	–1.331	
EXTERN	0.007	0.000	–0.006	
LINKS_TEC	–0.132*	–0.110	–0.029	
LINKS_MA	0.005	0.044	–0.007	
Obs:	295.000			
LR chi2(51)	96.28			
Log Likelihood	–289.879			
Pseudo R2	0.1424			

Another interesting finding is that firms in cluster NO-REGIONAL&SIMILAR and REGIONAL&VARIETY tend to be younger (AGE), which suggests that the substitutive relationship between cognitive and geographical proximity holds primarily for younger firms. Given their limited experiences, geographical proximity seems to be more important for these firms for maintaining successful collaborations with technologically dissimilar organizations. Cognitive proximity eases communication and as a result enables younger firms to interact with more geographically distant partners. Moreover, a substitutive relationship between the cognitive and geographical dimension tends to be more relevant for firms with low numbers of links concerning technological knowledge (LINKS_TEC). It might be attributed to the reluctance of firms to collaborate (having fewer links) and the caution they exercise while selecting partners. They may choose geographically proximate organizations because they can monitor them easily. Moreover, social relationships, which are more likely to exist between geographically proximate partners, might imply higher levels of trust. These relationships, in turn, may reduce the perceived obstacles to exchange knowledge between cognitively distant partners.

In contrast, the negative coefficient of EMPL for REGIONAL&SIMILAR/NO-REGIONAL&VARIETY suggests a complementary relationship between the two types of proximities that holds particularly for smaller firms. Accordingly, smaller firms (note that EMPL is not correlated with AGE) are well embedded in regional clusters of technologically similar firms. At the same time, they maintain significant relationships with dissimilar firms at larger distances. Apparently, this should make them gatekeepers that provide access to geographically distant and diverse knowledge sources. This finding clearly contradicts our expectation according to which this role should have been played by large firms (see also Lazerson–Lorenzoni 1999 and Graf 2011). We did not observe statistically significant differences between firms of cluster SIMILAR and our control group.

So far, our analysis demonstrated that systematic differences exist in the structures of firms' links along the geographical and cognitive dimension. In the subsequent sections, we focus on whether these structural differences can impact the innovation performance of firms.

Impact of link structures on firms' innovation performance

Empirical approach

As is common in innovation studies, we approximated the innovation performance of firms by means of the share of significantly improved products/processes on a firm's turnover (INN).[2] In addition to the geographical and cognitive composition of firms' partners, we considered all previously presented firm characteristics, which might influence the innovative success of firms.

Our dependent variable (INN) is a share within the interval [0,1]. The OLS regressions can be applied to such data when the variable is 'logit' transformed ($\hat{y} = log(y/\ (1 - y))$). However, Ferrari & Cribari-Neto (2004) point out that such an approach has several shortcomings, which include the problematic interpretation of coefficients and tendency towards asymmetric distribution of proportions (shares). Moreover, OLS regressions based on such data tend to be heteroskedastic. To overcome these drawbacks, they proposed *beta regression*, which is based on the assumption that the dependent variable is beta-distributed. Depending on the specification of some parameter values, beta's density can assume different shapes and it can be used to model left and right skewed distributions. We use beta regression with maximum-likelihood estimation and a 'logit' specification, as proposed by Cribari-Neto & Zeileis (2010), which implies that the obtained coefficients can be interpreted as change in the logit.[3] In some instances, our dependent variable is equal to, 0, or 1. To apply a beta regression, as suggested by Cribari-Neto & Zeileis (2010) and Smithson & Verkuilen (2006), we transformed the variable (y) using the following equation: $(y*(n\text{-}1) + 0.5)/n$, whereby n represents the number of observations (372).

Factors influencing firms' innovation performance

The cluster dummies created previously are employed to approximate differences in firms' choice of knowledge exchange partners. Using the beta regression approach, these cluster dummies and other firm characteristics are regressed on the share of a firm's turnover attributable to new or significantly improved products or processes that are not older than 3 years (INN). The results are shown in Table 7. We did not report the model with all cluster dummies because the overlap between the dummies (firms are classified into multiple clusters) causes linear dependency and multicollinearity issues. We tested each proximity dimension (cognitive and geographical) separately before considering the combined clusters (cognitive & geographical).

[2] This information was collected with the following question: 'Please indicate how much percentage of the turnover of your firm is attributed to new, dramatically improved products/processes introduced in the last three years?'

[3] We also ran models specifying a 'log' and a 'probit' link function, but the results remained stable.

Table 7

Impact on innovation performance of firms

Model	I	p-value	II	p-value	III	p-value	IV	p-value	V	p-value	VI	p-value	VII	p-value
CONST	-0.258	0.6186	-0.238	0.653	-0.182	0.729	-0.170	0.744	-0.124	0.813	-0.149	0.774	-0.163	0.752
Log(AGE)	-0.146	0.119	-0.142	0.136	-0.151	0.108	-0.152	0.108	-0.163*	0.084	-0.159*	0.092	-0.161*	0.083
Log(EMPL)	-0.072	0.2644	-0.090	0.178	-0.092	0.166	-0.111*	0.097	-0.112*	0.094	-0.097	0.139	-0.099	0.128
SKILL	0.005**	0.0493	0.005**	0.037	0.005*	0.052	0.005**	0.047	0.005**	0.044	0.005*	0.057	0.005**	0.040
ENG	0.003	0.2887	0.003	0.260	0.003	0.242	0.003	0.184	0.003	0.197	0.003	0.231	0.003	0.229
RD_INT	-0.596*	0.0887	-0.705*	0.057	-0.678*	0.068	-0.732**	0.048	-0.709*	0.055	-0.586*	0.099	-0.664*	0.056
AUTO	-0.109	0.7408	-0.335	0.360	-0.245	0.508	-0.305	0.403	-0.256	0.483	-0.070	0.836	-0.217	0.512
BIO	0.765*	0.0871	0.637	0.177	0.588	0.214	0.633	0.179	0.562	0.231	0.742	0.101	0.562	0.209
AERO	-0.155	0.5855	-0.304	0.337	-0.297	0.351	-0.342	0.277	-0.315	0.316	-0.169	0.563	-0.255	0.370
TEXT	-0.199	0.6105	-0.292	0.488	-0.281	0.504	-0.199	0.638	-0.232	0.581	-0.104	0.799	-0.262	0.501
MEDIA	-0.326	0.4025	-0.453	0.276	-0.417	0.318	-0.398	0.335	-0.411	0.318	-0.266	0.503	-0.420	0.277
SOFT	0.071	0.8365	-0.082	0.824	-0.054	0.884	-0.044	0.904	-0.096	0.794	0.063	0.857	-0.112	0.747
VIDEO	0.886**	0.0311	0.641	0.154	0.745*	0.097	0.685	0.123	0.608	0.171	0.822**	0.049	0.611	0.141
SWED_BIO	-0.852**	0.0185	-0.890**	0.016	-0.934**	0.011	-0.893**	0.014	-0.923**	0.011	-0.942***	0.009	-0.909**	0.012
CZECH_BIO	0.048	0.9144	0.131	0.772	0.115	0.800	0.163	0.718	0.215	0.633	0.183	0.684	0.170	0.702
EXTERN	-0.004*	0.0989	-0.004*	0.099	-0.005*	0.075	-0.004	0.104	-0.004*	0.084	-0.005**	0.069	-0.004*	0.092
LINKS_MA	0.011	0.6057	0.022	0.255	0.005	0.811	-0.001	0.945	-0.001	0.945	0.002	0.935	0.005	0.816
LINKS_TEC	0.022	0.2506	0.003	0.884	0.022	0.256	0.022	0.244	0.022	0.249	0.022	0.256	0.022	0.257
REL			0.058	0.824										
HIGH			0.186	0.515										
REL_HIGH			0.646**	0.019					0.376*	0.062	0.370**	0.041	0.519***	0.003
LOW			-0.004	0.989										
HIGH_REL			0.297	0.329										
LOW_REL			0.319	0.301										
MED					0.231	0.388								
REG					0.043	0.876								
REG_MED					0.210	0.426								
LARGE					0.069	0.847								
GEO_MIX					0.501*	0.065			0.269	0.140	0.306*	0.089	0.306*	0.083
NONREG_REL							0.483*	0.069	0.275	0.323	0.026	0.883		
NONREG_HIGH							-0.004	0.991	-0.090	0.792	-0.324	0.262		
REG_HIGH_NONREG_LOW							0.032	0.907	0.004	0.989	-0.210	0.330		
REG_REL							-0.213	0.570	-0.205	0.584	-0.391	0.251		
COM_REL							0.512*	0.068	0.234	0.440	-0.020	0.936		
COM_HIGH							0.494*	0.091	0.429	0.145			0.339	0.120
REG_LOW							0.006	0.987	-0.026	0.945	-0.256	0.431		
(phi)	1.056***	0.000	1.044***	0.000	1.027***	0.000	1.092***	0.000	1.109***	0.000	1.102***	0.000	1.100***	0.000
Log-Lik.			348 on 26 Df		345.3 on 25 Df		323.5 on 27 Df		326.4 on 29 Df		325.3 on 28 Df		324.9 on 23 Df	
Pseudo R^2			0.196		0.183		0.197		0.212		0.207		0.205	

Before discussing the results for the cluster dummies, we analyzed firm characteristics that gained significance in all model specifications (I–VII). The variable SKILL is positive and significant in all the models, implying that larger shares of highly qualified employees increase the innovation performance of firms. Surprisingly, the share of R&D employees (RD_INT) is negative and significant in all models. However, the variable is strongly positively correlated with SKILL (+0.33***) and INNO (+0.13**). It is therefore likely that SKILL captures the positive effect of RD_INT. As a result, the remaining variance of RD_INT that is not captured by SKILL might reflect some sector specifics (e.g. the negative correlation between TEXT and INNO).[4] Moreover, EXTERN is negative and significant in almost all models. It suggests that firms that attribute a high importance to external knowledge are less innovative. While the negative significances of EMPL in models IV and V as well as that of AGE in models V, VI, and VII are in line with literature (cf. Frenkel–Schefer 1998), they appear only in a limited number of models. Although there are no signs of multicollinearity, it was shown above that EMPL and AGE are significantly related to particular structures in the combined cognitive and geographic dimensions. As these variables become significant in those models including all cluster dummies of the combined cognitive and geographic dimensions, they probably capture some effects of the corresponding clusters. However, given the weakness of their effects, we refrained from disentangling these variables any further.

Here, we focus on the effects of the cognitive and geographical proximity dimensions. In model II, we included only the dummies for clusters representing differences in the cognitive dimension of the firms' partners. We expected firms in clusters SIMILAR, VARIETY, VARIETY/RELATED, and SIMILAR/RELATED to be outperformed by the firms in clusters RELATED and RELATED/SIMILAR because only these latter two have access to related knowledge, i.e. contacts that offer sufficient variety of knowledge but still allow for effective communication. Only one dummy (RELATED/SIMILAR) was positive and significant. The results meet our expectations as they show that having a mixture of links to related and technologically similar firms is conducive for innovation performance of firms. However, we also expected RELATED to be positive and significant, which is not the case.

In model III, only the dummies of the geographical proximity dimension are included. We expected firms with a mix of linkages at different spatial levels (MIX-DIST and REGIONAL/MEDIUM-DIST) to outperform firms that focus mainly on only one type of spatial level for interaction (REGIONAL, MEDIUM-DIST, LARGE-DIST). Only the MIX-DIST cluster was positive and significant. This confirms our hypothesis that a balance of links at all three spatial levels facilitates the innovation performance of firms.

[4] It might however also account for a potential negative size effect because, when excluding EMPL from the regression, RD_INT loses its significance.

In Model IV, the cluster dummies for the combined cognitive and geographical proximity dimensions are tested. Three of these have positive and significant effects (NO-REGIONAL&SIMILAR, RELATED, and SIMILAR). It suggests that innovation performance is particularly enhanced when firms have significant links with technologically similar organizations outside their region (NO-REGIONAL&SIMILAR), links with technologically similar organizations located at different spatial levels (SIMILAR), and linkages with technologically related organizations at various spatial scales (RELATED). A prominent observation is that no cluster dummies are significant with primarily regional linkages.

It is apparent that the geographical dimension plays only a minor role in defining these clusters. Therefore, we confront the dummies of the combined cognitive and geographical proximity dimensions with the two dummies that were found significant in the previous models (RELATED/SIMILAR, MIX-DIST). As a result, as shown in model V, NO-REGIONAL&SIMILAR, RELATED, and SIMILAR lose their significance. Their effect is captured by RELATED/SIMILAR. However, the previously significant MIX-DIST also loses its significance and regains significance when SIMILAR is removed (model VI). Although the overlap between the two clusters SIMILAR and MIX-DIST is small, their positive impact on innovation seems to be interrelate'. To disentangle their effects, we excluded all insignificant dummy variables and estimated model VII that includes only RELATED-SIMILAR, SIMILAR, and MIX-DIST. The RELATED/SIMILAR, SIMILAR, and MIX-DIST clusters remained highly significant, insignificant, and positive and significant, respectively. Therefore, the innovation performance of firms is enhanced when they have access to knowledge sources at different spatial scales or link with a combination of organizations that offer related and similar technological knowledge. This implies that the best practices facilitated by the geographical structure of knowledge exchange partners (with a combination of regional and non-regional linkages) is an alternative to the best practice facilitated by the cognitive structure (with a combination of cognitively similar and related organizations).

Conclusions

The paper investigated firms' embeddedness in knowledge networks. Taking a proximity approach, we focused on firms' knowledge exchange partners and the roles of cognitive and geographical proximities. Our investigation contributed to the literature in two ways.

First, we showed that the structure of firms' links differs systematically along the cognitive and geographical dimensions. With respect to the cognitive dimension, we found a large number of firms having relationships primarily with one type of organization (technologically dissimilar, related, or similar). Nevertheless, somewhat more than 50% of firms exchange knowledge with organizations possessing varying degrees of knowledge similarity. For the geographical dimension, similar findings were obtained with about 53% of firms having links to organizations at various spatial scales. We also found that the two proximity dimensions share a relationship. For younger firms and those with fewer technological knowledge links, geographical proximity helps to overcome cognitive distance. In contrast, geographical and cognitive proximities are complementary for smaller firms (but not younger firms), which implies that these firms tend to be well embedded in regional clusters of technologically similar firms. At the same time, they maintain significant relationships with dissimilar firms at larger geographical distances.

Second, we tested the effects of these different partner structures on innovation. The results clearly confirmed that firms increase their innovation performance by primarily linking with technologically related firms as well as technologically similar organizations. This is in line with Boschma and Frenken's findings (2010) and confirms the findings of other studies on this issue (cf. Broekel–Boschma 2012, Fornahl et al. 2011). In accordance with existing research, it also helps firms to access knowledge sources at various geographical scales (cf. Arndt–Sternberg 2000). However, in our study, we did not find an indication for the interrelated effects of cognitive and geographical proximity (in either a substitutive or a complementary sense) on firms' innovation performance.

This study has several shortcomings. We only measured links that are realized by the firms, not the interaction possibilities that firms may possess. For instance, the absence of technologically dissimilar organizations in a region fails to facilitate communication between such organizations at the regional level. Accordingly, they might be driven to search for such collaboration partners at greater geographical distances. The study does not consider that the existing potential for collaborating with a particular organization at a particular spatial level might influence the results on the relationship between cognitive and geographical dimensions for the establishment of knowledge links.

Our measure of geographical proximity might also have captured the effect of other types of proximity, like social proximity (Boschma 2005). In this study, we cannot entirely rule out this potential of spurious correlation. Therefore, we clearly see the need for further research to include data on all proximity types that were not considered in this paper, namely social, institutional, and organizational proximity. This would certainly increase our understanding of what drives the structure of the knowledge exchange partners of firms, whether proximity dimensions act as substitutes or complements, and the manner in which different structures of knowledge exchange partners influence the innovation performance of firms.

REFERENCES

ARNDT, O.–STERNBERG, R. (2000): Do manufacturing firms profit from intraregional innovation linkages? An empirical based answer *European Planning Studies* 8 (4): 465–485.

AZZALINI, A.–TORELLI, N. (2007): Clustering via nonparametric density estimation *Statistics and Computing* 17 (1): 71–80.

BALLAND, P.-A. (2011): Proximity and the evolution of collaboration networks: Evidence from Research and Development projects within the Global Navigation Satellite System (GNSS) industry *Regional Studies* 46 (6): 741–756.

BALLAND, P.-A.–SUIRE, R.–VICENTE, J. (2013): Structural and geographical patterns of knowledge networks in emerging technological standards: Evidence from the European GNSS industry *Economics of Innovation and New Technology* 22 (1): 47–72.

BATHELT, H.–MALMBERG, A.–MASKELL, P. (2004): Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation *Progress in Human Geography* 28 (1): 31–56.

BIN, R. D.–RISSO, D. (2011): A novel approach to the clustering of microarray data via nonparametric density estimation *BMC Bioinformatics* 12:49.

BOSCHMA, R. A. (2005): Proximity and innovation: A critical assessment *Regional Studies* 39 (1): 61–74.

BOSCHMA, R.–FRENKEN, K. (2010): The spatial evolution of innovation networks. A proximity perspective In: BOSCHMA, R.–MARTIN, R. (eds.) *The Handbook of Evolutionary Economic Geography* pp. 120–135., Edward Elgar, Cheltenham.

BRESCHI, S.–LISSONI, F. (2003): *Mobility and social networks: Localised knowledge spillovers revisited* CESPRI Working Paper, No. 142., Università Commerciale "Luigi Bocconi", Milano.

BROEKEL, T.–BOSCHMA, R. (2012): Knowledge networks in the Dutch aviation industry: the proximity paradox *Journal of Economic Geography* 12 (2): 409–433.

CANTNER, U.–GRAF, H. (2004): The network of innovators in Jena: An application of social network analysis *Research Policy* 35 (4): 463–480.

CASSIMAN, B.–COLOMBO, M. G.–GARRONE, P.–VEUGELERS, R. (2005): The impact of M&A on the R&D process: An empirical analysis of the role of technological- and market-relatedness *Research Policy* 34 (2): 195–220.

CRIBARI-NETO, F.–ZEILEIS, A. (2010): Beta regression in R *Journal of Statistical Software* 34 (2): 1–24.

FERRARI, S. L. P.–CRIBARI-NETO, F. (2004): Beta Regression for modeling rates and proportions *Journal of Applied Statistics* 31 (7): 799–815.

FORNAHL, D.–BROEKEL, T.–BOSCHMA, R. A. (2011): What drives patent performance of German biotech firms? The impact of R&D subsidies, knowledge networks and their location *Papers in Regional Science* 90 (2): 395–418.

FRENKEL, A.–SCHEFER, D. (1998): Local milieu and innovations: Some empirical results *The Annals of Regional Science* 32 (1): 185–200.

GLUCKLER, J. (2007): Economic geography and the evolution of networks *Journal of Economic Geography* 7 (5): 619–634.

GRAF, H. (2011): Gatekeepers in regional networks of innovation *Cambridge Journal of Economics* 35 (1): 173–198.

KAUFMAN, L.–ROUSSEEUW, P. J. (1990): *Finding groups in data* John Wiley and Sons Ltd. West Sussex, England.

LAZERSON, M. H.–LORENZONI, G. (1999): The firm that feed industrial districts: A return to the Italian source *Industrial and Corporate Change* 8 (2): 235–266.

Martin, R.–Moodysson, J. (2011): Innovation in Symbolic Industries: The Geography and Organization of Knowledge Sourcing *European Planning Studies* 19 (7): 1183–1203

MEYER-KRAHMER, F. (1985): Innovation behaviour and regional indigenous potential *Regional Studies* 19 (6): 523–534.

NOOTEBOOM, B. (2000): *Learning and innovation in organizations and economics* Oxford University Press, Oxford.

NOOTEBOOM, B.–VAN HAVERBEKE, W.–DUYSTERS, G.–GILSING, V.–VAN DEN OORD, A. (2007): Optimal cognitive distance and absorptive capacity Research Policy 36 (7): 1016–1034.

PHELPS, C. C. (2010): A longitudinal study of the influence of alliance network structure and composition of firm exploratory innovation *Academy of Management Journal* 53 (4): 890–913.

PONDS, R.–VAN OORT, F.–FRENKEN, K. (2007): The geographical and institutional proximity of research collaboration *Papers in Regional Science* 86 (3): 423–443.

POWELL, W. W.–WALTER, W.–KOPUT, K. W.–SMITH-DOERR, L. (1996): Interorganizational collaboration and the locus of innovation: Networks of learning in Biotechnology *Administrative Science Quarterly* 41 (1): 116–145.

SINGH, J. (2005): Collaborative networks as determinants of knowledge diffusion patterns *Management Science* 51 (5): 756–770.

SMITHSON, M.–VERKUILEN, J. (2006): A better lemon squeezer? Maximum-Likelihood Regression with Beta-Distributed Dependent Variables *Psychological Methods* 11 (1): 54–71.

STERNBERG, R.–ARNDT, O. (2001): The firm or the region: What determines the innovation behavior of European firms? *Economic Geography* 77 (4): 364–382.

TER WAL, A.–BOSCHMA, R. (2009): Applying social network analysis in Economic Geography: Framing some key analytical issues *Annals of Regional Science* 43 (3): 739–756.

TORRE, A.–RALLET, A. (2005): Proximity and localization *Regional Studies* 39 (1): 47–59.

UZZI, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect *American Sociological Review* 61 (4): 674–698.

APPENDIX

Table 1

Descriptives

Variables	Mean	Sd	Median	Min	Max	Skew	Kurtosis
INNO	0.29	0.31	0.20	0.00	1.00	1.04	–0.01
AGE	27.12	41.62	16.00	1.00	502.00	5.83	52.75
EMPLOY	220	2129	30.00	1.00	40877	18.72	361
HIGH	43.12	37.60	30.00	0.00	100.00	0.39	–1.46
ENG	48.19	40.16	50.00	0.00	100.00	–0.02	–1.67
RD_INT	0.25	0.31	0.08	0.00	1.00	1.11	–0.18
EXTERN	37.06	25.83	30.00	0.00	100.00	0.66	–0.11
CAR	0.16	0.36	0.00	0.00	1.00	1.89	1.64
BIO	0.19	0.39	0.00	0.00	1.00	1.61	0.64
AERO	0.19	0.39	0.00	0.00	1.00	1.57	0.50
TEXT	0.08	0.28	0.00	0.00	1.00	3.00	7.20
MEDIA	0.10	0.30	0.00	0.00	1.00	2.67	5,25
SOFT	0.08	0.27	0.00	0.00	1.00	3.14	8.04
VIDEO	0.05	0.23	0.00	0.00	1.00	3.94	13.86
FOOD	0.08	0.26	0.00	0.00	1.00	3.21	8.50
CZECH_BIO	0.04	0.20	0.00	0.00	1.00	4.49	18.56
SWED_BIO	0.08	0.27	0.00	0.00	1.00	3.07	7.61
LINKS_TEC	5.31	5.57	4.00	0.00	36.00	1.80	4.78
LINKS_MA	5.37	5.13	4.00	0.00	35.00	2.10	7.04

An improved radiation model and its applicability for understanding commuting patterns in Hungary

Levente Varga
Department of Physics,
Babeş-Bolyai University,
Cluj-Napoca, Romania
E-mail:
vargal@cs.ubbcluj.ro

Géza Tóth
Hungarian Central
Statistical Office,
Budapest, Hungary
E-mail:
geza.toth@ksh.hu

Zoltán Néda
Department of Physics,
Babeş-Bolyai University,
Cluj-Napoca, Romania
E-mail:
zneda@phys.ubbcluj.ro

Keywords:
human mobility models,
commuters data,
population income, and
density

Several empirical models aimed at describing human mobility have been proposed in the past. Most of them are based on an unjustified analogy, with a focus on gravity and physical vector or scalar fields. Recently, however, statistical physicists introduced a new category of models that are theoretically motivated by a few simple and reasonable socio-economic assumptions. *The Radiation Model* (Simini et al. 2012) and the *Radiation Model with Selection* (Simini–Maritan–Néda 2013) are such successful approaches. Here, we introduce a new version of the radiation model, the *Travel Cost Optimized Radiation Model*, and test its applicability for describing the commuting patterns in Hungary. We compare critically the performance of this model with the results of the previous radiation type models.

Introduction

Human mobility patterns form the focus of many studies in sociology and human geography (Varga et al. 2016), owing to their immediate impact on the economy (See literature review, for example, Ritchey 1976). Several electronically available large-scale and freely downloadable data sets help researchers unveil universalities and test their models (Barthélemy 2010, Bazzani et al. 2010). Recently, the statistical physics community has shown an interest in the impact of human mobility patterns on the economy, exploiting the fact that many models and modelling methods that are well known to physicists can be used for understanding the spatiotemporal universalities observed in human mobility (Brockmann–Hufnagel–Geisel 2006, González–Hidalgo–Barabási 2008).

The continuous growth of big cities (Bettencourt–West 2010), spatial concentration of industries around them, and rapid development of quick and affordable travelling mediums make commuting an increasingly important social phenomena, affecting a large portion of the population (Cohen et al. 2008). An understanding of the statistical laws governing the commuting flux serves both practical and scientific purposes. Uncovering some universally observable features will facilitate better planning for new travel routes, traffic optimisation, and prevention of failures in mobility fluxes. From the viewpoint of fundamental research, a successful modelling paradigm will reveal important psychological and socio-economic factors governing the job and housing selection of individuals.

Physics motivated the most popular models that attempted to fit the distribution of the mobility fluxes between the settlements. In this context, one can refer to the gravity model (for a review, please consult Barthélemy 2010) that along with Newton's law of gravitation exploits a phenomenological analogy. According to the assumption of the gravity models, the commuter fluxes between two settlements are proportional to their populations and inversely proportional with the square of the commuting distance. Generalized versions of the model propose power-law dependencies with freely defined fitting exponents for the population of the host and target cities and the distance between them. The generalized potential model is also an attempt to use the classical methodology of physics for understanding commuting fluxes (Anderson 1956, Lukermann–Porter 1960). Other well-known models that are not necessarily motivated by physics are the intervening opportunities model (Stouffer 1940) and the random utility model (Block–Marschak 1960). The radiation model, which was recently proposed by Simini et al., is built on a new paradigm and it starts by using a sociological driving force for the individuals—an optimization to improve their current income (Simini et al. 2012). The initial model, which in principle is a one-parameter model, was formulated for a discrete settlement structure. It was later generalized (Simini et al. 2013) for a continuous population distribution and made more complex by allowing a realistic job selection for the individuals. The new version of the model, the radiation model with selection, is a two-parameter model that offered an improved fit for the commuting data in the USA.

In the present work, we offer a new generalization of the radiation model by taking into account the travel cost involved in the job selection process. Our model (the travel cost optimised radiation model) leads again to a compact analytical formula involving two fitting parameters. Here, we test the applicability of this model for the fluxes derived from a complete commuting database in Hungary. We also compare the performance of this new model with the previously known radiation type models.

The model

The **original radiation model (RM)** starts from the assumption that commuting fluxes reflect, in some manner, optimisation reasoning in the jobseekers' choice for employment. The main hypothesis is that a jobseeker will accept the closest job offer that has a better salary than the one that it is currently available in his/her location. Owing to this requirement each individual travels to the nearest location where she/he can improve her/his current benefits. Let us assume that the cumulative distribution function of the incomes (benefits) is given by $p_{\leq}(z)$ (the probability that the income of one person is smaller or equal than z). Mathematically written, the main assumption of the original radiation model is that the probability $P_{>}(z \mid a)$ that an individual with benefit z refuses the closest a number of job offers is:

$$P_{>}(z|a) = p_{\leq}(z)^{a}, \tag{1}$$

In writing this probability, we assumed that the rejection of the a number of closest job offers with benefits less or equal to z are independent events. This is true in case there are in total more number of job offers than the number a.

Using the incomes probability density function, $p(z) = \frac{\delta p_{\leq}(z)}{\delta z}$ we can now calculate the $P_{>}(a)$ probability of not accepting the closest a number of job offers by integrating over the benefits:

$$\begin{gathered} P_{>}(a) = \int_{0}^{\infty} P_{>}(z|a)\, p(z)\, dz = \\ \int_{0}^{\infty} P_{>}(z|a) \frac{\delta p_{\leq}(z)}{\delta z} dz = \int_{0}^{1} p_{\leq}(z)^{a}\, dp_{\leq}(z) = \frac{1}{a+1} \end{gathered} \tag{2}$$

If we assume now that the jobseekers are selective in their choices and they are willing to accept better offers only with a probability 1–λ, we get (Simini et al. 2013):

$$P_{>}(z|a) = \left(1 - (1-\lambda)(1 - p_{\leq}(z))\right)^{a} = \left(\lambda + (1-\lambda)\, p_{\leq}(z)\right)^{a} \tag{3}$$

Alternatively, the assumption behind this generalization can be interpreted as a fact that the jobseekers are aware only of a fraction 1–λ of the available job offers. In such conditions, we obtain:

$$\begin{gathered} P_{>}(a) = \int_{0}^{\infty} \left[\lambda + (1-\lambda)\, p_{\leq}(z)\right]^{a} p(z)\, dz = \\ \int_{0}^{1} \left[\lambda + (1-\lambda)\, p_{\leq}(z)\right]^{a} dp_{\leq}(z) = \frac{1-\lambda^{a+1}}{(a+1)(1-\lambda)} \end{gathered} \tag{4}$$

This yields a more complex formula for the commuting fluxes, since an addition λ parameter appears. This model is known as the **radiation model with selection (RMwS)**. For λ =0, we recover the original radiation model. Simini et al. in 2013 showed that this model describes the distribution of the commuting fluxes for the USA better than the simple radiation model.

Subsequently, we introduced another extension for the original radiation model, which we named as the **travel cost optimized radiation model (TCORM).** The radiation model assumes that commuters will accept jobs that offer gains, relative to their present income, independently, on the distance between their residence and workplace. It presents a very weak condition for commuting, and our proposal is that individuals must choose to commute if there is a net gain in their salaries after subtracting the travel costs. However, travel costs depend on the distance travelled. Therefore, in principle, the radiation model has to be modified so that, in addition to the transited job offers *(a)*, the travel distance *(r)* is considered as an important factor. In this view, the working paradigm of the radiation model is reconciled with basic assumptions in the radiation and potential models. According to this assumption and *following the reasoning of the radiation model,* we can modify the starting hypothesis (1) as:

$$P_{>}(z|a) = p_{\leq}(z)^{a_0} \cdot p_{\leq}(z+\alpha r_1)^{a_1} \cdot \ldots \cdot p_{\leq}(z+\alpha r_q)^{a_q} \quad (5)$$

So that $\sum_{j=1}^{q} a_j = a$ (6)

In the above hypothesis, α is the parameter that governs the distance-dependent travel costs. We denoted the job openings inside settlement j by a_j and distance between originating point to the destination point inside the settlement by r_j.

According to this, in order to proceed we had to make some assumptions about the nature of the $p_{\leq}(z)$ cumulative income distribution. Definitely, $p_{\leq}(z)$ has to be a monotonically increasing function, with the condition $p_{\leq}(0) = 0$ and $p_{\leq}(\infty) = 1$. Empirical data on an estimated income distribution for Hungary shows that a kernel of the form $p_{\leq}(z) = 1 - e^{-\gamma z}$ gives a reasonable approximation (see Figure 1 where we have plotted $1 - p_{\leq}(z)$ on lognormal scale, suggesting a linear dependence). This form also satisfies the above-imposed conditions. Working with such kernel, one can write:

$$p_{\leq}(z+\alpha r_i)^{a_i} = (1 - e^{-\gamma(z+\alpha r_i)})^{a_i} \quad (7)$$

We assumed that the cost of travel is small in comparison to the salary, and hence $\alpha r_i << z$. It also means that $\alpha << 1$. In such case, we can make a Taylor expansion around the value of z:

$$\left(1 - e^{-\gamma(z+\alpha r_i)}\right)^{a_i} \approx \left(1 - e^{-\gamma z}\right)^{a_i} \left(1 + \alpha \gamma r_i a_i \frac{e^{-\gamma z}}{1 - e^{-\gamma z}}\right) \quad (8)$$

According to this, keeping only up to first order terms in α, we obtained:

$$P_{>}(z|a) \approx \left(1-e^{-\gamma z}\right)^{a}\left(1+\alpha\gamma\frac{e^{-\gamma z}}{1-e^{-\gamma z}}\sum_{j=1}^{q}a_{j}r_{j}\right)=\left(1-e^{-\gamma z}\right)^{a}\left(1+\alpha\gamma a\langle r(a)\rangle\frac{e^{-\gamma z}}{1-e^{-\gamma z}}\right) \quad (9)$$

Where, $\langle r(a)\rangle$ denotes the average distance of the a number of job openings from the centre.

Based on the above derivation, we expressed the $P_{>}(a,r)$ probability as:

$$P_{>}(a)=\int_{0}^{1} p_{\leq}(z)^{a}\left(1-\alpha\gamma a\langle r(a)\rangle\left(1-\frac{1}{p_{\leq}(z)}\right)\right)dp_{\leq}(z)= \int_{0}^{1} x^{a}\left(1-\alpha\gamma a\langle r(a)\rangle\left(1-\frac{1}{x}\right)\right)dx \quad (10)$$

Performing the integral, we obtained:

$$P_{>}(a)=\frac{1+\alpha\gamma\langle r(a)\rangle}{(a+1)} \quad (11)$$

As it was considered in previous work (Simini et al. 2012), we also assumed that job offers in a territory is linearly proportional with the population, i.e. $a=\mu w$, where w is the population and μ is a proportionality constant. According to this, we obtain again a two-parameter model, yielding the following with $\beta=\alpha\gamma\mu$:

$$P_{>}(w)=\frac{1+\beta\langle r(w)\rangle}{(\mu w+1)} \quad (12)$$

Thus, this gives another two-parameter model extension to the original radiation model.

In case of assuming a uniform distribution of the population in the considered territory, we obtained the following:

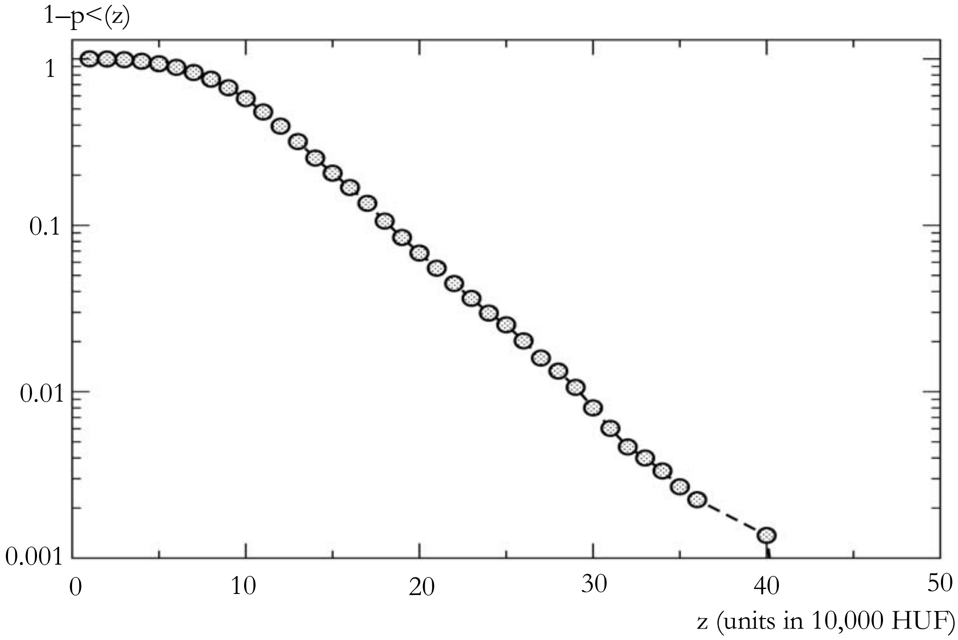
$$\langle r(w)\rangle \propto \sqrt{w} \quad (13)$$

which yielded the following:

$$P_{>}(w)=\frac{1+\lambda\sqrt{w}}{(\mu w+1)} \quad (14)$$

Where, λ is again proportionality constant.

Figure 1
Estimated cumulative distribution of incomes in Hungary in the year 2014 (please note the logarithmic vertical axes) The linear trend suggests an exponential decay



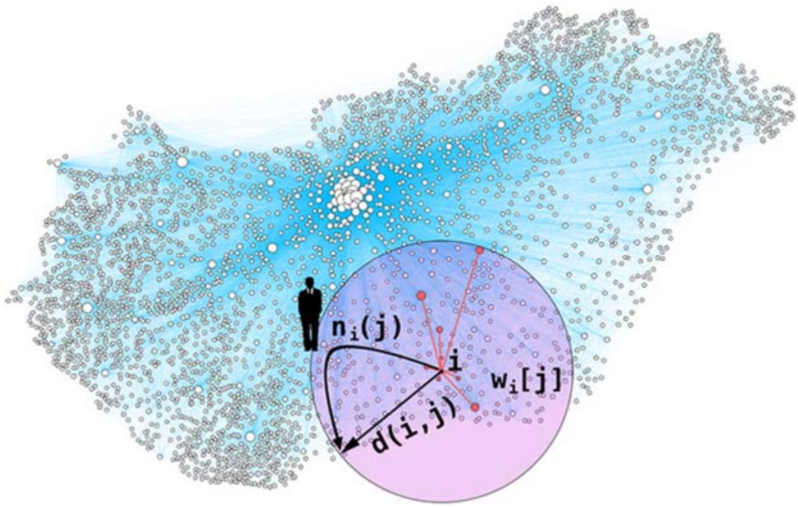
Source: HCSO.

Data analyses

As a basis for our experimental investigation, we processed a complete commuter and population database from Hungary. We analysed the 2011 population census data using ***Q*=3.176** settlements (nodes) (see white filled circles in Figure 2) and **81.664** commuter routes (edges) (see cyan lines between white filled circles in Figure 2) on the territory of Hungary. In our calculation, we also used data with a square kilometre resolution for population distribution in the census year 2011. The total population of Hungary, according to the 2011 census year, is **W=9.972.000**.

During our data processing, we select one by one the settlements i from where commuting starts and construct the boundaries of the disks with radius *d(i, j)*, reaching up to those settlements *j*, where the commuting is done (see Figure 2). We counted the population $w_i[j]$ inside the specified disk with radius *d(i, j)* and recorded the number of commuters $n_i(j)$ starting from settlement *i* and traveling to settlement *j*.

Figure 2
Settlements in the commuting network of Hungary. Here, we illustrate the data processing method. Disks of different radius *d(i, j)*, starting from a given settlement and reaching the other *j* settlements, are constructed. The population $w_i[j]$ inside these disks and the commuter number, starting from settlement *i* and traveling to settlement *j*, $n_i(j)$, is recorded.



Source: Edited by the authors.

With *d(i, j)*, $n_i(j)$, and $w_i[j]$ data for all the settlement pairs *(i, j)*, where $n_i(j) \neq 0$, we proceeded to the construction of the $P_>(w)$ probabilities. The number of workers that have their residence in settlement *i* are denoted by N_i. We have expressed this as follows:

$$N_i = \sum_{j=1}^{Q} n_i(j) \tag{15}$$

We denoted the total population in the considered territory as *W*.

We ordered the settlements according to their distance relative to *i*. Let $h_i^{[k]}$ be the index of the settlement that is the k^{th} one in this row (for example, $h_i^{[1]}$ is the index of the settlement that is the closest to settlement *i* and $h_i^{[2]}$ is the index of the settlement that is the second closest to *i*). Here, we introduced the *s(i, w)* integer number in the following manner:

$$s(i,w) \in \{1,2,3,...,Q\} \tag{16}$$

which is defined in the following manner for $w \leq W$:

$$w_i\left[h_i^{[s(i,w)-1]}\right] < w \leq w_i\left[h_i^{[s(i,w)]}\right] \tag{17}$$

(We record that Q denotes the total number of settlements and W the total population in the studied territory.) If no such number exists, then we will consider $s(i, w)=Q$.

In layman terms, $s(i, w)$ denotes the smallest number of settlements that have to be considered so that the total population is larger (or equal) than W inside the disk that contains these settlements (centred to i).

The probability that the commuters from i are a transitioning population W can be written as:

$$P_{>}^{i}(w)=\frac{1}{N_i}\left[N_i-\sum_{k=1}^{s(i,w)} n_i\left(b_i^{[k]}\right)\right] \tag{18}$$

The $\langle r_i(w)\rangle$ curve can also be constructed as:

$$\langle r_i(w)\rangle=\frac{\sum_{k=1}^{s(i,w)}\left[w_i\left[b_i^{[k]}\right]-w_i\left[b_i^{[k-1]}\right]\right]\cdot\frac{1}{2}\left[d(i,b_i^{[k]})+d(i,b_i^{[k-1]})\right]}{w_i[s(i,w)]} \tag{19}$$

After obtaining the quantities for all settlements, we constructed the desired curves as follows:

$$P_{>}(w)=\frac{1}{Q}\sum_{i=1}^{Q}P_{>}^{i}(w) \tag{20}$$

$$\langle r(w)\rangle=\frac{1}{Q}\sum_{i=1}^{Q}\langle r_i(w)\rangle \tag{21}$$

After constructing a data analyses algorithm based on the above calculations, we proceeded to construct experimentally the $P_{>}(w)$ and $\langle r(w)\rangle$ curves.

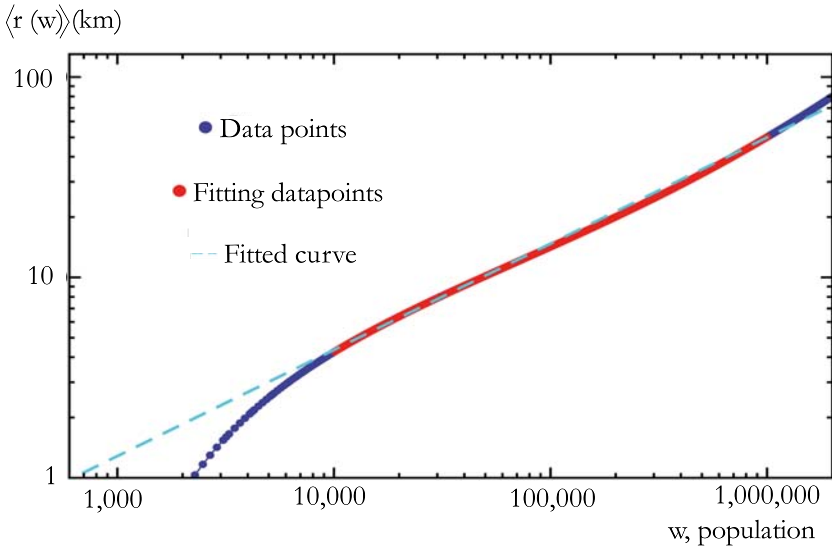
Results and discussions

The values for $P_{>}(w)$ computed from the census data are compared with the best fits obtained with the original radiation model (RM), the Radiation model with selection (RMwS), and the travel cost optimized radiation model (TCORM). The boundary effects become important for large w values (the disks centred on the settlements are incomplete due to the fact that they extend over the borders of Hungary). In addition, to minimise these effects we considered the data only up to w=1.000.000. This cut does not affect the commuters originating from Budapest, since the capital is not considered as a single settlement in this study, but is divided into sectors.

In order to use the TCORM model, we first had to investigate experimentally the $\langle r(w)\rangle$ function and check whether our assumption (13) holds. This is calculated by following the formulae (19) and (21) and it is plotted in Figure 3. The results reveal that the approximation (13) works well for distances that are larger than 3 km and smaller than 100 km (which is a relevant distance scale in commuting). If one would like to get an even better theoretical verification for formula (12), then the numerically computed $\langle r(w) \rangle$ could be used instead of (13).

Figure 3

The mean distance of a population *w*, in a disk-like geometry, relative to the centre of the disk (the mathematically defined $\langle r(w) \rangle$ as a function of *w*). The blue line is calculated using formulae (19) and (21) from census data for Hungary in year 2011. The red line shows the region where the power-law dependence is very good. The dashed light blue line shows a power-law trend corresponding to an exponent 0.53 (please note the logarithmic axes).



Source: Edited by the authors.

In order to verify the applicability of different variants of the radiation model we first compute the $P_>(w)$ experimental curve from the 2011 census data by following the steps specified in (18) and (20). We compare than visually the best fits to this line by using the:

RM prediction: $P_>^{RM}(w) = \frac{1}{\mu w + 1}$ (22) (one parameter fit)

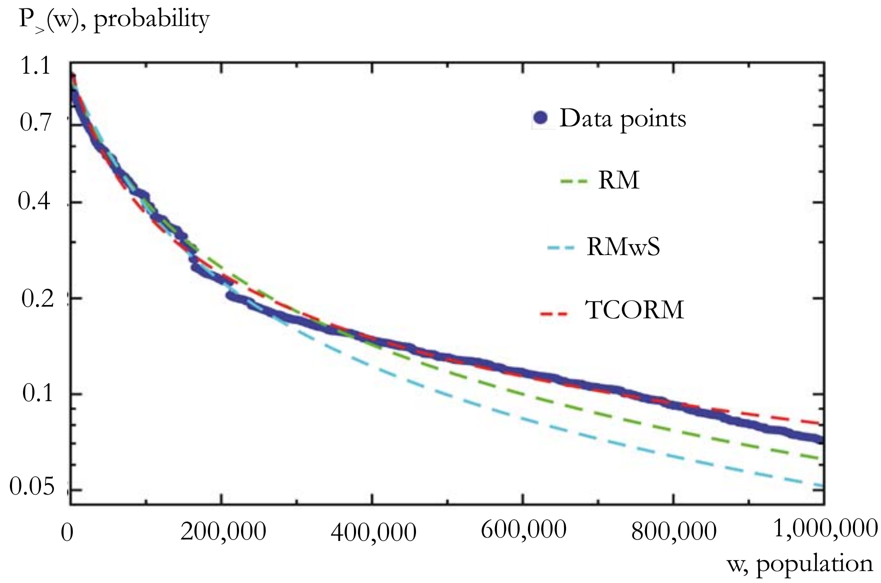
RMwS prediction: $P_>^{RMwS}(w) = \frac{1 - \lambda^{\mu w + 1}}{(\mu w + 1)(1 - \lambda)}$ (23) (a two parameter fit)

TCORM prediction: $P_>^{TCORM}(w) = \frac{1 + \lambda\sqrt{w}}{(\mu w + 1)}$ (24) (a two parameter fit)

The best fit to the experimental curve is done numerically and here we rely only on a visual comparison. The results are plotted on Figure 4. In Figure 4, we use a lognormal scale in order to visualize better the differences at small values.

Figure 4

Comparison of radiation type models prediction with experimental data for Hungary. The thick blue line $P_>(w)$ is calculated for commuters' data in Hungary using the 2011 census data and following the algorithm summarized by equations (18) and (20). These results are visually compared with the best fits obtained for the RM model (eq. 22), RMwS model (eq. 23) and TCORM model (eq. 24). Please note the logarithmic vertical axes.



Source: Calculated by the authors.

The best fits parameters obtained are as follows: μ=0.000015 for RM (eq. 22), μ=0.000028 and λ=0.33 for RMwS (eq. 23), and μ=0.00003 and λ=0.0015 for TCORM (eq. 24).

A visual examination also clearly indicates that the TCORM model performs better than the simple RM and RMwS model. The TCORM model has the capability of offering a visually good fit for the whole interval, while the simple RM model and RMwS are only capable of describing a portion of the $w \in [1,1.000.000]$ interval. It is apparent that a that a two-parameter fit is more successful than a one-parameter fit. Accordingly, it is also natural that TCORM would perform better than the simple RM. It is important to note that the TCORM seems to be a much better approximation than RMwS for describing commuter data for Hungary, although both the models are two-parameter regressions. However, future commuter data for other geographical regions must also be checked in order to affirm that TCORM is a better approximation for commuter fluxes than RMwS.

Conclusion

We studied the distribution of commuter fluxes between the settlements of Hungary using the census data from 2011. In order to explain the observed data, we derived a novel extension for the classical radiation model (RM). In our travel cost optimized radiation model (TCORM), we take into account the travel cost for commuting in the benefit optimisation. Realistically, where the travel cost is much smaller than the benefits, we obtained an analytical solution and a compact formula for the commuter fluxes. The complexity of this description is comparable with the complexity of the radiation model with selection (RMwS), wherein both the approaches focus on fitting a two-parameter model to experimental data. We find that our novel TCORM model performs better than the RMwS model in describing the commuting data for Hungary. This observation has to be thoroughly examined in the future for other geographical regions as well.

Acknowledgement

Work supported by the Romanian UEFISCDI grant PN-II-ID-PCE-2011-3-0348. The research of Z.N. was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TAMOP 4.2.4.A/2-11-1-2012-0001 'National Excellence Program'.

REFERENCES

ANDERSON, T. R. (1956): Potential models and the spatial distribution of population *Papers in Regional Science* 2 (1): 175–182.

BARTHÉLEMY, M. (2010): Spatial networks *Physics Reports* 499: 1–101.

BAZZANI, A.–GIORGINI, B.–RAMBALDI, S.–GALLOTTI, R.–GIOVANNINI, L. (2010): Statistical laws in urban mobility from microscopic GPS data in the area of Florence *Journal of Statistical Mechanics: Theory and Experiment* 2010: P05001.

BETTENCOURT, L.–WEST, G. (2010): A unified theory of urban living *Nature* 467: 912–913.

BLOCK, H.–MARSCHAK, J. (1960): Random orderings and stochastic theories of responses *Contributions to probability and statistics* 2: 97–132.

BROCKMANN, D.–HUFNAGEL, L.–GEISEL, T. (2006): The scaling laws of human travel *Nature* 439: 462–465.

COHEN, J. E.–ROIG, M.–REUMAN, D.C.–GOGWILT, C. (2008): International migration beyond gravity: A statistical model for use in population projections *Proceedings of the National Academy of Sciences* 105 (40): 15269–15274.

GONZÁLEZ, M.C.–HIDALGO, C.A.–BARABÁSI, A. L. (2008): Understanding individual human mobility patterns *Nature* 453: 779–782.

LUKERMANN, F.–PORTER, P.W. (1960): Gravity and potential models in economic geography *Annals of the Association of American Geographers* 50 (4): 493–504.

RITCHEY, P. N. (1976): Explanations of migration *Annual Review of Sociology* 2: 363–404.

SIMINI, F.–GONZÁLEZ, M.C.–MARITAN, A.–BARABÁSI, A.L. (2012): A universal model for mobility and migration patterns *Nature* 484: 96.

SIMINI, F.–MARITAN, A.–NÉDA, Z. (2013): Human Mobility in a Continuum Approach. *PLoS ONE* 8(3): e60069.

STOUFFER, S.A. (1940): Intervening opportunities: a theory relating mobility and distance *American Sociological Review* 5 (6): 845–867.

VARGA, L.–KOVÁCS, A.–TÓTH, G.–PAPP, I.–NÉDA, Z. (2016): Further We Travel the Faster We Go *PLoS ONE* 11 (2): e0148913.

Commuting patterns in Romania: Case study on Cluj County

József Benedek
Babeş-Bolyai University,
Faculty of Geography,
Cluj-Napoca, Romania
E-mail:
jozsef.benedek@ubbcluj.ro

Iulia Hărănguş
Babeş-Bolyai University,
Faculty of Geography,
Cluj-Napoca, Romania
E-mail:
iulia_harangus@yahoo.com

Titus Man
Babeş-Bolyai University,
Faculty of Geography,
Cluj-Napoca, Romania
E-mail:
titus.man@.ubbcluj.ro

The study examines the spatial and economic characteristics of commuting to work in one of the most dynamic areas of Romania, Cluj County. Based on the 2011 census data, the study reveals a strong connection between accessibility and commuting intensity, while the urban network determines the spatial orientation of the dominant commuting flows. However, we found no significant relation between dynamic economic performance and commuting intensity.

Keywords:
commuting to work,
accessibility,
economic development,
Cluj County

Introduction

During socialism, apart from the Central and Eastern European (CEE) countries, a restricted and controlled spatial mobility regime was established in Romania, characterised by a low motorisation rate and migration (Sandu 1984). Under these circumstances, commuting was reduced to the closer suburban areas of the large urban agglomerations. While at the beginning of the transition period a large amount of the active population was employed in agriculture, the establishment of capitalist market structures determined a radical change in the structure of economic activities, which has influenced the intensity of the commuting. As a result of the post-socialist economic restructuration, the workplaces became spatially more concentrated. However, housing is still strongly dispersed, partly due to the inherited settlement network (more than 13,000 localities) and partly due to on-going suburbanization (Soaita 2013). There are two ways of overcoming this inconvenience: migration and commuting. The main advantage of commuting is related to the satisfaction of different preferences offered by the proximity between workplace and residential location The commute between rural residential locations and urban workplaces is seen as a striking form of urban-rural integration (Partridge et al. 2010). In addition, labour can easier follow the change on the workforce market. However, commuting may cause major difficulties due to pollution, traffic jams, and costs.

However, the Romanian population still has an overall low spatial mobility–both internal migration and commuting have a low intensity, even by regional comparison. Characterized by a strong economic growth, although the period 2002–2011 has witnessed an increase in the number of active population, the inner-county commuting has registered a small decrease (table 1). It means that the strong economic growth was not enough to trigger a significant increase in commuting. In fact, we have not found any significant correlation neither between the county level GDP per capita and the commuting rate nor between the county level economic growth rate and the commuting rate, contrary to the empirical findings of international literature (Östh and Lindgren 2012).

In parallel, both inter-county and international commuting have almost doubled. It implies that there is little change in commute over short distances, but – somehow surprisingly – an important change in both middle- and long-distance commutes. The strong spatial concentration of economic growth and reducing importance of political borders for labour mobility in trans-border regions in Romania explains the latest situations in the country, which is forcing people from peripheries to travel longer distances for work (Nagy 2012). However, overall the commuting rate in 2002 was only around 16.8% and approximately 21% in 2011. This relative low growth of internal mobility is compensated by the strong international mobility of the Romanian population, with an estimated 2–3 billion Romanians being registered in different EU countries, mainly in Italy and Spain (Boboc et al. 2012). It means that more than 10% of the total population of Romania in 1992 has left the country, a particular situation in the present European context, which may explain the lower figures of internal mobility.

Table 1

Commuting categories in Romania

Year	Active population	Workplace in other localities of the county		Workplace in other county		Workplace in other country	
			%		%		%
2002a)	7,811,733	1,000,798	12.81	187,176	2.40	124,248	1.59
2011b)	8,507,759	1,086,614	12.77	470,963	5.53	247,270	2.91

a) http://www.insse.ro/cms/files/RPL2002INS/vol2/tabele/t31(576–584).pdf
b) www.recensamantromania.ro/wp-content/uploads/2015/05/vol3_t33.xls

In this broader spatial and national context, the main aim of the paper is to analyse commuting in one of the economically wealthiest areas of Romania, Cluj County (NUTS-3), situated in the north-western part of Romania. We assume that commuting plays an important role in dynamic economics, and that economic development, accessibility, and commuting intensity are interconnected.

Data and Methodology

The study relies largely on data from 2011 census, which is the only source of data on commuting in Romania. For the general context, we compared commuting data on county level (județ in Romanian, corresponding to the Nomenclature of Territorial Units for Statistics (NUTS) 3 level), while the main part of the study focuses on commuting data at local level (towns and communes, corresponding to the NUTS 5 level). The focus is on spatial differentiation in commuting, therefore we extensively visualised the analysed results. In addition, we analysed the commuting characteristics recorded in the census data, which include the intensity and spatial orientation of commuting flows and economic profile of the commuters. Although we do not intend to go beyond the rather descriptive and exploratory nature of this study, we assume that commuting is strongly connected with economic development and accessibility. Concerning economic development, we provided only qualitative and descriptive information at the local level, while the GDP per capita is used for county level economic comparisons. For determining the accessibility of population to the urban centres of the examined county, particularly its main centre, Cluj-Napoca (Kolozsvár in Hungarian), we used the function of the GIS extension Network Analyst, 'Closest Facility'. Accordingly, for each locality, the shortest travel time was determined using the road and/or railway network. For the road network, we considered the average travel speed calculated in Rusu et al. (2013): 110 km/h for highways, 70 km/h national and European roads, 50 km/h for county roads, and 30 km/h for local roads. The average speed for train travels was determined through the function according to the travel times between the railway stations, as given in the official circulation program of the Romanian Railways.

Spatial and economic characteristics of commuting in Cluj County

In the general context described in section 1, we expected to have a different situation for Cluj County, which registered one of the highest economic growth rates in Romania. This dynamic economic performance in the area and the improvements realised in the transport infrastructure have influenced active commuting in the county as compared to Romania. Despite the fact that the commuting rate of 21.8% is only slightly above the country average, a careful observation of the commuting categories in each county (Table 2) for 2011 provides a different picture: middle- and long-distance commuting (measured by the share of population with workplaces in other counties and the share of population with workplaces in other countries, respectively) are common for peripheries, while in the counties with highest economic performance, which is the case of Cluj County, short distance commuting (expressed as the share of population with workplaces in other localities of the county) is more significant. Particularly, we observed a short distance commuting rate of 17.5% for Cluj County, which is still low in international comparison but is well above the national average. The value of short-distance commuting places Cluj County in the 13th position (among 41 counties) compared to Romania, which is well below its economic position (4th position in the GDP per capita ranking). It is due to the stronger spatial concentration of population in the county's centre, Cluj-Napoca, (comprising 66% of the county's population) and in its larger urban influence zone (comprising 75% of the county's population), which reduces the commuting intensity.

Table 2

Commuting categories in the Romanian counties (NUTS-3 units) in 201

County	Active pop.	Pop. with work-place in other localities of the county	%	Pop. with work-place in other county	%	Pop. with work-place in other country	%	Comm. rate %
Alba	141,389	25,673	18.16	5,906	4.18	2,195	1.55	23.89
Arad	176,087	37,544	21.32	5,328	3.03	3,865	2.19	26.54
Argeș	283,772	67,113	23.65	10,924	3.85	3,513	1.24	28.74
Bacău	269,136	32,297	12.00	9,038	3.36	11,686	4.34	19.70
Bihor	240,521	47,933	19.93	4,553	1.89	4,482	1.86	23.68
Bistrița-Năsăud	133,797	15,353	11.47	3,261	2.44	4,765	3.56	17.47
Botoșani	192,647	9,615	4.99	5,924	3.08	10,915	5.67	13.74

(Table continues the next page.)

(Continued.)

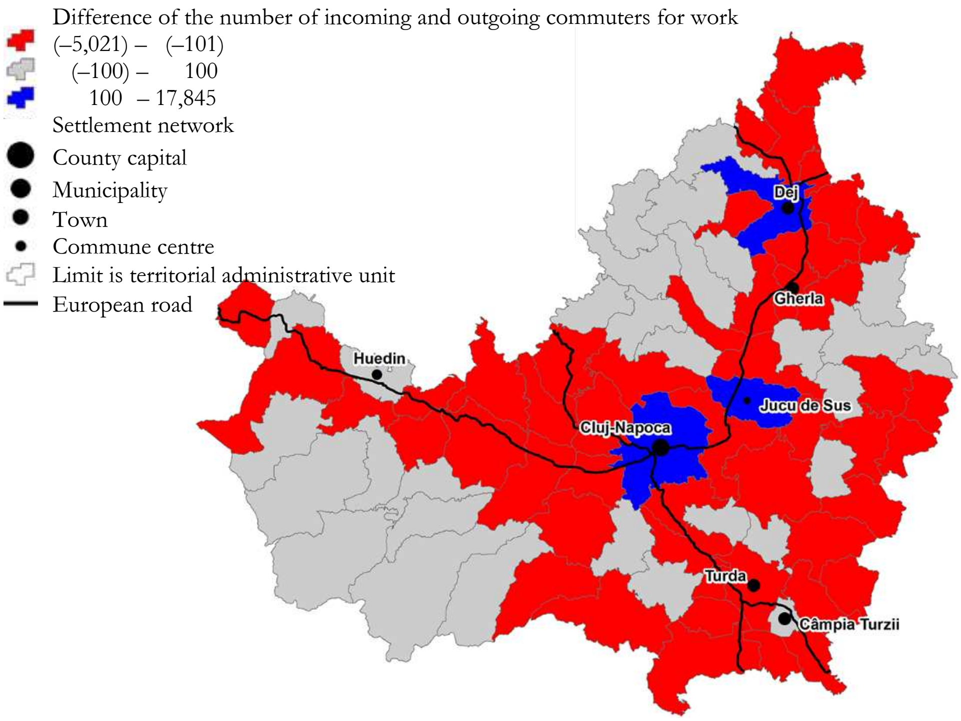
County	Active pop.	Pop. with work-place in other localities of the county	%	Pop. with work-place in other county	%	Pop. with work-place in other country	%	Comm. rate %
Brăila	135,648	7,223	5.32	6,546	4.83	2,007	1.48	11.63
Braşov	215,485	42,065	19.52	7,326	3.40	4,616	2.14	25.06
Buzău	192,681	26,752	13.88	10,214	5.30	3,738	1.94	21.12
Călăraşi	128,771	10,127	7.86	20,801	16.15	2,277	1.77	25.78
Caraş-Severin	107,092	12,374	11.55	5,708	5.33	4,397	4.11	20.99
Cluj	270,972	47,551	17.55	6,870	2.54	4,640	1.71	21.8
Constanţa	255,767	50,080	19.58	9,405	3.68	7,212	2.82	26.08
Covasna	90,138	11,812	13.10	5,719	6.34	3,682	4.08	23.52
Dâmboviţa	237,178	38,574	16.26	28,350	11.95	15,907	6.71	34.92
Dolj	263,356	22,232	8.44	7,276	2.76	7,232	2.75	13.95
Galaţi	230,444	15,818	6.86	7,855	3.41	7,803	3.39	13.66
Giurgiu	120,474	6,229	5.17	24,652	20.46	852	0.71	26.34
Gorj	149,169	29,141	19.54	6,480	4.34	2,544	1.71	25.59
Harghita	123,217	20,809	16.89	3,388	2.75	4,570	3.71	23.35
Hunedoara	173,577	36,786	21.19	6,387	3.68	3,356	1.93	26.8
Ialomiţa	104,771	11,063	10.56	14,804	14.13	1,746	1.67	26.36
Iaşi	341,850	36,528	10.69	9,499	2.78	15,938	4.66	18.13
Ilfov	166,005	15,203	9.16	83,353	50.21	601	0.36	59.73
Maramureş	198,081	24,588	12.41	4,147	2.09	17,923	9.05	23.55
Mehedinţi	114,438	10,115	8.84	5,910	5.16	2,011	1.76	15.76
Bucureşti	893,501	–	–	40,386	4.52	4,268	0.48	5.00
Mureş	208,841	38,036	18.21	6,769	3.24	7,525	3.60	25.05
Neamţ	189,466	24,705	13.04	5,667	2.99	9,614	5.07	21.10
Olt	182,011	15,451	8.49	9,432	5.18	4,834	2.66	16.33
Prahova	311,289	79,400	25.51	27,039	8.69	4,539	1.46	35.66
Sălaj	96,427	12,003	12.45	4,265	4.42	3,254	3.37	20.24
Satu Mare	129,828	21,028	16.20	4,033	3.11	11,190	8.62	27.93
Sibiu	161,275	29,986	18.59	4,671	2.90	3,503	2.17	23.66
Suceava	277,536	30,859	11.12	4,616	1.66	17,928	6.46	19.24
Teleorman	157,810	12,447	7.89	15,309	9.70	2,488	1.58	19.17
Timiş	290,393	52,635	18.13	6,805	2.34	4,673	1.61	22.08
Tulcea	91,140	7,173	7.87	4,184	4.59	3,303	3.62	16.08
Vâlcea	148,145	25,844	17.45	7,223	4.88	3,132	2.11	24.44
Vaslui	170,448	10,391	6.10	7,018	4.12	3,976	2.33	12.55
Vrancea	143,186	16,058	11.21	3,922	2.74	8,570	5.99	19.94

Source: Own compilation based on census data.

The difference between absolute incoming and outgoing commuter numbers in Figure 1 illustrates the balance in commuting at the NUTS 5 level (cities and communes). The cities Cluj-Napoca, Dej, and the Jucu commune registered the highest positive balance. Both Cluj-Napoca and Dej underwent reindustrialization during the 2000s and witnessed the establishment of new industrial parks, and consequently became home to important industries. In addition, the counties have developed a strong service sector, based on their regionally important demographic size (Cluj-Napoca and Dej comprise 324,000 and 31,000 inhabitants, respectively). The situation of Jucu (4,300 inhabitants) is explained by the spatial distribution of the new economy, which transformed the commune. As a result, Jucu became the only rural location to be selected for the development of one of the six new industrial parks planned for the county. The commune witnessed the launch of Tetarom III in 2008, generating around 2000 new jobs (Benedek et co. 2013). Generally, all cities have shown a positive balance, with exception of Gherla (21,000 inhabitants) and Turda (48,000 inhabitants), which are situated in the influence zone of Dej and Cluj-Napoca.

Figure 1

The spatial distribution of absolute commuting balance



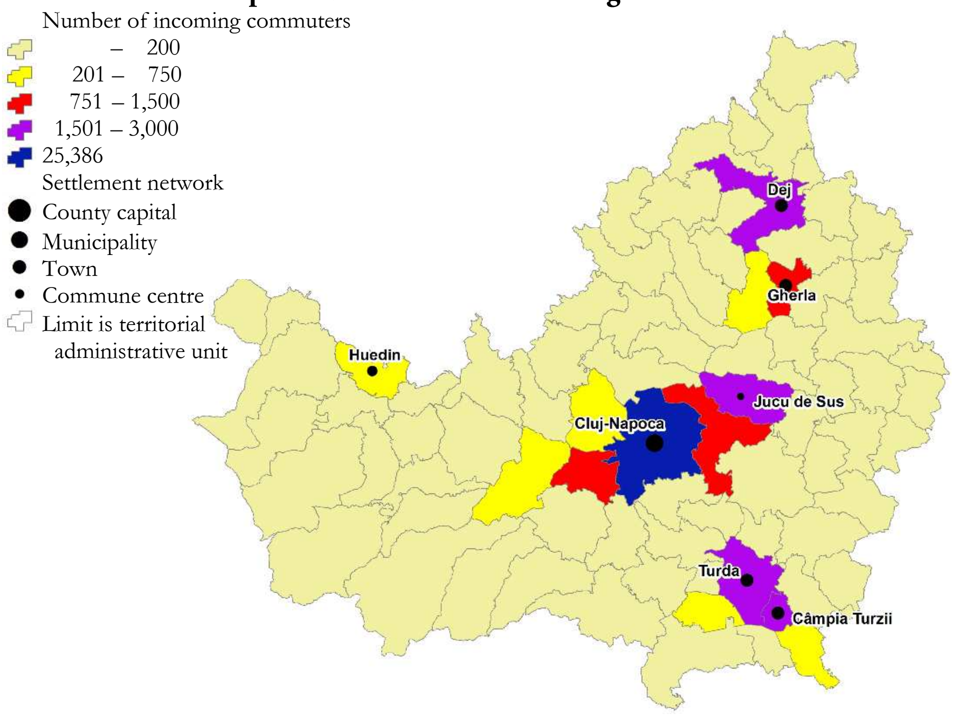
Source: Own draft, based on 2011 census data.

On the opposite site, the suburban commune Floreşti has the highest negative balance. Although its population growth is impressive (the highest in Romania, reaching 21,000 inhabitants in 2011), it has not been successful in developing employment opportunities. These rural communes have become a dormitory area of Cluj-Napoca. High values of negative commuting balance are also registered by other suburban communes of Cluj-Napoca (Baciu, Apahida, and Aghireş) and two cities (Gherla and Turda). These cities have good transport connectivity, and therefore provide opportunities for commuting to the main concentration areas of the workforce represented by Cluj-Napoca, Dej, and Jucu.

The high negative commuting balance of large rural areas (indicated in orange in Figure 1) might be related to their high accessibility levels, while the three compact peripheral rural areas (shown in green in Figure 1) are rather indifferent to commuting, registering low values of absolute commuting (between 100 and −100 commuters).

Figure 2

The spatial distribution of incoming commuters



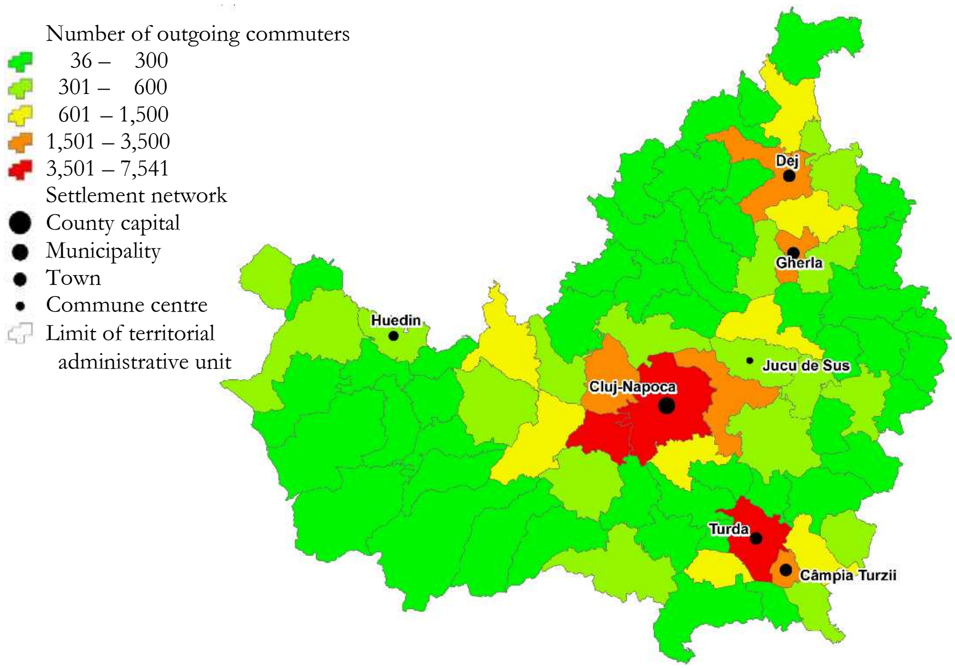
Source: Own draft, based on 2011 census data.

Figures 2 and 3 depict the two main components of the commuting balance: the incoming and outgoing commuting, in absolute numbers. Cluj-Napoca is the leading destination for incoming commuters, attracting around 25,000 commuters daily, followed by Dej (2,800 commuters), Jucu (2,200 commuters), and the urban agglomeration formed by the neighbouring cities of Turda and Câmpia Turzii (4,400 commuters together). In fact, all the cities are important attractors of incoming commuters; in addition, the suburban communes like Florești (1,500 commuters, more than cities like Gherla or Huedin), Apahida, and Baciu all attracting an important number of commuters.

The picture is slightly different concerning the outgoing commuters. Cluj-Napoca leads with around 7,500 outgoing commuters; closely followed by Florești (6,500), and at some distance by Turda (3,700); Apahida and Baciu (each with 2,600 commuters); and Dej, Gherla, and Câmpia Turzii (each with around 2,000 commuters). It implies that we cannot define the dominant unidirectional commuting flows; they are rather oriented in both directions inside the economically active areas like the urban agglomerations of Cluj (including the three suburban communes), Turda-Câmpia Turzii, Dej, and Gherla. The only exceptions are the Jucu and Aghireșu communes. While incoming commuters flow is dominant in Jucu due to its industrial park, the outgoing commuters flow is dominant in Aghireșu. The overall findings reflect the profound changes determined by the intense post-socialist socio-economic transformation, typical commuting groups including the suburbanized social middle-class groups around the Cluj County, well-educated employees commuting from cities to rural areas, or well-qualified workers commuting to the new industrial parks from Cluj, Jucu, or Dej.

Figure 3

The spatial distribution of outgoing commuters

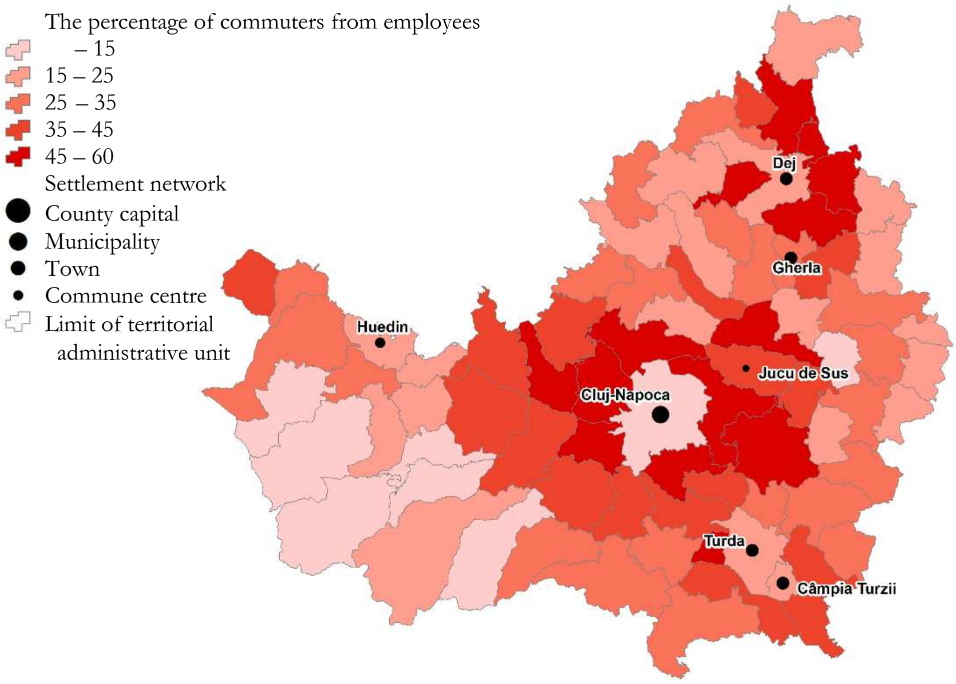


Source: Own draft, based on 2011 census data.

Figure 4 presents the spatial distribution of relative commuting. It implies that the share of commuters form the total number of employees in each town and commune. The scenario differs in details when compared to the absolute commuting. Cluj-Napoca, with the largest number of absolute commuters, has the lowest commuting rate due to its large employee base, totalling to 143,000 employees, which represents 53% of the total number of employees in Cluj County. This situation and explanation is also valid for the other two major commuting areas represented by Dej and Turda-Câmpia Turzii.

Figure 4

Relative commuting in Cluj County



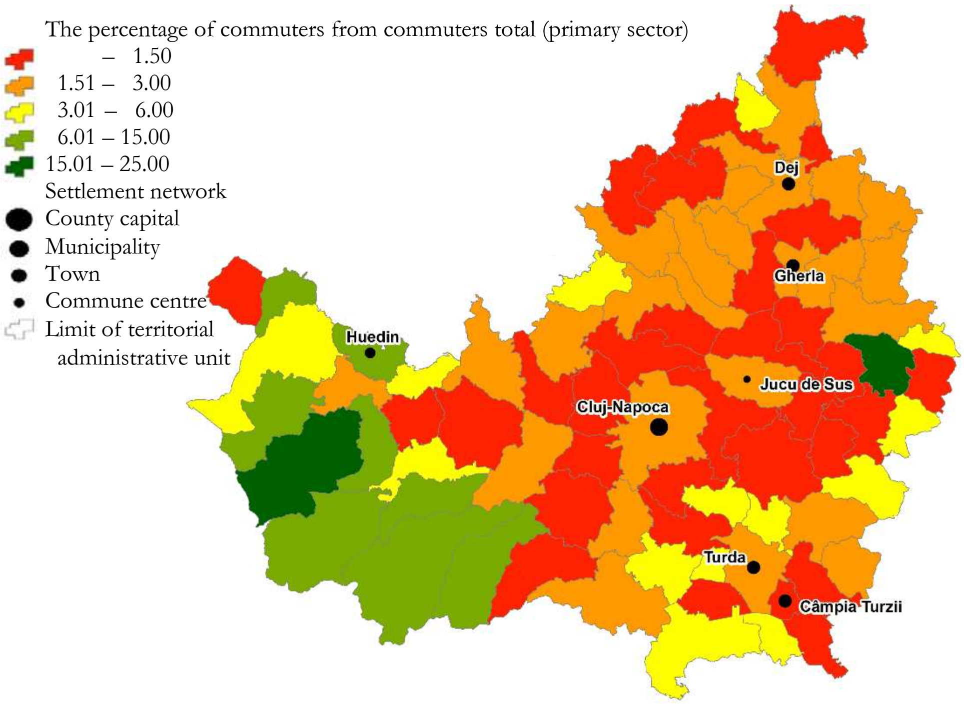
Source: Own draft, based on 2011 census data.

Unlike in the case of relative commuting, the role of Cluj-Napoca as a major organiser of the commuting flows in the county is expressed by the disposal of two commuting belts around it, with the rate of commuting exceeding 45%, respective between 35–45%, where the intensity of commuting decreases with distance and accessibility. The northern and southern parts of the county, the city of Dej and the urban agglomeration Turda-Câmpia Turzii, respectively, are playing a similar role, but on a lower scale and intensity.

Owing to the radical shift in the economic structure of the county from agriculture and industry towards services, the service activities are attracting a large number of commuters. Commuting in industry still plays an important role as a result of the reindustrialisation in the 2000s, while commuting in agriculture shows low values in the county, with the notable exception of the western, mountainous part of the county (Figure 5). The urban centre of this area, Huedin (9,300 inhabitants), has one of the largest primary sector (5%) among the cities in Cluj county, overtaken only by Dej (6%).

Figure 5

The spatial distribution of commuters in primary activities

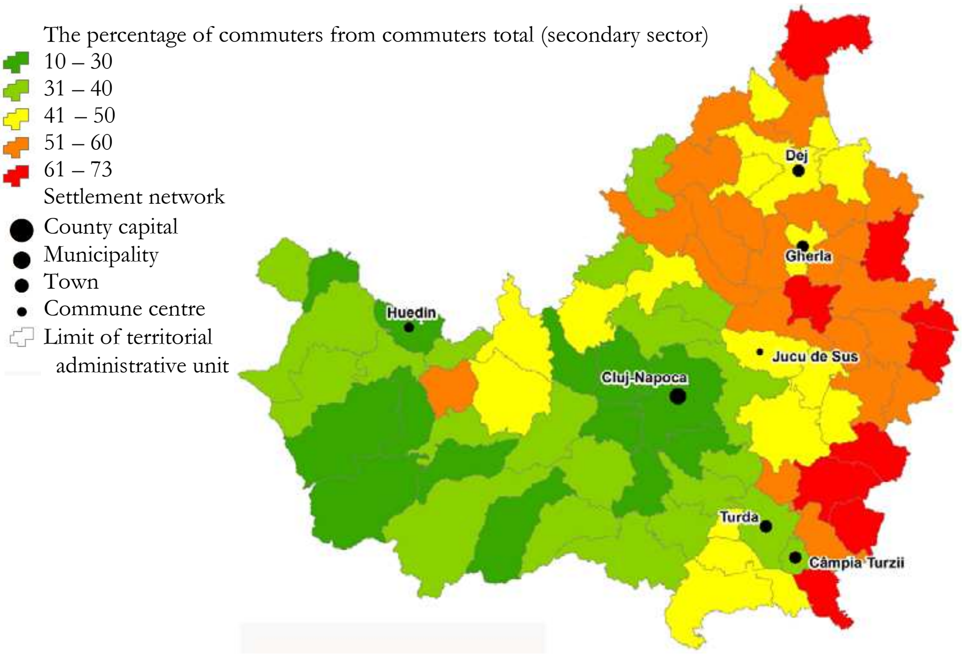


Source: Own draft, based on 2011 census data.

Figures 6 and 7 reveal that the specialisation of commuters is highly adapted to the economic profile of the main commuter attracting areas. Therefore, commuters in the larger suburban zone of Cluj-Napoca have workplaces engaged in different service activities, corresponding to the high share of services in the economic profile of the city. Annex 1 provides additional information on the professional profile of commuters compared to the professional categories of the attractor localities. Approximately, 75% of the workforce in Cluj-Napoca is engaged in the services sector, with IT, retail trade, education, and health services being the largest employers of this sector. The people commuting to Cluj-Napoca have the largest share of employees engaged in retail, construction, and transportation, while the employees in IT, education, and health services have smaller proportions in comparison to Cluj, reflecting the lower education and status of the commuters.

Figure 6

The spatial distribution of commuters in industry

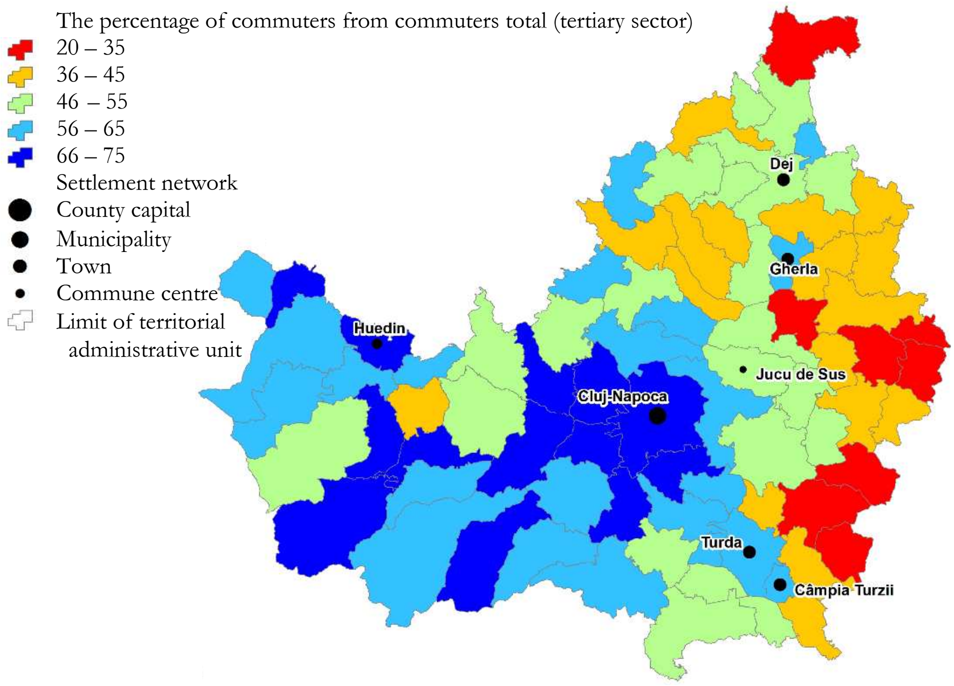


Source: Own draft, based on 2011 census data.

Commuters in the northern and southern part of the county, attracted by the labour market of the urban centres, Dej, Gherla, and Turda-Câmpia Turzii, are mainly occupied in industrial activities, according to the high industrial profile of these cities. Annex 1 also provides detailed data on this aspect. For example, in the case of commuters to Gherla, the largest group (20%) of commuters consists of employees engaged in the manufacture of wood products. In the case of commuters to Câmpia Turzii, employees engaged in the metal manufacture activity represent the largest group (14%), while for Dej and Turda we have the same economic profile of commuters as in the case of Cluj-Napoca. Câmpia Turzii belongs to the few cities in Romania where the largest amount of local workforce is employed in the industrial sector (50%) due to the small demographical size of the city (22,000 inhabitants) and the dominant position of a single large industrial factory in the local labour market.

Figure 7

The spatial distribution of commuters in the service sector



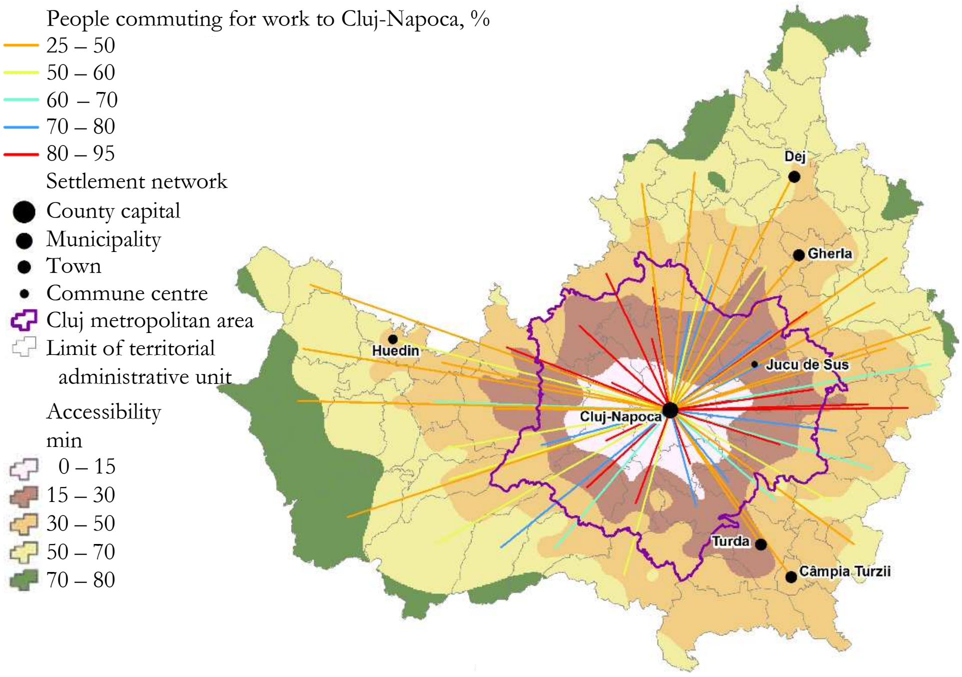
Source: Own draft, based on 2011 census data.

Figure 8 presents arguments in favour of the connection between accessibility and commuting intensity. The area delimited by the isochrone of 30 minutes travel time registers the highest commuting intensity. In addition, we have three cases where the isochrones of 30–50 minutes do not represent an obstacle for intensive commuting to Cluj-Napoca. It might be attributed to the fact that, during the post-socialist transition, the means of commuting have changed. There is a shift from bus and train to private car, which may lead to an increase in the commuting distance.

The same figure reveals that the Metropolitan Zone of Cluj-Napoca, created as an association with an aim to promote integrated spatial planning and development, functions like a highly integrated labour market, whose limits can probably be even extended based on commuting intensity and accessibility.

Figure 8

Accessibility and commuting oriented to Cluj-Napoca

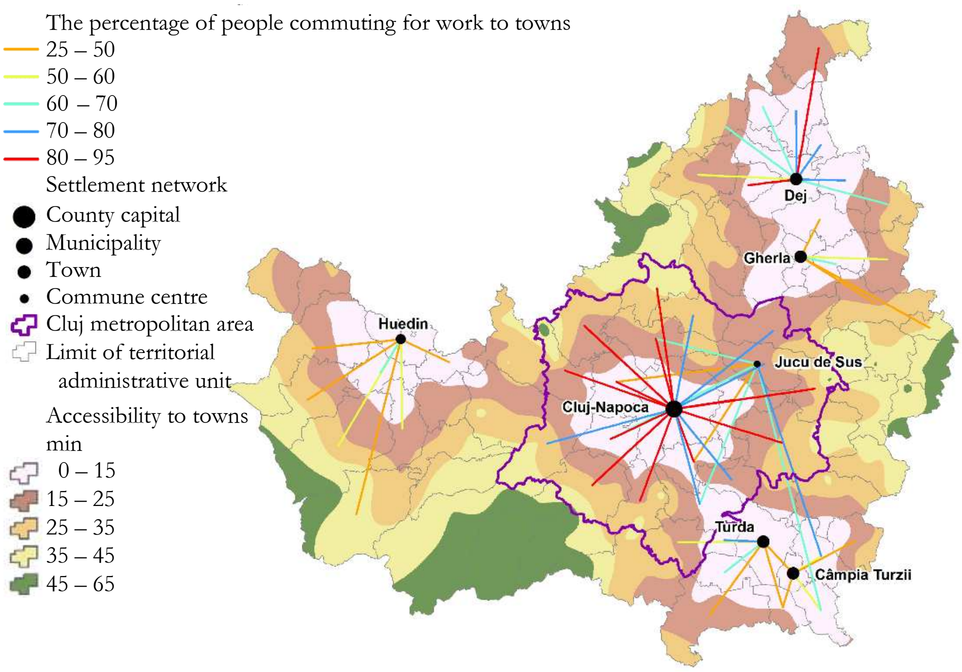


Source: Own draft, based on 2011 census data.

When considering the relationship between accessibility and commuting to all cities of the county, three secondary commuting systems around Turda-Câmpia Turzii, Gherla-Dej, and Huedin became evident (Figure 9). Intensive commuting could be identified only in the cases of Cluj-Napoca and Dej, while long-distance commuting was present in case of Huedin and Gherla (isochrones of 35–45 minutes), revealing their particular role as local workplace suppliers and in geographical space configuration. Figure 9 displays the particular role played by the commune of Jucu to attract significant commuting flows from both the suburban area of Cluj and Turda-Câmpia Turzii. As a general picture, Figure 9 reflects the delimitation of the urban influence zones in relation to commuting intensity and accessibility. Facilitated by short distance and the good accessibility, the overlapping tendencies between the commuter zone of Cluj-Napoca and that of Turda are evident, while longer distance has favoured the maintenance of the local urban influence zones in the cases of Huedin and Gherla. The peripheral areas situated between the isochrones of 45–65 minutes show that an insignificant amount of commuting has been left outside the functional urban labour markets. It can be viewed as the threshold where migration decisions may prevail to commuting.

Figure 9

Accessibility and commuting oriented to Cluj-Napoca



Source: Own draft, based on 2011 census data.

Conclusions

The study has produced some interesting results on present day commuting to work in a dynamic economic environment, illustrated through the Cluj County. First, although we were unable to measure the economic output at a local level accurately, the county level comparison of GDP per capita and commuting rate did not reveal any significant correlation between commuting and economic development, contrary to the findings in international literature. We have explained this particular situation in the context of intensive international migration. Second, we found a similar situation at the county level. In this context, we were able to provide qualitative information on the location of new industrial parks in relation with the establishment of spatial commuting attractors. We also presented new facts on the relationship between accessibility and commuting, the economic profile of commuters, and the important role played by settlement networks in the spatial orientation of commuting flows.

REFERENCES

BENEDEK, J.–CRISTEA, M.–LANG, T. (2013) Fallstudienbericht Rumänien In: Müller, A.–Plöger, J. (eds.): *Wieder erstarkte Städte: Strategien, Rahmenbedingungen und Ansätze der Regenerierung in europäischen Groß- und Mittelstädten* pp. 16–30., (Werkstatt: Praxis, Heft 82)., Bundesministerium für Verkehr, Bau und Stadtentwicklung, Bundesamt für Bauwesen und Raumordnung, Berlin.

BOBOC, C.–VASILE, V.–TODOSE, D. (2012) Vulnerabilities associated to migration trajectories from Romania to EU countries *Procedia – Social and Behavioral Sciences* 62: 352–359.

NAGY, E. (2012) The Situation of the Workforce Market in the Southern Part of the Romanian-Hungarian Border *Studia Universitatis „Babeş – Bolyai", Series Geographia* 57 (2): 153–160.

ÖSTH, J.–LINDGREN, U. (2012): Do changes in GDP influence commuting distances? A study of Swedish commuting patterns between 1990 and 2006 *Tijdschrift voor Economische en Sociale Geografie* 103 (4): 443–456.

PARTRIDGE, M. D.–ALI, K.–OLFERT, M., R. (2010) Rural-to-urban commuting: three degrees of integration *Growth and Change* 41 (2): 303–335.

RECENSĂMÂNTUL POPULAŢIEI ŞI AL LOCUINŢELOR (2011) *Clasificarea activităţilor din economia naţională* Bucureşti.

RUSU, R.–MAN, T.–MOLDOVAN, C. (2013): The GIS-Based Road Distance and Time Connectivity Index of the Settlements within the West Region of Romania *Studia Geographia* 58 (1): 141–150.

SANDU, D. (1984): *Fluxurile de migraţie în România* Editura Academiei, Bucureşti.

SOAITA, A., M. (2013): Romanian suburban housing: home improvement through owner-building *Urban Studies* 50 (10): 2084–2101.

WEB REFERENCES

http://www.insse.ro/cms/files/RPL2002INS/vol2/tabele/t31(576-584).pdf
www.recensamantromania.ro/wp-content/uploads/2015/05/vol3_t33.xls

The transport in our time-budget

Tamás Fleischer
Institute for World Economics, Centre for Economic and Regional Studies of the Hungarian Academy of Sciences (IWE CERS HAS), Budapest, Hungary
E-mail: tfleisch@vki.hu

Melinda Tir
Institute of Economics, Centre for Economic and Regional Studies of the Hungarian Academy of Sciences (IE CERS HAS), Budapest, Hungary
E-mail: tir.melinda@krtk.mta.hu

Do we save time with our faster transport modes? The answer is, no. The authors answered this question after comparing the average daily per capita transport time-use based on the 1986/87, the 1999/00, and the 2009/10 Hungarian time-budget survey. The average time-use remained between 60 and 65 minutes, same as it was in 1977. During the period studied, the share of the motor/car time-use approximately doubled in the total transport time-use, while the other modes (walking, cycling, and public transports) decreased proportionally. In the same period, there was a wide distribution in the per capita daily transport time-use data influenced by geographical destination choice (in space and time)—and by demographic (age, gender), spatial (county, settlement status), and social (activity, qualification) variables. The paper analysed the effect of the latter explanatory variables on the heterogeneity of the transport time-use. The gender and activity variables can explain motor/car time-use differences; geography and settlement status the bicycle-, and the settlement status also the public transport time-use differences. However, all the explanatory variables analysed could only explain 10% of the divergences.

Keywords:
transport,
time-use,
time-budget survey,
Hungary,
trends, distributions,
transport modes.

Introduction

Transport performance is traditionally measured in vehicle-km, passenger-km, and tonne-km—as a product of transported volume (vehicles, passengers, or tonnes) and distance. It is well justified by the fact that the cost, fuel consumption, or energy input of the transport are to a certain extent proportional to the combined transport performance index.

Besides the volumes, this study focuses on the time spent on transportation rather than on the distance covered by the transport. This seems to be an *anti-spatial* approach, since we do not take into account distance in our analysis. However, we will see that the time-use orientation also emphasises important spatial connections and consequences.

In a next section of the introductory chapter, we show different myths and suppositions that are related to the time used for transportation. We also present a background to the time-use approach and its application in the field of transportation. A separate section discusses the Hungarian background literature on the subject. Subsequent to this discussion, we present the subject and methodology of our study, the sources of the data, and the selected data series: time-balance data for the Hungarian transport, including the different transport modes, and the social and territorial background data for people using the different transport modes. Our aim is to analyse the variance and mutation of the consumed transport time in time series (of the past decades), on the one hand, and transport time-use by the related demographic, social, and territorial patterns, on the other hand. The following chapter reports the main results, and the summary of the main findings concludes the study.

A key-effort of the modernist transport planning: saving transit time by higher speed

People perform different mandatory and discretionary activities on a daily basis. The mandatory activities are classified further by Gerike et al. (2015). They distinguish three types of activities and trips: (1) subsistence (work, education), (2) non-discretionary or maintenance (shopping, errands, chaperoning, care, voluntary, personal care, other), and (3) discretionary (leisure). Regardless of whether we like or dislike an activity (the activities can be either pleasant or monotonous), we have to invest time in the activities we choose to perform.

Jones and Boujenko (2009) distinguish 'Place' and 'Link' functions in an urban area. As a Link, a street facilitates movement of people, thereby minimising their travel time. A Place is always a destination in its own right, where people are encouraged to spend time performing different activities[1]. Modernist planning always considered the transit spaces as a Link; in addition, people, the users, also agreed that the time spent on transportation is a loss that has to be shortened. While estimating the cost-benefit of a new transport project, the time that can be saved (e.g. by an electrified rail or new overhead crossing) is considered as part of the benefits. Time-

[1] The quoted authors underline that even an urban street can become a Place, and their main aim was to show which planning tools are necessary for facilitating this transformation. This is important, but it is not the main topic of our paper. Here, we used their expressions to describe the time-spending and the time-losing features of spaces distinguished as Place and Link.

saving was definitely considered as a social gains originating from various investments.

The existence of that social gains from time saved during a transit was questioned in the 1970s.

Current status: The transport time budget is invariant and channelize transport time-savings for transportation

Ivan Illich was one of the pioneers who realised that time-savings, which technological developments aimed at achieving in transport, do not exist for the society (Illich 1974). The speed that streamlined transport modes are able to achieve, relative to walking, 90–95% of the transport time should have already been saved.

Where is that time? The answer was shocking: we use the time saved though transport for transportation. In other words, we choose distant destinations and use expensive travel modes that entails paying the extra costs through extra work. We can learn that the destination points of the trips are not fixed in the space but rather our time budget seems to be stable: that is the time amount that a society is willing to spend with transport does not change, or change only very slowly.

This argument is justified for the people as they feel that need to spend the same amount of time on transit as in the past. However, the question is whether it can be proven with data. This issue challenged the authors of this study to focus on the Hungarian transport time budget trends.

Early time budget studies

Although there were studies that dealt with the time-use of different activities, the organised time-use surveys did not begin until the 1960s. Hungary was among the early promoters of this activity, and there is a wide consensus that one of the main starting documents of the uniform and internationally comparable time-use surveys was that of Alexander Szalai (1972). Since its first release, several other international organisations, such as International Association for Time Use Research (IATUR), Multinational Time Use Study (MTUS), Centre for Time Use Research (CTUR)[2], have collected and prepared databases and elaborated methodologies for comparable surveys. Here, it is worthy to mention about the Eurostat based, HETUS, (Harmonized European Time Use Surveys, see Eurostat 2009) project that promoted and developed similar corporations.

Even if the time-use surveys always contained *transport* data, the time-orientation thinking arrived from other disciplines into the transport sector. One starting point for time-use surveys can be traced back to a seminal paper of Hägerstrand (1970), where the author underlined the importance to raise the human factor and the time

[2] See Jara-Díaz – Rosales-Salas (2015) and http://www.timeuse.org/; http://www.timeuse.org/mtus

dimension into the regional approach. Fifteen years later, he published a whole book about time-geography (Hägerstrand 1985). The other early transport time researcher, Illich, belonged to the field of sociology. It is typical that regional and social experts recognised the importance of the time-use surveys in transportation: their approach markedly differ from the traditional transport thinking, where the origin and destination points are considered as fixed and the time appears but as the possibility of acceleration to save time during a trip. Gerike et al (2015) call this approach *trip-based* one and sharply distinguish it from the *activity-based* approach that 'conceptualise travel as an activity within an individual's daily activity schedule which creates the demand for that travel' (ibid. p. 4.). The national travel surveys focus on the movement and related information in terms of trip stages, distances, modes, and circumstances; however, they rarely link the destinations to the trips. Time-use surveys embed the trip into the daily activity in a better manner, thereby making facilitating a wider analysis.

Hungarian time-balance researches

As mentioned above, the Hungarian Statistical Office had started time balance surveys in the decade 1960s. The studies that elaborated the data focused first on the way of living and considered other social analyses. As part of the study, commuting and transport time-use results were also presented, but they were not detailed (we will use these findings in the sections below).

An exception is the detailed work of Lakatos (2013) about the time-use of the employed people during their commuting/travelling to their workplace. While the author offers a general picture about the transport time-use, the main question analysed was the manner in which the long and extra-long (> 90 minute) commuting time influences the life of the people and families, their leisure time, and other time-use. The main groups involved in the comparison were the employed and the other people, and the object of the comparison was their general (non-transport) activity influenced by their work and commuting time. In our study a significant portion of the background data will be similar as our data sources were same as that of Lakatos. We present that the employed people use about the 60% of the total transport time-use. However, as Lakatos already presented the time-use of these employed people is not representative for the whole population.

Our study: What is the difference?

While the aforementioned study by Lakatos focused on employed population, and on the social and behavioural consequences of their fixed work and commuting time, our study turns around that approach. Our study focuses on time spent with transport and with different transport modes and considers different social, demographic and spatial background as the cause, and not as consequences of the varying time-use patterns. Our first aim was to describe how the transport time varies by the following categories (explanatory variables):

- 2 genders (male, female);
- 15 age groups (10 to 84 years old divided into five-years age groups);
- 4 settlement statuses (Budapest, county seats, other towns, villages);
- 20 counties (Budapest and 19 Hungarian counties);
- 4 academic qualifications (finished primary school or less, skilled, graduate, professionals);
- 6 activity statuses (employed, student, pensioner, child-care, unemployed, and other); and at another scale:
- 3 periods of surveys (1986/87, 1999/2000, 2009/10).

Dependent variables were the *total time spent on transportation*, and/or the time spent travelling in *six different transport modes* (walking, cycling, local public transport, regional/long distance public transport, motor/car transport, and other modes).

Theoretically the whole problem can be described in an eight-dimensional space, where the axes are not independent variables: we can easily present the time-use differences by gender, but these differences are also consequences of differing activity, education, etc. patterns of the two genders. However, as a first and easily understandable step, we present here two-dimensional diagrams by single explanatory variables.

Methodology

The data for the study was extracted from the time-balance surveys performed by the Central Statistical Office of Hungary (KSH) in the years 1986/87, 1999/2000, and 2009/10. Although the time-balance surveys by the office began in 1963, the comparability of the results from the earlier surveys is relatively low. Therefore, we have taken data from the last three surveys. The detailed database of the KSH survey, comprising elementary data, is also available in the databank of the Centre for Economic and Regional Studies (CERS), which enabled us to base our analyses and calculations on elementary (people level) data, and not just on the prepared KSH publications (KSH 2012).

Concerning the description of the methodology of the time-balance surveys and data preparations, the Office issued a detailed guide (KSH 2010). Although it was important to learn and apply the entire methodology in our study, here we have reviewed those details that are indispensable to understand this paper.

The record of the previous time-budget surveys is in the form of a diary entry covering the 1440 minutes of a day. In the latest survey, between October 2009 and September 2010, a representative sample of about 10,000 people filled the diary on one random chosen day of the week, and employed people and students were allotted an extra day during the weekend to complete the survey. The 1986/87 and 1999/2000 sample also comprised 10,000 people, but they filled out four diaries: namely one day for each season. The survey was activity-based and more than 500 elementary activities were distinguished. The main activity groups include the *physiologically fixed activities,* such as sleeping or eating; *leisure-time activities,* such as sports or cultural activities; and the *socially fixed activities*, such as working, learning, housework, shopping, and transportation.

Transportations that exceeded 5 minutes appeared in the time-balance as a main activity, also pointing to the destination activity. If the transport time was shorter than 5 minutes, then it was added to the destination activity, i.e. to the reason behind the transportation activity.

Not every kind of movement is considered as transportation activity. Running, swimming, or hill climbing are leisure or sports activities, and not transportation. However, a trip to a bath or a hill (to swim or climb there) *is* transportation.

Waiting can also be considered as a possible transportation activity, (but it might not always be connected to transportation). In the survey, waiting time that was connected to transportation and exceeded 10 minutes was classified as a separate activity. However, shorter waiting periods during the transit evidently became part of the transport time.

Our study distinguishes between six transport modes[3] namely walking[4], cycling, local public transport[5], regional or long distance public transport[6], motor or car transport[7], and other[8] transport modes.

Based on the collected data series, the KSH also offers weight coefficients that adjust the representation of the used sample more accurately to the total population proportions.

As a first step these processes provide the average time spent daily on a given activity by the participants of the activity. These findings are called the 'C' value series. Besides participants of a given activity, others were not active in that case. Their time spent with that specific activity is zero. The 'B' value series show the percentage of the active participants within the given population. A third series, the series 'A' provides the time that an average member of the society spends daily on a given activity, by calculating both the participants and the non-participants in the whole population. We will also refer to it as average *social* time-use. To understand the importance of the three value series and the difference between them in a clear manner, we present the case of bicycle users by settlement status in 2009/10.

[3] While this paper uses statistical data and statistical categories as standard deviation or arithmetic mean, etc., the notion of *mode* as a statistical category is not taken into account. We use the word *mode* always in the meaning of the aforementioned transport modes.

[4] Walking includes handicapped cart use.

[5] Tram, bus, trolley bus, local train, funicular, etc.

[6] Train, coach, airplane.

[7] Motor, car, taxi, shuttle-bus.

[8] Horse-cart, waggon, van, truck, lorry, agricultural tools, and any other mode.

Figure 1

Distribution of 'C'-type values of cyclists ('users') by settlement status

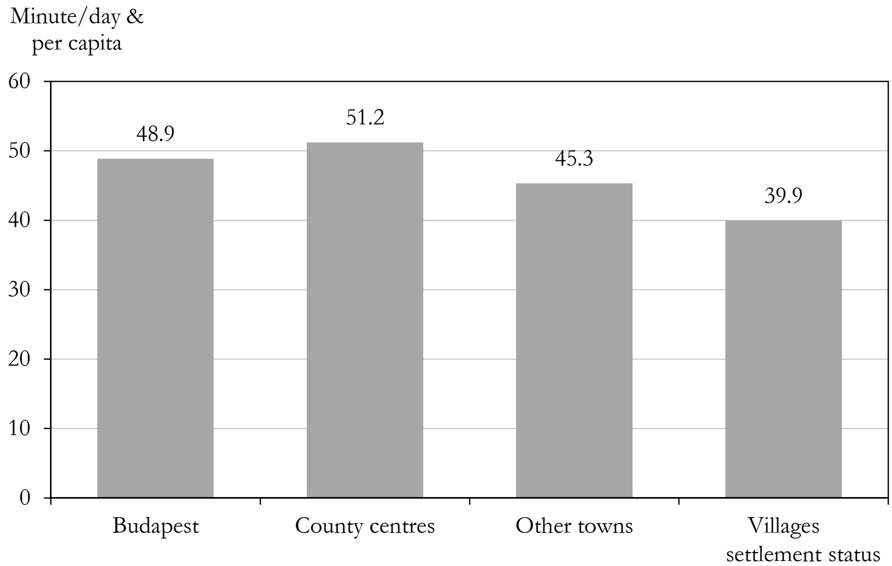


Figure 1 shows that the cyclists use their bicycles for approximately 40–50 minutes per day: 40 minutes in villages, 45 minutes in smaller towns, or about 50 minute in the county centres and the capital.

Figure 2

Percentage of users ('B'-type values of cyclist) by settlement status

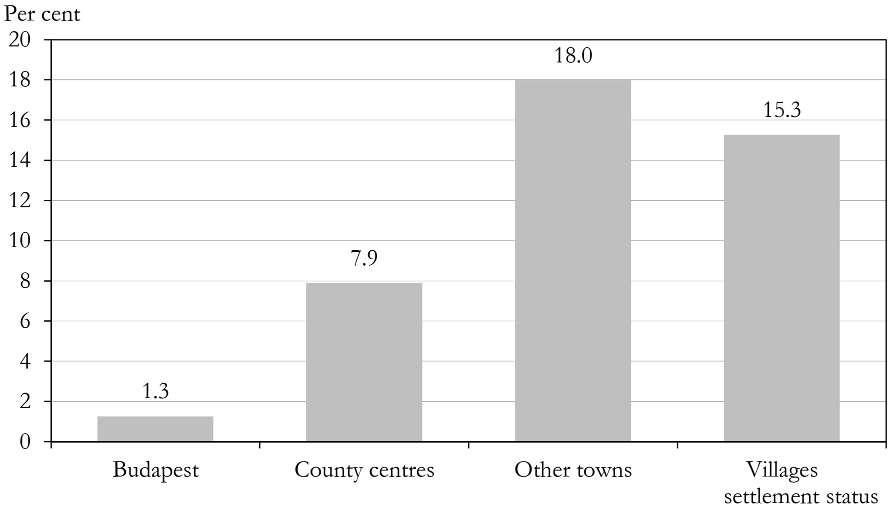
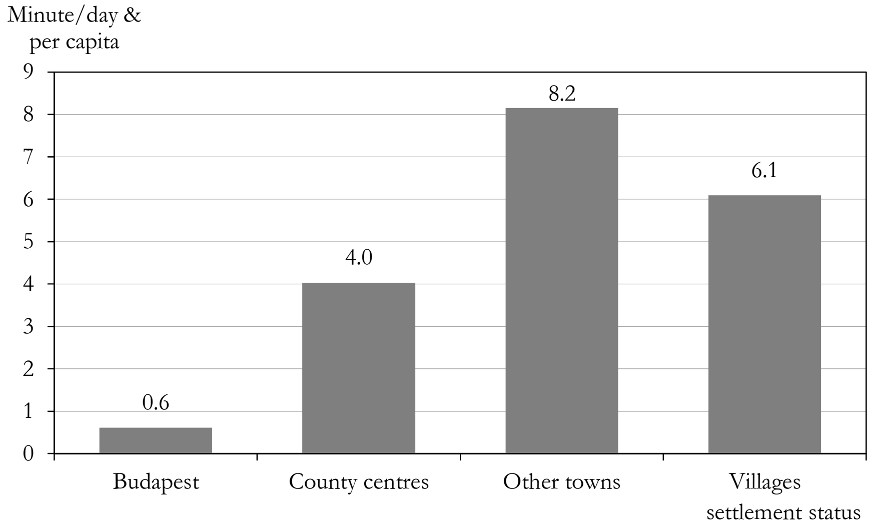


Figure 2 shows that there is much bigger difference in bike use proportions between settlement categories than between the time-use habits of the users. While the users in bigger settlements spent more time on their bikes than those in the smaller settlements, the proportion of cyclist population in the capital is still very small (1.3%), six times bigger in the county centres, 18% in other towns, and more than 15% in villages. These differences are dominant when we observe the average social time spent on bike in Figure 3.

Concerning the time spent on cycling, Figure 3 shows that 'average residents' of Budapest (biker or non-biker), county centres, villages, and small towns spend approximately 36 seconds (0,6 minute), 4 minutes, 6, and 8 minutes, respectively. A superficial observation of these values will fail to make sense, as the 'average resident' does not exist. However, only the 'A'-type social time values can be compared with the time-use of the other transport mode users; the 'A'-type diagram can show us the social proportions between different transport mode's time-use habits or provide a comparison between the social time-use of the transport and other activities.

Figure 3

Distribution of 'A'-type values of cyclist ('average resident') by settlement status

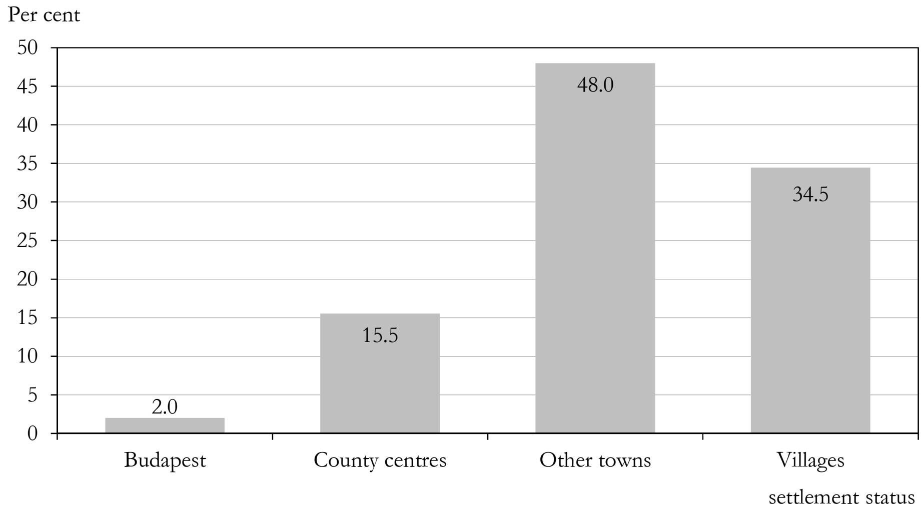


All the above values describe the behaviour of the average transport user, presenting the per capita daily time-use of the average resident ('A') or the average user ('C') of a transport mode. In certain cases, we also wanted to show the volume of the total transport time, multiplying the per capita time-use with the number of the people using the given mode(s) or belonging to the aforementioned categories. These results help us to see the composition of the daily traffic that we can encounter on a given day. Concerning the sample of the study, the population of participants between

10–84 years in Budapest, county centres, other towns, and the villages are 1 536 464, 1 802 868, 2 760 377, and 2 650 376, respectively. Multiplying the per capita cycling time of the people with these values, we get the proportions as shown in Figure 4. The differences between settlement groups have been growing even bigger than the per capita differences that were indicated in Figure 3.

Figure 4

Share of the total time the population spent cycling (100%) by settlement status (2009/10)



Out of the easily traceable diagrams, we developed the multidimensional database. We found that the distribution of the time spent on transportation is not normal, but an asymmetric one and it is the logarithm of the time spent that shows a normal distribution. We ran an ANOVA multivariance analysis for both the original and logarithmic scale to examine the extent to which the different social, demographical, and territorial variables are able to explain the deviations. The two approaches we will present in the last chapter have shown very similar results. This portion of the study would need further analysis and control before making final statements. Concerning the multivariance analysis, had to decrease the range of the variables: we classified the 20 county variables into 5 geographical groups and the 15 age groups into 4 new categories: 15–29, 30–44, 45–64, and 65+ years.

The Hungarian time budget for transport: Main trends during the period 1987–2000–2010

As mentioned above, the first aim of the study was to control the amount of time spent in transport, and the manner in which the time spent has changed over time.

The targeted population in the survey of 1986/87 was between 15–74 years of age, and therefore the comparative time series had to refer to that age group of the population. Since the age groups 10–15 and 75–84 travel less than the average, the average time spent on transit for the smaller population range shows bigger values than for the 10–84 age groups. Especially for 2009/10, the related values are the following:

Table 1

Comparison of the total transport time-budget and user share depending on the range of the age groups

Total transport time 2009/10	Average daily social transport time-use for the whole population (Series 'A') [minute/day & per capita]	Share of transport users in the whole population (Series 'B') [per cent]	Average daily transport time for the transport user population (Series 'C') [minute/day & per capita]
10–84 years	62.3	78,6	79.3
15–74 years	65.2	80,6	80.8
Difference	2.9 min	2,0	1.5 min

The stability of the total transport time budget

Comparing the data for the entire transportation activity (all modes) of the three survey period, the most stable values were identified for the share of the people using transport ('B' index) at a given day: 81.2% for 1986/87; 80,0% for 1999/00; and 80.6% for 2009/10. The standard deviation (SD) is 0.59. To make it comparable to the different dimension other indexes we used rather relative standard deviation (RSD) that is standard deviation projected to the arithmetic mean. The RSD is 0.7%.

The average daily time spent on transportation by those who used some transport mode ('C' index) was 76.2 [min/day & per capita] in 1986/87; 74.9 [min/day & per capita] in 1999/00; and 80.8 [min/day & per capita] in 2009/10. The RSD is 4.4%.

Adding to this user population those, whose transport performance was zero on the given day, the average daily social time spent in transportation by the whole population ('A' index, 'social time-use') was **61.8** [min/day & per capita] in 1986/87; **59.4** [min/day & per capita] in 1999/00; and **65.2** [min/day & per capita] in 2009/10. The RSD is 4.7%.

Although methodology, sample size, age range, etc. are different, the Central Statistical Office conducted time-budget surveys already in 1963 and 1977. In 1963, the surveyed population was between 18 and 60 years old. In 1977, the age limit was 15 and 69 years. Using the comparison of Andorka et al. (1990), for a population between 15–69 years of age, the average time spent on transport by the whole population was **63** [minute/day & per capita] in 1976/77 (74 min for male and 55 min for female). The data for the same survey conducted in 1986/87 was 61 [minute/day & per capita]. Owing to other methodical differences, this result can hardly be compared with the recent findings. However, at first sight, we can say that **the hypothesis of the relative stability of the total social time spent in transport was supported by the Hungarian surveys. The average daily per capita transportation time-budget has remained within narrow limits in the same band, since the last 35 years**.

Distribution function of the daily time-use values

Before we think about the average time spent on transportation as a characteristically symbolic value describing the behaviour of the population, it is worthwhile to see the distribution function of this value. *Figure 5a* shows the distribution curve of the per capita daily transport (all modes) time-use of the people in Hungary in 2009–10. This diagram is built on unprocessed sample data, otherwise relating to series 'A', but the average time-use here is 59.6 minute, while the relative standard deviation is 106%. The figure shows that about 22% of the people do not travel on an average day. The people who travel follow an asymmetric distribution (similar to a scale-free distribution shape): rapidly decreasing from 15–16% to extremely small values while achieving 250–300 minutes time-use. The 59.6–minute average time-use does not represent any distinguished or characteristic point on the diagram.

Figures 5a and 5b.

Distribution of the per capita daily transport time-use of the people in Hungary in 2009-10 ('A'-type values): Normal scale and logarithmic scale

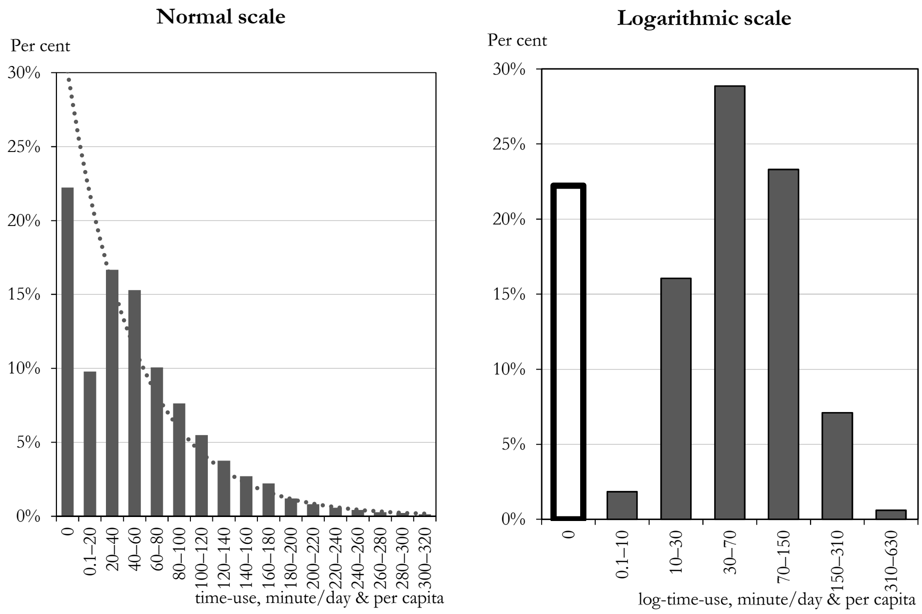


Figure 5b shows the same data, but the time-use axis has a logarithmic scale (the '0' point does not make sense at that scale, not part of the diagram). The active transport time-users present a normal distribution after such transformation.

While the study does not illustrate, but similar diagrams have also been made for the total transport time-use for 1986/87 and 1999/00. We also calculated the median values, as the distance between the arithmetic mean and the median offers a rough estimate on the steepness of the diagram. We saw earlier that the average values were relatively stable in decades, which leads to the question whether the distribution has undergone a change.

Table 2

Comparing arithmetic mean and median values of the transport time-use surveys of the 80s, 90s and 00s

Comparing survey averages and medians Surveys	Average daily social transport time-use for the whole population (Series 'A') 15–74 years [minute/day & per capita]	Median value of the daily transport time for the whole population (Series 'A') 15–74 years [minute/day & per capita]	Difference between average and median (Series 'A') 15–74 years [minute]
1986/87	61.8	50,0	11.8
1999/00	59.4	45,0	14.4
2009/10	65.2	50,0	15.2

Table 2 shows that even if there is a tiny increase in the differences between averages and medians, the main conclusion is that the definite asymmetric distribution of the transport time-use function is a permanent feature; in addition, a very big share of the heterogeneity follows from that character: people prefer short trips that need little time. **The frequency of the daily travel is more-or-less inversely proportional to the duration of the trips**.

While this function would be clearly an important part of the transport time-use analysis, a detailed study should connect the trips to the geographical location of the destinations and would lead out from our time-budget statistics domain. Our study does not analyse this problem further and focuses on the effects that can be discussed within the frame of the time-budget statistics.

In the next section, we have continued with a detailed analysis of the components of the time-budget using the previous, 2009/10, survey data series. Occasionally, we have referred to the time series data in case of specific topics, but otherwise the study focuses on the modal differences in transport time-use and on the effects of the social, spatial, and demographic differences on the time-use.

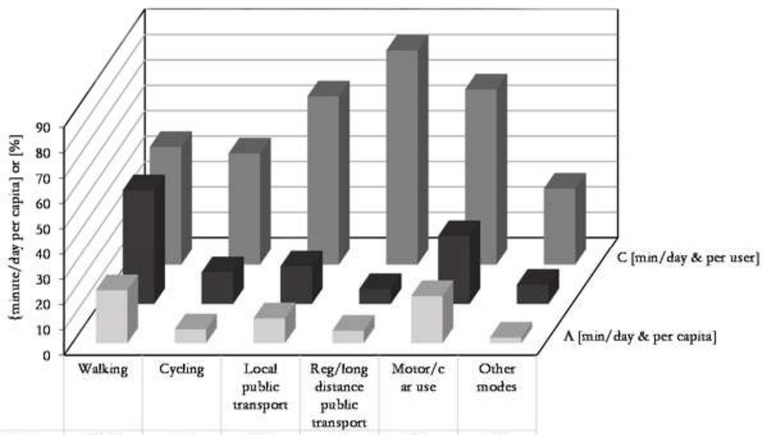
The Hungarian time budget for transport: Detailed study of 2009/10 focussing on the main causes of the deviations

Even if the average total transport time-budget does not change much in time, we know that this average covers very different patterns of time-use.

Composition of the transport time-use from different transport modes

Figure 6/Table 3

Transport time-use composition of the different transport modes



	Walking	Cycling	Local public transport	Reg/long distance public transport	Motor/car use	Other modes
A [min/day & per capita]	21.0	5.5	9.9	5.0	18.6	2.3
B [user % of the population]	45.0	12.5	14.8	5.9	26.9	7.6
C [min/day & per user]	46.6	44.0	66.4	84.6	69.2	30.1

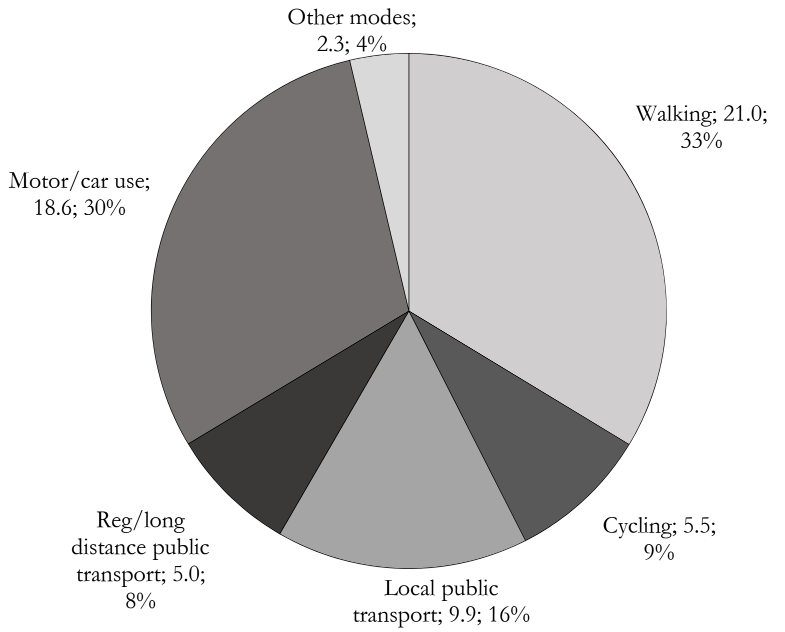
As seen above, 78.6% of the population uses transport on an average day. Row 'B' in Table 3 shows in detail that 45% of the population walks on an average day, 26.9% uses car or motorcycle, 14.8 + 5.9% uses public transport, 12.5% uses cycle, and 7.6% uses other transport mode. We can also see that the shares of the modes-use are overlapping, and they cannot be added.

While only 5.9% of the population uses regional or long distance public transport, those who use the mode spend maximum amount of travelling time, i.e. an average 84.6 minutes a day per head (see row 'C'). Car users and local passengers spend 69.2 minutes and 66.4 minutes a day, respectively, travelling in their modes. Concerning people that choose to walk, on an average people walk 46.6 minutes a day, while bikers spend 44 minutes cycling. These time-use data overlap more than four times and they cannot be added.

Row 'A' shows the amount of time that a person in the society (that is users and non-users together) spends with each transportation mode. The highest time is spent on walking, which totals to 21 minutes per day and per capita, closely followed by car use (18.6 minutes). Public transport, cycling, and other modes total to 9.9 + 5,0 minutes, 5.5 minutes, and 2.3 minutes respectively. On the one hand, these values enable us to compare time-use in other areas of the society spanning 1440 minutes in a day. On the other hand, an addition of these values gives us the total transport time that average people spend on transport (Figure 7 shows that it was 62.3 minutes in 2009/10).

Figure 7

Dividing the average per capita daily transport time-use (62.3 minutes) between different transport modes in 2009/10

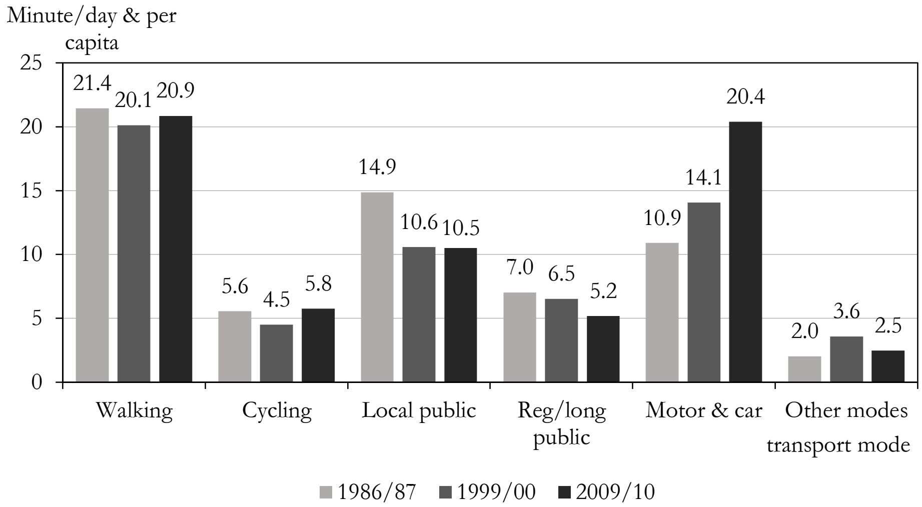


Here, it is useful to look back and examine how these modal shares have been changed over time. As the data here refers to the 15–74 years age range, the numbers differ slightly from the above data in the case of the period 2009/10.

The most characteristic change seen in Figure 8 is an increase in the car-use time: the time almost doubled from a daily 10.9 minutes in 1986/87 to 20.4 minutes in 2009/10. During the same period, both local public time-use and regional/long distance public time-use decreased from 14.9 minutes to 10.5 minutes and 7,0 minutes to 5.2 minutes, respectively. There was no significant change in the rest of the transport modes: walking and cycling both showed the same time-use as earlier (about 21 minutes and about 5.5 minutes, respectively).

Figure 8

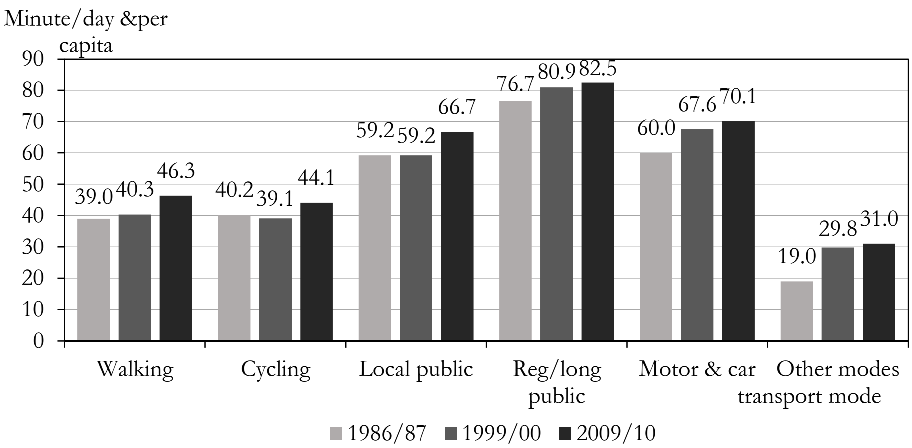
Manner in which the social time spent on different transportation modes has changed in the past two decades



It is interesting to note how these changes were based on the behaviour of the users and on the changes in the share of the users of the different modes. Figure 9 helps us to understand that the users of each mode spend more time on their respective modes now as compared to the time spent earlier. Even in the case of public transport, the time spent by the users of local public transport increased to about 10% and only little less for the distance public transport users.

Figure 9

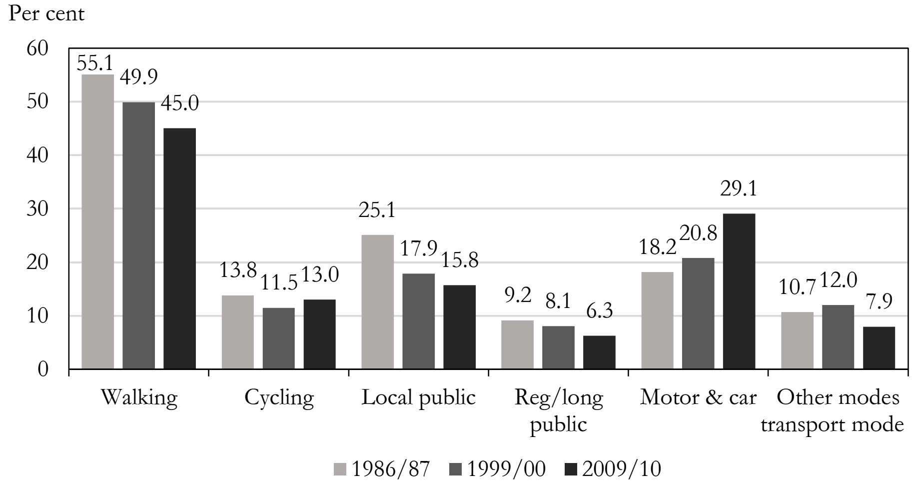
Manner in which the time spent by each transport mode user has changed in the past two decades



During the same two decade period, the share of those using the given transport modes (Figure 10) increased only in the case car (from 18.2% to 29.1%), while at the share for all the other modes decreased (e.g. from 25.1% to 15.9% for local public transport and 55.1% to 45,0% for walking). As a result, the sometimes reversed tendencies produced relative moderate change in the modal time-use of these modes as presented above in Figure 8.

Figure 10

Manner in which the share of each transport mode users has changed in the past two decades



Time-use bias by demographic, spatial, and social background

Time-use bias by gender in 2009/10

Turning back to the 2009/10 survey, the gender share of 62.3 minutes of average daily transport time-use ('A') comes from an average of 67.7 minutes time-use by males and 57.3 minutes time-use by females. This produces an 11.8% RSD value. The significant share of the 10 minutes difference comes from car use, where the daily car-use by males totals to 25.1 minutes, while the daily use by females totals to 12.8 minutes. There is also a similar directional difference in distance public transport and cycling (5.3 minutes versus 4.7 minutes and 6.2 minutes versus 4.9 minutes, respectively). Reverse directional difference is observed in case of walking and local public transport (19.3 minutes for male versus 22.4 for females and 9.5 minutes for males versus 10.2 for females, respectively).

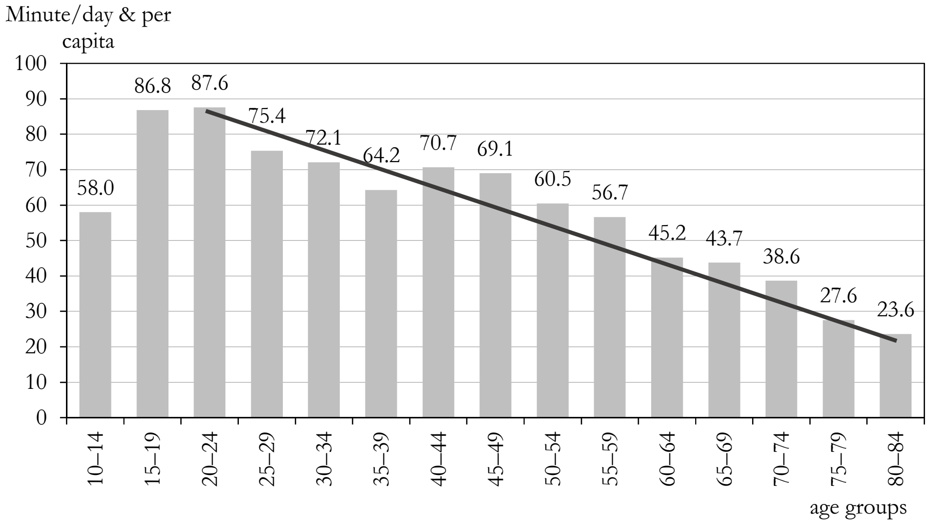
The share of users by gender (percentage, 'B') and the average time spent by the users themselves (min/day & per capita 'C') presented parallel directional asymmetry for the modes: they together contributed the described 'A' results.

Time-use bias by age group in 2009/10

For **transport activity as a whole** and for the total population ('A' type data), Figure 11 presents the average daily social time-use values by five-years long age groups. The youngest (10-14 y) group spends 58 minutes on transportation and the value quickly increases to close to 90 minutes. From the 20-24 years age, there has been an almost monotonous decrease in the average time-use. The black line below depicts the change in the sampled 60 years, between ages 20 and 80, which shows a decline in daily transport time-use from about 90 minutes to about 30 minutes, i.e. a one-minute social time decrease per age-years.

Figure 11

Social time spent on transport by age group in 2009/10



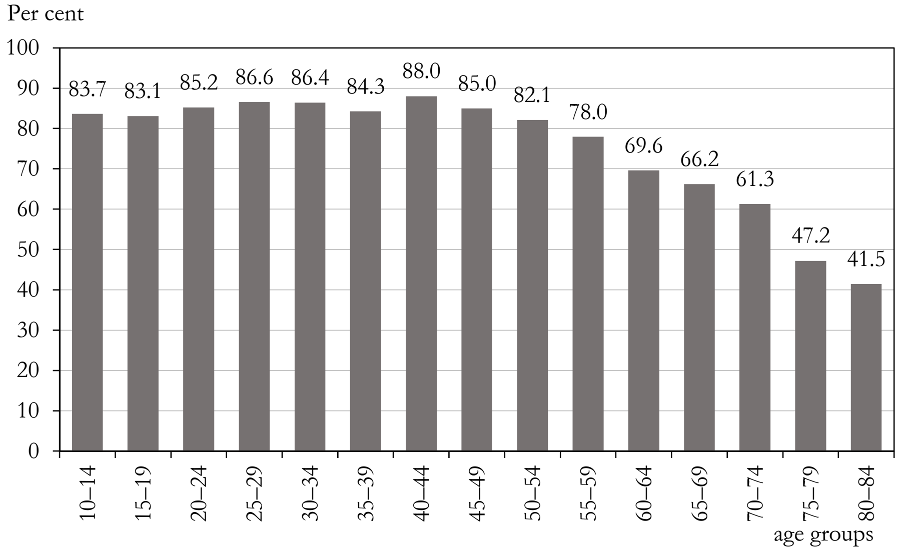
A more detailed observation reveals that nearly 90 minutes of time-use in young users aged 15–24 years, is followed by a plateau-shaped curve with a time-use of 70 minutes in users aged 30–49 years, following which started the smooth decrease in time-use.

This age group based time-use shows a high RSD of 31.5%.

It is also worthwhile to study the 'B' and 'C' components of this change by age group. Concerning the share of users within the population ('B'), it stays almost constant around 85% until age 50 and from this point it decreases to 45% along the next 30 years (Figure 12). The RSD of this 'B' component is 19% and an additional 18% RSD is explained by the 'C' component. The actual users daily spend 103-105 minutes on transport at age 15–24 years, which reduces to 87 minutes from 25 years of age and decreases further to 57 minutes in the next 60 age-years.

Figure 12

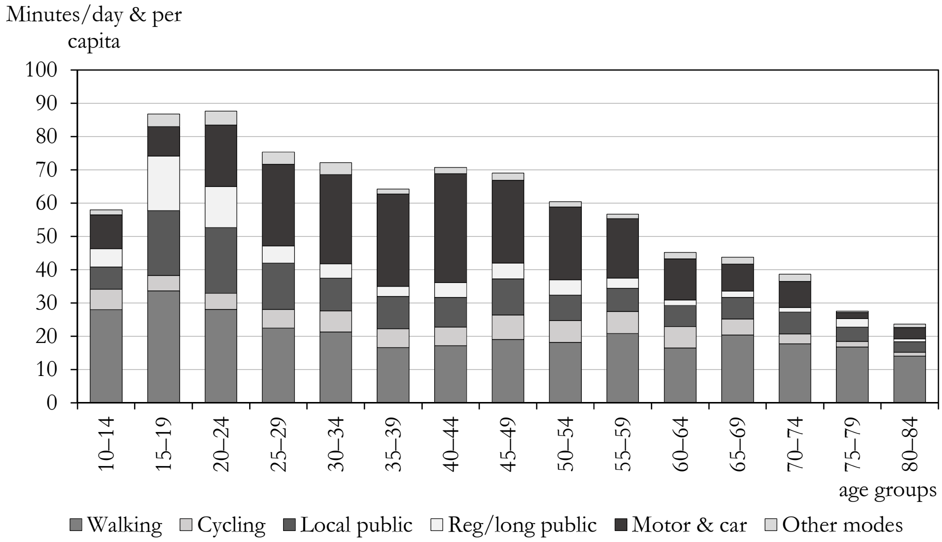
Share of the transport user population by age group in 2009/10



As for the modal share of the age-based time-use, the results can be easily presented (for the 'A' type data) through a single figure (Figure 13).

Figure 13

Social time spent on transport by age group and by transport mode in 2009/10



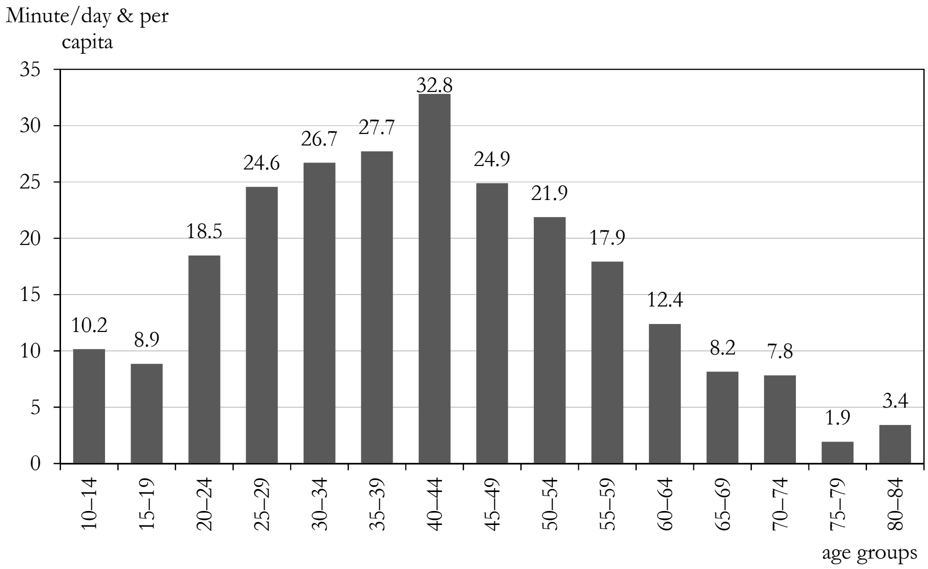
While the contour of Figure 13 is identical to Figure 11, it again shows the extent to which walking and car use dominates the transport time-use. As already shown in

Figure 7, the third notable time user is the local public transport. However, here we can also follow that car use time increases with increasing age and pushes down all the other transport modes until the early forties; while car use time begins to decrease from the forties, the proportion—sometimes even the value—of all the other modes increases.

It is worthy to highlight the biggest time-users and present them separately. Figure 14 shows the average social time in car and motor use ('A').

Figure 14

Social time spent on motor/car transport by age group in 2009/10



Until the age 19, the average daily car use is 10 minutes. This time-use increases from 20 to 33 minutes in the age group of 20–44 years; however, this time goes back below 10 minutes from 45 years onwards until the succeeding 30 age-years. The RSD is 52%.

The deviation in the time-use among the car users ('C') is much less. The users below 19 years, 20–59 years, 60–74 years, and above 75 years spend daily 55 minutes, about 70 minutes, 60 minutes, and lesser, respectively, in car. The RSD of this component is 16.8%. The main cause of deviation of time-use along the age years is the share of the users. The shape of this 'B' diagram is quite similar to that of Figure 14. The share of the car/motor users below age 19 is 17%. This percentage rises from 25% to 42% in users aged 20–44 years, while this percentage goes back to 13% and then reduces to 5% in the successive 30 age-years. The RSD is 44.3%.

Figure 15 points to the fact that the characteristic time-use rise in the middle age-years originated more from the male population than from the females. Males spend about two times more time on motor/car transport in their active years than females.

Figure 15

Social time spent on motor or car transport by age group and by gender in 2009/10

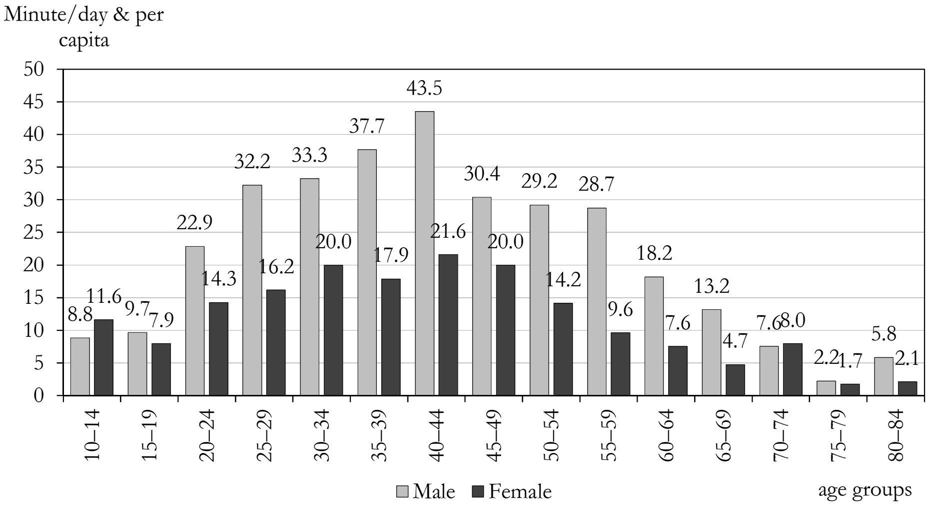
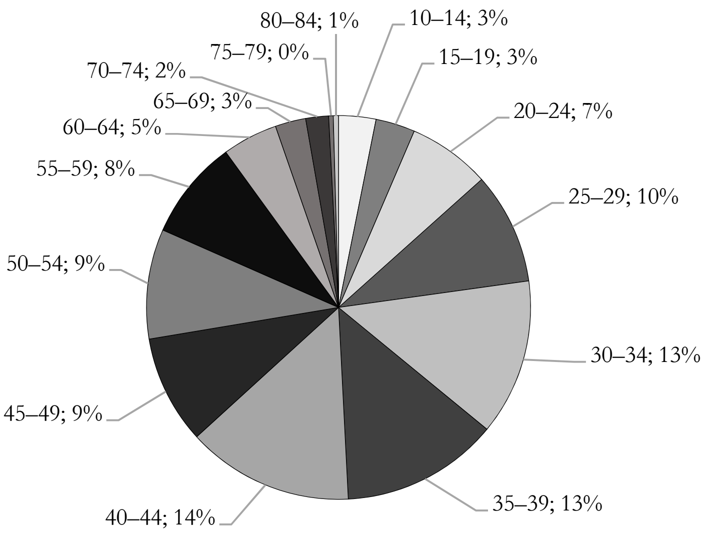


Figure 16

Share of the total social time spent on motor or car transport by age group in 2009/10

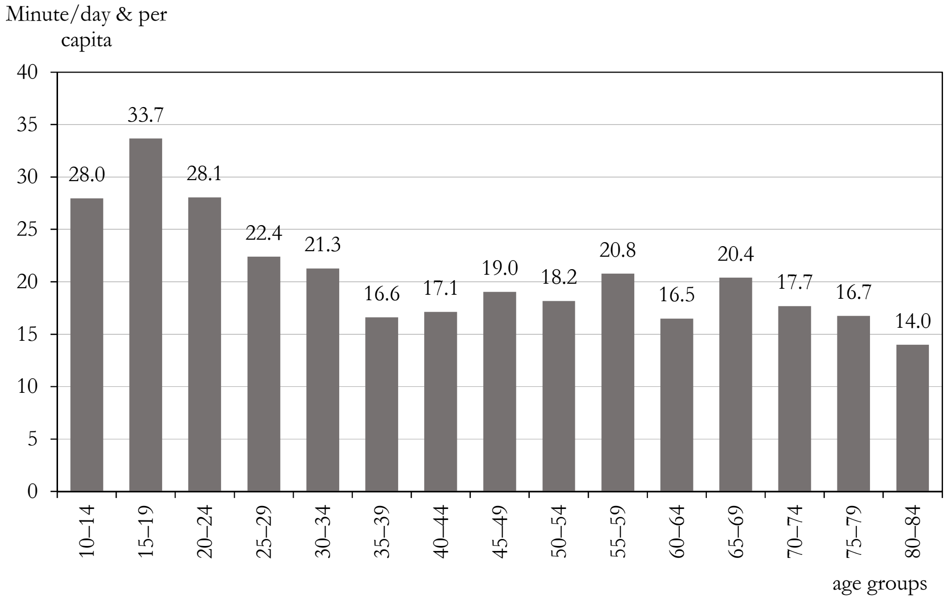


In concluding this section, it is meaningful to look at Figure 16, wherein a pie diagram presents the share of the age groups appearing in cars or motorcycles on an average day. To get these proportions, the per capita time data used in Figure 14 had to be weighed by the population share of the given age group. The Hungarian census was held in 2011. During this period, the people involved in the time-budget survey had become a year older, and hence the census age groups fitted to the time-use data were 11–15, etc. instead of 10–14, etc. Figure 16 shows that **motor/car-users aged 25–44 years** (who represent 33.2% of the 10–84 years old population) **give half (49.8%) of the time that the population spent with this mode on an average day**.

In case of **walking,** an average of 40–50 minutes a day is spent by people who walk ('C'), almost independent of their age (RSD = 7.2%). Concerning the share of people who walk ('B'), about 60% of the people in the youngest age group spend time on walking. This value goes down to 40% until 35 years of age and stays at that level until the age of 50. At this point, another peak is observed as the value rises to 45%, declines back to 40% by the age of 70, and to 30% over 80 years. This 'B' deviation is dominant at a RSD value of 19%. The shape of the diagram with the social time spent on walking ('A'; *Figure 17*) is similar to that described for the 'B' component (RSD = 25.6%).

Figure 17

Social time spent on walking by age group in 2009/10

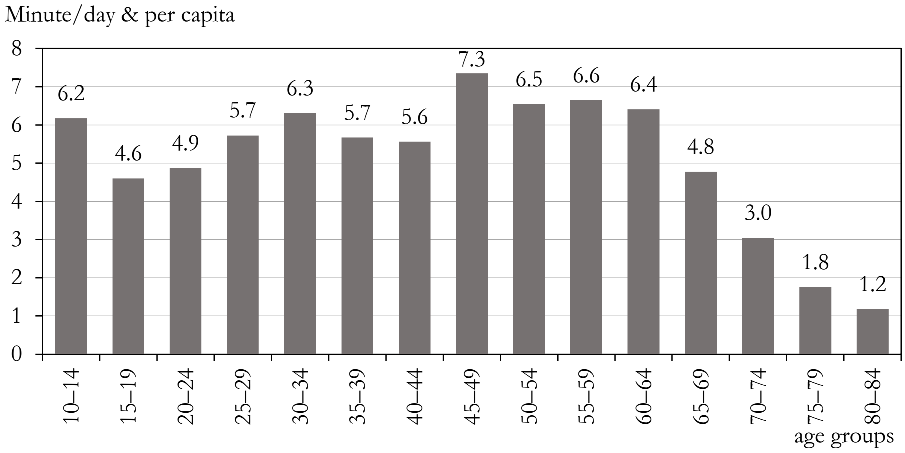


People who use **bike** ('C'), cycle for an average of 40–50 minutes a day, almost independent of their age (this value is quite similar to the people who walk). The RSD is 6.5%. The 'B' value in the diagram shows that, apart from the youngest age group, where 14% use bike, the value starts at 10% and increases to 17%. arriving to the age group 45–49. At this point, the value declines slowly (14%, age 64) and then continues to decline rapidly (~2%, age 80–84). The RSD is 33.9%.

Putting together the picture ('A'), Figure 18 shows that, contrary to the case of walking, the groups that spend most time with this mode are not the younger groups, but the groups between age 45 and 64. The RSD is 33%.

Figure 18

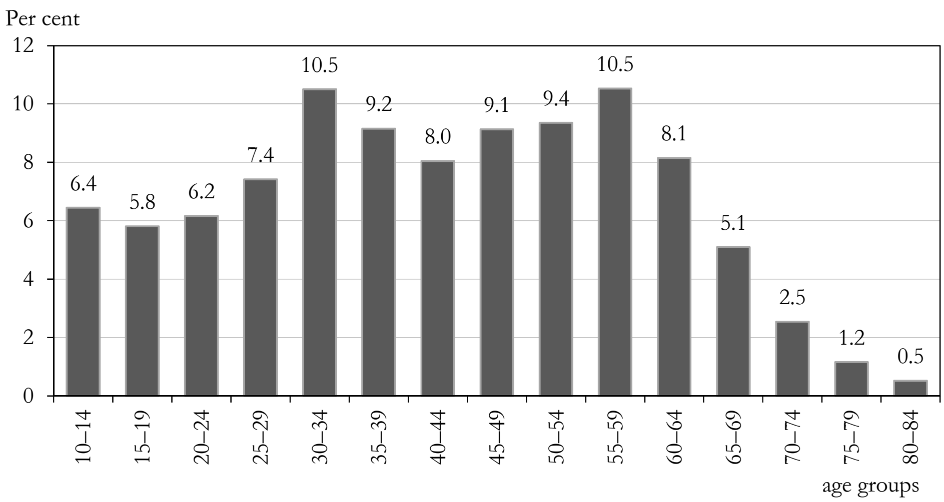
Social time spent on cycling by age group in 2009/10



To obtain the amount of time that an age group spends on cycling, we also have to calculate the headcount of the given cohort, beside the per capita time-use. Figure 19 shows the real share of the time between age groups in the total daily biker time-use (per capita time-use multiplied by the headcount of the age group). The total volume of the users increases the relative significance of the 30s age cohort'.

Figure 19

Share of social time [percentage] spent cycling by age group in 2009/10 (the total cycling time used by all members of all age groups is 100%)

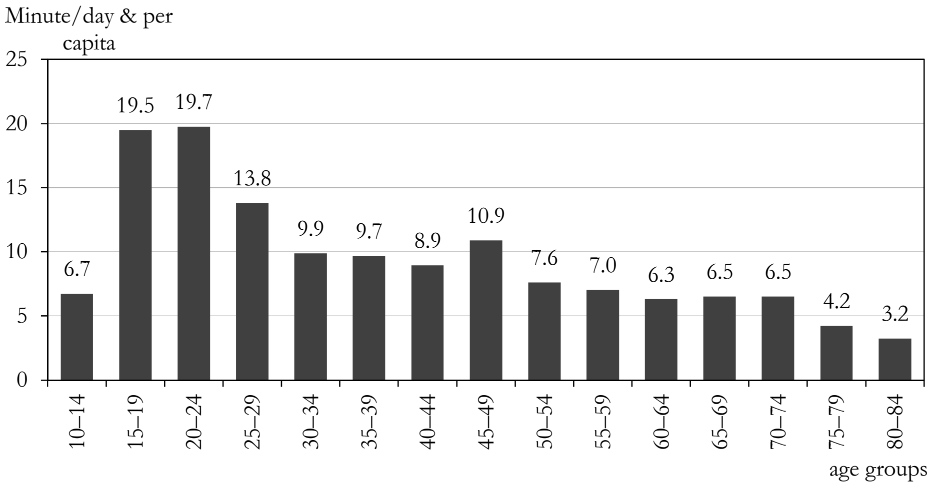


The users of the mode **local public transport** ('C') travel for an average of 55–75 minutes a day. The RSD is 9.6%. Concerning the share of the users ('B'), it is high in ages 15–29, decreasing from 30% to 20%. At this point takes the form of a plateau between the ages 30–64 at about 14%; followed by another plateau at 10% until the age of 74. The RSD is 46.5%.

The share of users with its high RSD value is the dominating factor and the contour of the 'A' type diagram (Figure 20.) follows that of 'B'. Apart from age 10–14, the main users of the local public transport are the young people (age 15–24) with almost 20 minutes of daily use. The successive age groups witness a significant and monotonous decline from 14 minutes to 3 minutes. The RSD is 49.6.

Figure 20

Social time spent on local public transport by age groups in 2009/10



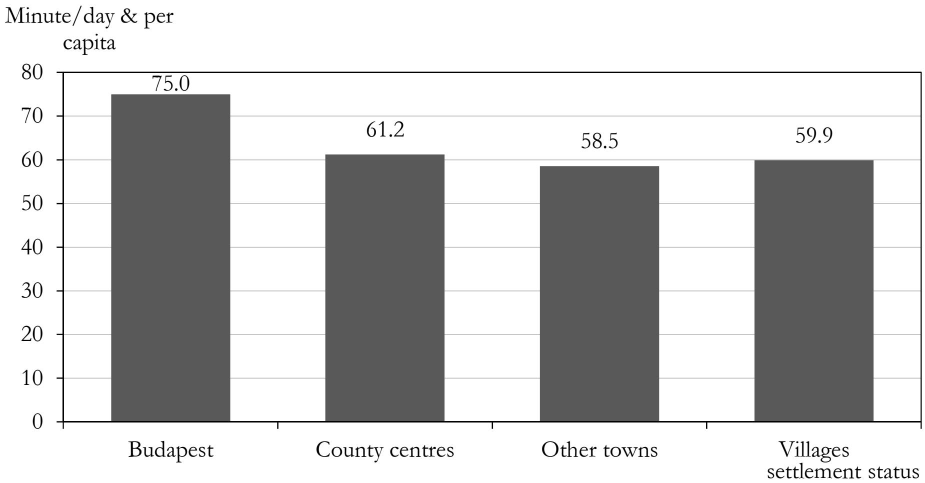
The main characteristics of the **long distance public transport** are similar to the local public transport, but the time-use values are smaller (an average of 5 minutes versus 9.9 minutes in case of local public transport) and the RSD values are bigger (84% versus 49.6% in the 'A' series).

Time-use bias by settlement status in 2009/10

We distinguished four categories by settlement status: Budapest, county centres, other towns, and villages. Considering **all the transport modes,** the social time-use (for the whole population, 'A') in Figure 21 shows that Budapest markedly differs with its average 75 minutes daily time-use, while all the other places use about 60 minutes (RSD = 12.3%). This is caused by the high travel time-use of 95 minutes by people in Budapest as opposed to the 75-77 minutes of time-use in other places ('C') (RSD = 12%). The share of the users ('B') is almost the same (77–80%) in each site (RSD = 1.7%).

Figure 21

Social time spent on transport by settlement status in 2009/10



As the above spatial units cover different volumes of population aged 10–84 years, the per capita time-use expresses the behaviour of the residents, but not the share of the total time the society spends on transport. The Table 4 shows that the share of the population in the groups varies from 17.6% (Budapest) to 31.5% (other towns), and this difference dominates the total time-use shares. The villages and smaller towns each use 29–30% of the transportation time, while the county seats and Budapest use 20% and 21% of the total transport time, respectively.

Table 4

Per capita and total social time spent on transport in different settlement groups in 2009/10

'A'	Time-use/cap [minute/day & per capita]	Population [head]	Population [per cent]	Total time-use [minute]	Total time-use [per cent]
Budapest	75,0	1 536 464	17.6	115250930	21.1
County centres	61.2	1 802 868	20.6	110302223	20.2
Other towns	58.5	2 760 377	31.5	161556649	29.6
Villages	59.9	2 650 376	30.3	158754230	29.1
Total		*8 750 085*	*100*	*545 864 032*	*100*

Concerning **walking**, no difference was observed between the places: people walk for 20–22 minutes ('A') per capita per day. In this case, Budapest shows an average time-use of 21 minutes.

The difference is also small in case of **motor/car transport** use: 20 minutes ('A') in Budapest and in county centres and 18 minutes in the other two settlement groups

(RSD = 6.8%). A speciality of the settlement status case is that the variables, share of users ('B') and time spent by the users ('C'), show an inverse-shaped deviation for almost each mode.

Figures 22 b and c

Percentage of population using motor or car transport [%, 'B'] and time spent on motor/car transport by users ['C'] by settlement status in 2009/10

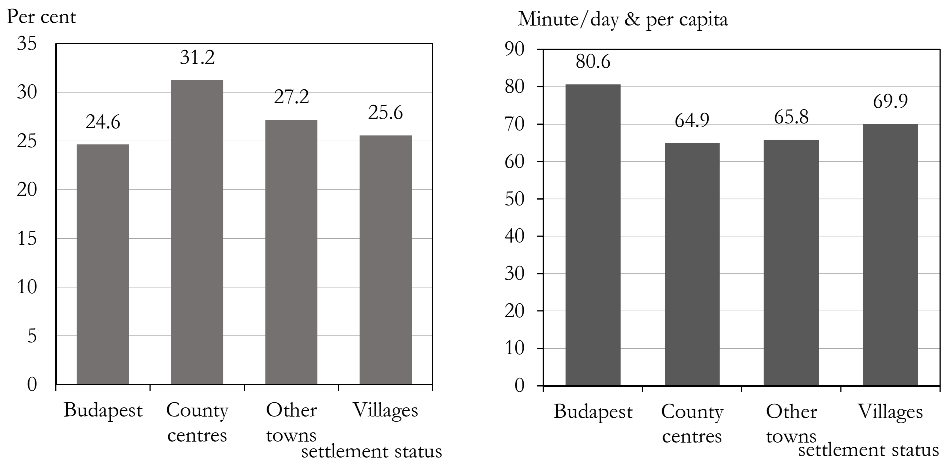


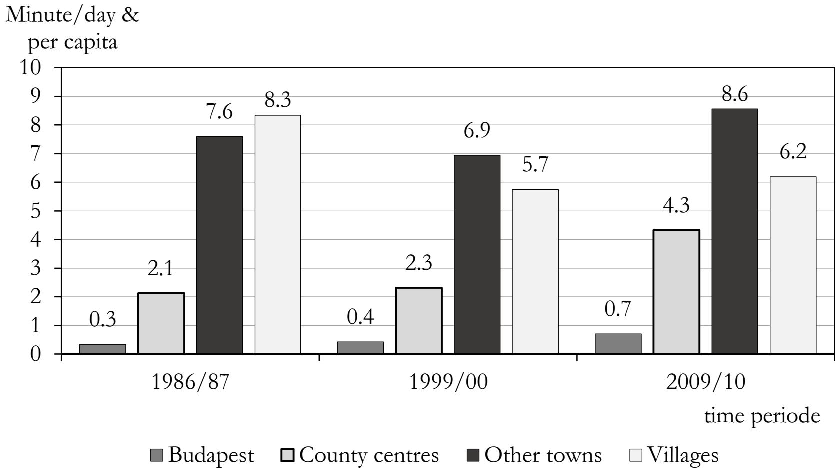
Figure 22 shows that while the share of the motor/car users ('B') in the county centres is significant (33.2%) followed by other towns, villages, and Budapest (24.6%), the average daily time-use by motor/car users ('C') shows a totally reverse ranking. Especially, even the RSD of 'B' and 'C' cases of the motor or car transport mode was almost the same (10.8% versus 10.4%), while at cycle and public modes the RSD show that the share of users ('B') became dominant.

In the case of **cycling,** the time used by bikers ('C') varies between 40 and 51 minutes, which is more in the bigger cities including Budapest – as it was seen in Figure 1. The share of users ('B') vary between 1.3% (Budapest) and 18% (other towns), as we already presented in Figure 2; in addition, the social time used by the whole population ('A') also follows the contour of Figure 2 (see Figure 3).

Considering the small values seen in cycling for Budapest and county centres, it is worthwhile to examine the earlier periods. Figure 23 presents three survey periods together. (The different (15–74) age ranges used for comparison led to higher values for 2009/10, relative to Figure 3.) We can see that the social time spent on bicycle has more than doubled in the past two decades in Budapest and the county centres. The share of the people who cycle tripled during the same period in Budapest: from 0.5% to 1.5% (for users aged 15–74 years). One fact is that there has been a revolutionary change in cycling over the past few years in Budapest; however, both the share of cycle users and the social time an average resident spends on cycling in Budapest is one tenth or less than tenth of those in villages or in smaller towns.

Figure 23

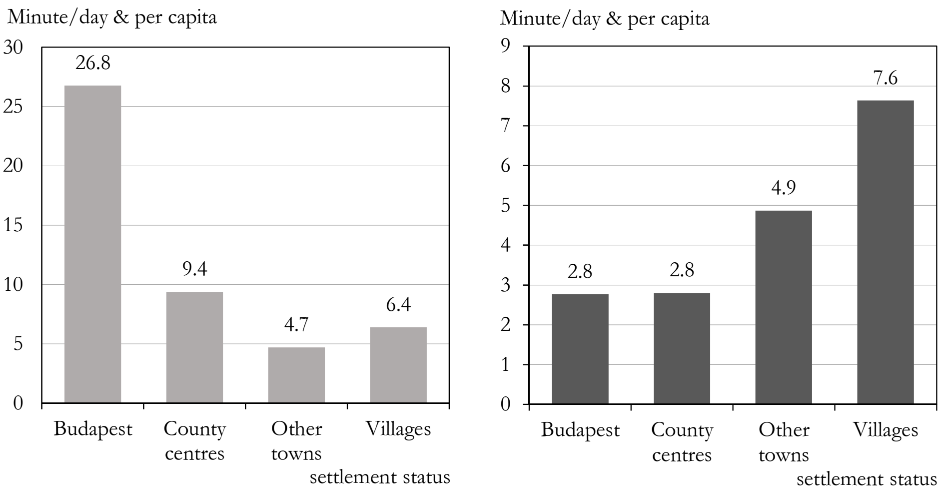
Social time spent on cycling by settlement status in 1986/87, 1999/00, and 2009/10 (people aged 15–74 years)



In the case of **public transport,** the situation of local and distance transport is quite different. Figure 24a presents 'A'-type local transport time-use: daily 27 minutes in Budapest versus 5–9 minutes in other places (RSD = 102%). Naturally, this is significantly determined by the availability of local public transport slowly followed by the necessity to use the mode.

Figures 24 a and b

Social time spent on local (a) and regional or long distance (b) public transport by settlement status in 2009/10



The Figure 24b presents 'A'-type regional or long distance public transport time-use. The range is between 2.8 minutes and 7.6 minutes (RSD = 46%). Evidently, the necessity to use the transportation mode affects the differences between the places largely.

Time-use bias by counties in 2009/10

While settlement status aggregates settlements in a virtual space, studying county level differences connects us with the geographic space.

Figure 25 shows the differences between social time-use of people by counties (**all transport modes** together). The range of the average per capita daily transport time-use spreads from 50 minutes (Békés county) to 75 minutes (Budapest) (RSD = 9.9). Figure 25 ranked the counties by their growing per capita time-use, and also presents a diversified range of modal components of time-use.

Figure 25

Social time spent on transport by counties in 2009/10

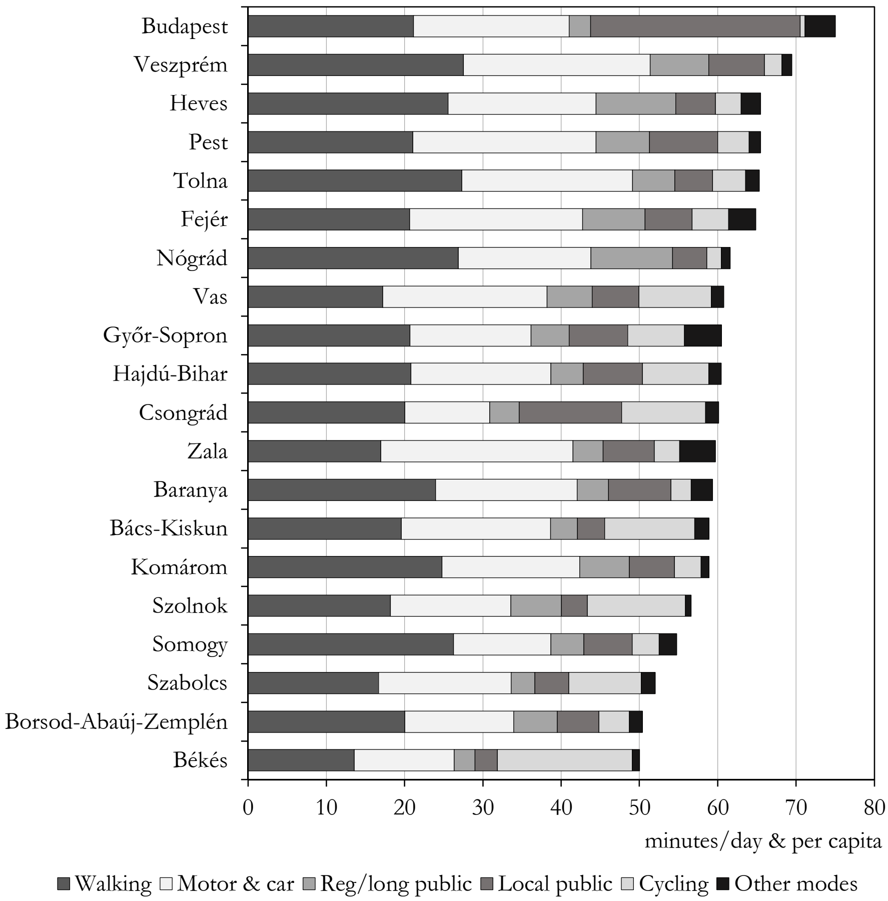


Figure 26

Per capita average daily social time spent on transportation in minutes Hungarian counties in four quadrilles in 2009/10

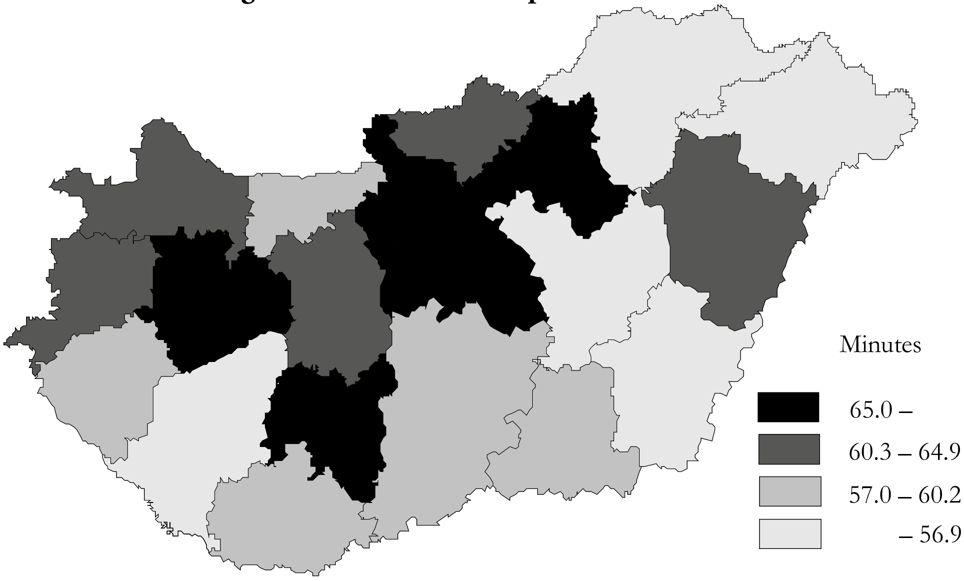


Figure 27

Time spent on different transportation modes by counties in 2009/10

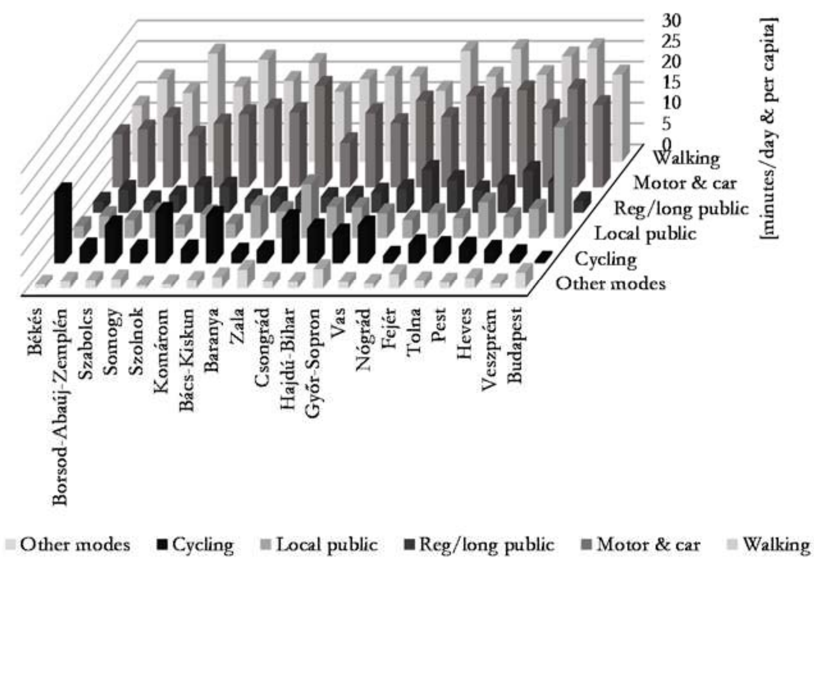


Figure 26. interprets the same social transport time data on a map. The 20 spatial units were classified into four quadrilles by the rank of their average per capita daily transport time-use. The quadrilles do not seem to form any salient spatial structure.

Figure 27 attracts attention to the fact that the main components of the time spent on transportation are walking and motor/car transport in each county, while other modes of transport like local public transport in Budapest or cycling in Békés County peaked at certain locations. Relatively high local public transport and cycling together seem to have led to a significant decline in motor or car transport in Csongrád County.

The overall picture indicates that **bicycle** transport, with its amplitude variance, is also worthy of a countywide study (RSD = 79.5%).

Figure 28

Social time spent cycling by counties in 2009/10. Ranked by increasing average daily per capita time-use values

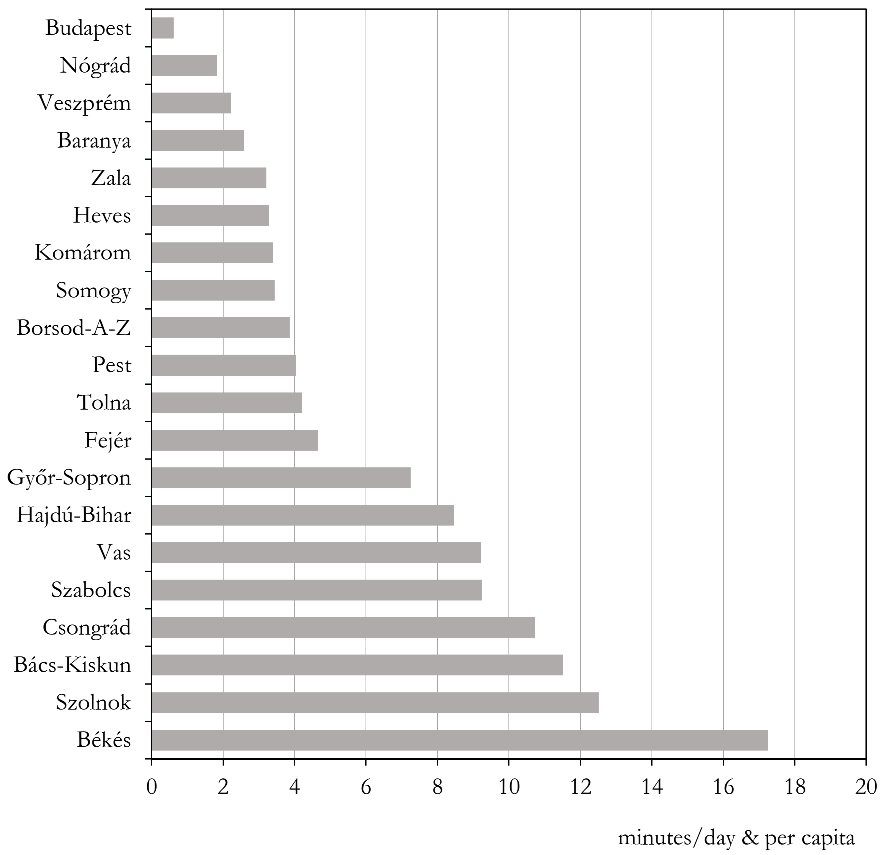
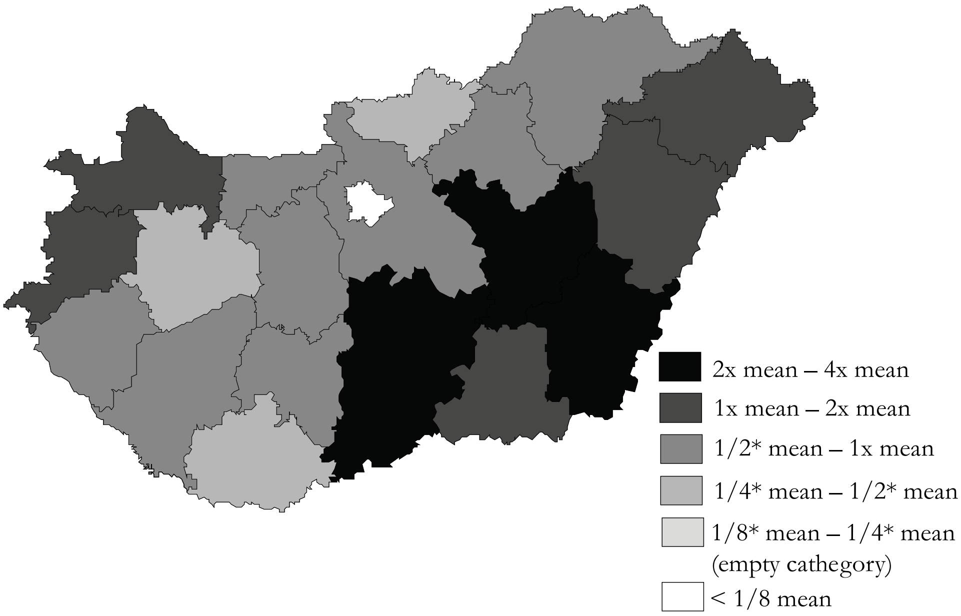


Figure 28 presents the social time spent on cycling in different counties, ranked by the increasing per capita time-use values. While the average cycle time-use in the country is 5.5 minutes/day, it is interesting that just around that average there is a gap in the values and the counties present characteristically lower or higher per capita cycle use times.

Figure 29

Per capita average daily social time spent cycling by counties in 2009/10



The differences can be well studied on the map of Figure 29. As it was mentioned above, average cycling time-use for the whole country is 5.5 minutes per day per capita. There are three counties in the southern part of the Grand Plain/Alföld where people use bike between twice to four times more than the average; and five counties (three in the east and two at northwest Kisalföld) use cycle between 1–2 times the average time. Eight counties along the hilly north-east/south-west line use bike between half and one time the average, while three others, Nógrad, Veszprém, and Baranya, use it between one-fourth and half of the average time. Bike use between one-eighth and one-fourth of the average time is empty and in Budapest, the per capita daily cycling time-use is less than one-eighth of the country average. This geographical pattern follows the varied topography of the country—at least at the countryside.

Time-use bias by activities in 2009/10

This section first presents the **transport time-use** differences by activity status of people that is divided into six categories. Figure 30 shows that students and employed people spend a significant amount of time ('A') on transport, namely 80 and 73 minutes per capita, while the rest of the categories lead to less transport time-use than the 62.3 minutes country average. . (RSD = 24.3).

Figure 30

Per capita social time spent on transport by activity groups in 2009/10

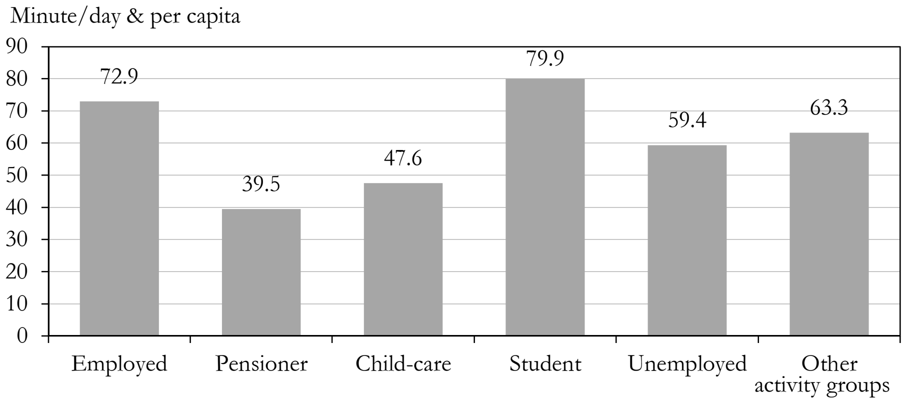
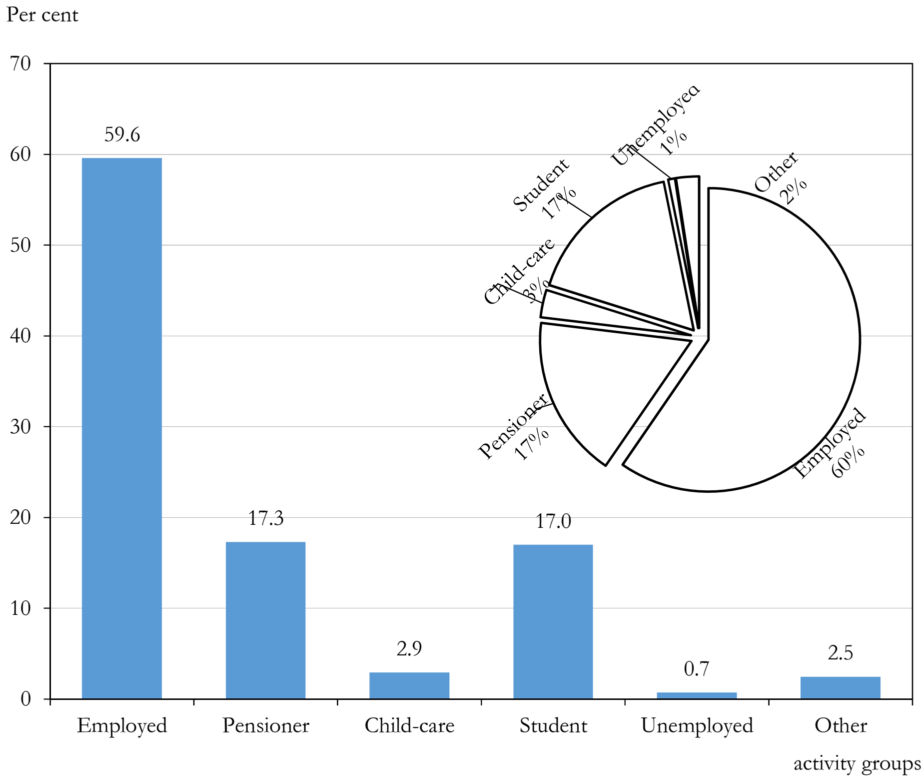


Figure 31

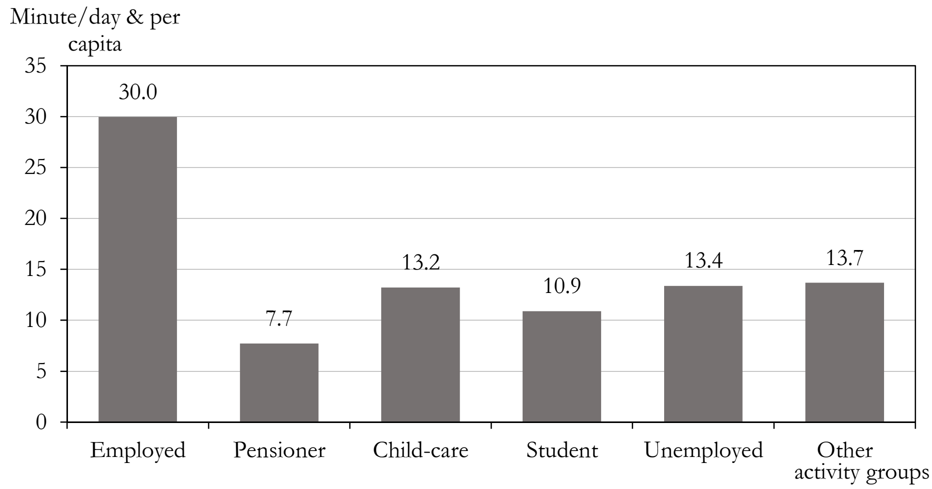
Share of the total time spent on transport by activity groups in 2009/10 (100% is the total time that the population spends on transport)



Since there are big differences in the number of people belonging to different activity groups, the per capita time-use fails to represent their real share that appears as daily traffic. The real traffic share can be presented by multiplying per capita values with the volume of the given activity group. As 51.7% of the 8 750 000 people in the 10–84 age group are employed, 27.7% pensioners and 13.4% students, Figure 31 shows quite a different share of the total transport time-use between the activity groups. The employed people use 60% of the total transport time.

Figure 32

Per capita social time spent on motor or car transport by activity groups in 2009/10



In the case of **motor or car transport** mode, Figure 32 shows that the weight of the employed group is even bigger; the employed people at an average use motor/car transport two-and-a-half to three times longer than people in any other activity group, in the case of the per capita social time-use. Multiplying the per capita use with the population share, it can be inferred that **employed people use 77.3% of the total time used for motor/car transport**.

In the case of **walking,** the per capita social time used by employed people ('A' series) is 17.4 minutes, while other activity groups spend more time walking, especially students who spend 31.6 minutes. Multiplication of per capita time used for walking by different activity groups with the respective population data reveals that employed people, pensioners, and students use 44.4%, 26.4%, and 21% of the total walking time respectively (Table 5).

Table 5

Per capita and total social time spent on walking by the activity groups in 2009/10

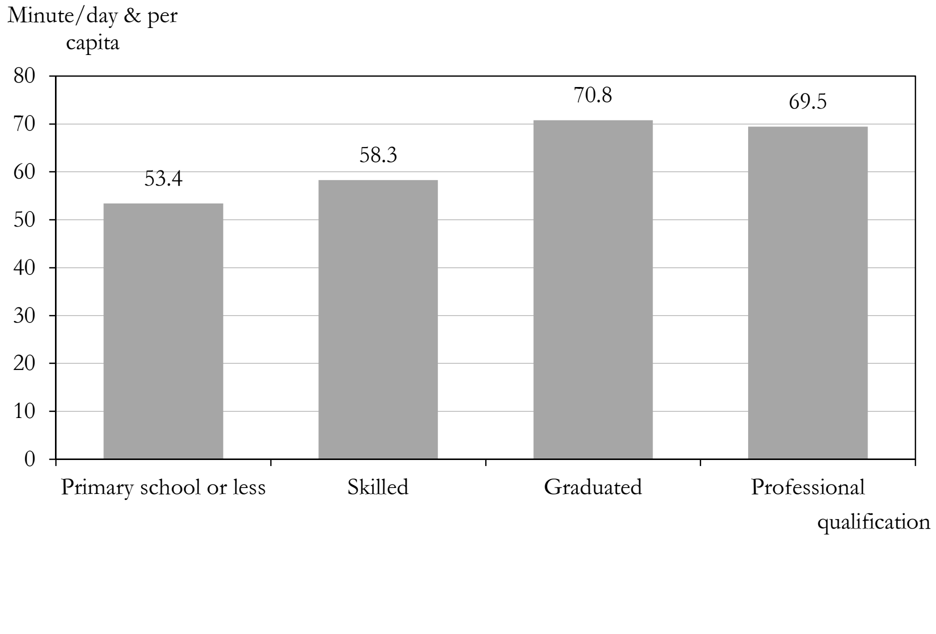
A	Daily time-use [minute/cap]	Population [11–85y range/2011]	Total time-use [minute]	Total time-use [per cent]
Employed	17.4	4,524,603	78,599,176	**44.4**
Pensioner	19.2	2,425,217	46671606	**26.4**
Child-care	22.7	340,669	7,722,523	**4.4**
Student	31.6	1,176,502	37,180,406	**21,0**
Unemployed	26.1	68,111	1,780,852	**1,0**
Other	22.7	214,983	4,882,532	**2.8**
Sum		*8,750,085*	*176,837,094*	*100.0*

Time-use bias by educational qualification in 2009/10

The last series of the explanatory variables present **transport time-use** differences by four educational qualification categories. *Figure 33* shows that people with more education generally spend more time ('A') on transport. While those with finished or unfinished primary school use daily 53 minutes for transportation, skilled personnel use 58 minutes, and graduate or professionals use about 70 minutes. (RSD = 13.6).

Figure 33

Per capita social time spent on transport by qualifications in 2009/10



As there are differences in the number of people belonging to the different qualification groups, the per capita time-use fails to represent their real share that appears as daily traffic. The real traffic share can be presented by multiplying the per-capita values with the volume of the people with the given qualification (Table 6). In the age range 10–84 years, 34% of the 8 750 000 people have a primary school certificate or less; 28.5% of them are graduates; 20.6% skilled personnel; and 16.9% are professionals. Figure 34 shows the share of the total transport time-use between the above groups, where the results are close to the above proportions. People with a primary school certification and graduates use around 30%–30% of the total transport time, while professionals and skilled personnel use close to 20%–20%.

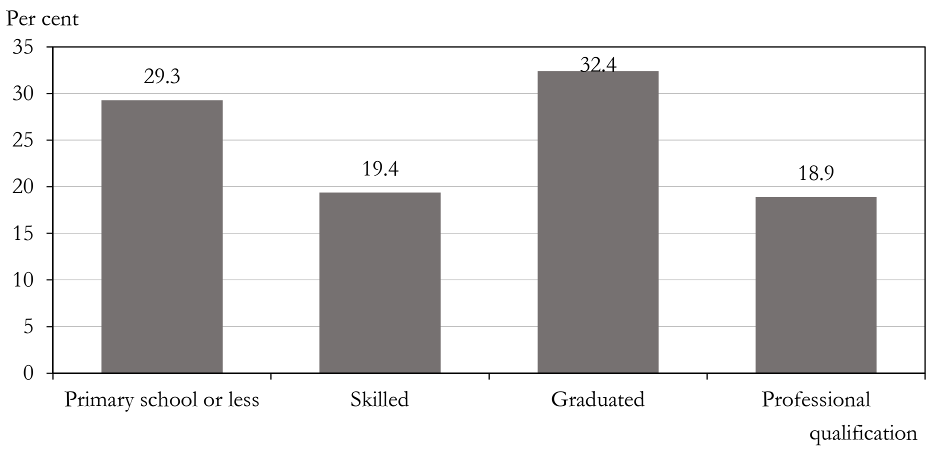
Table 6

Per capita and total social time spent on transport by educational qualification differences in 2009/10

'A'	Daily time-use [minute/cap]	Population [11-85y range/2011]	Total time-use [minute]	Total time-use [per cent]
Primary school or less	53.4	2,977,507	159,022,049	**29.3**
Skilled	58.3	1,805,051	105,266,263	**19.4**
Graduated	70.8	2,489,424	176,201,523	**32.4**
Professional	69.5	1,478,103	102,669,416	**18.9**
Sum		*8,750,085*	*543,159,252*	*100.0*

Figure 34

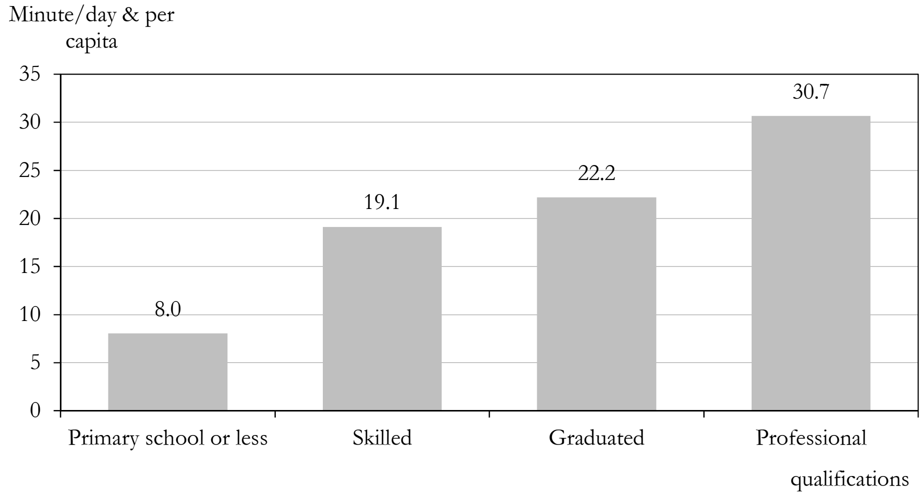
Share of the total time spent on transport by people's qualifications in 2009/10 (100% is the total time that the population spends on transport)



In the case of the **motor/car transport** mode, Figure 35 shows that the bias due to the qualification level is much more definite: the per capita daily social time-use varies between 8 minutes and 31 minute (RSD = 50.2%).

Figure 35

Per capita social time spent on motor or car transport by qualifications in 2009/10



Multiplication of per capita social time spent on motor or car transport with population share owing to bigger population ratio reveals that the graduates occupy a significant share of social car time-use: 34.7% as it is presented in Table 7. This group is followed by professionals with 28.5% time-use; skilled population covers 21.7% of the time; and those not educated after primary school use 15.1% of motor/car time.

Table 7

Per capita and total social time spent on motor/car transport by qualifications in 2009/10.

'A'	Daily time-use [minute/cap]	Population [11-85y range/2011]	Total time-use [minute]	Total time-use [per cent]
Primary school or less	8,0	2,977,507	23,957,051	**15.1**
Skilled	19.1	1,805,051	34,546,757	**21.7**
Graduated	22.2	2,489,424	55,231,968	**34.7**
Professional	30.7	1,478,103	45,316,047	**28.5**
Sum		*8,750,085*	*159,051,823*	*100.0*

Concerning **cycling,** the per capita social time ('A' series) is more-or-less the inverse of the motor or car transport case. While people with primary school education and skilled personnel cycle for an average of 7 minutes daily, graduates and professionals cycle for 4 minutes and 3 minutes, respectively (RSD = 38.3%). These are average estimations; those who really use their bicycle use it for about 45 minutes, independent of their level of education (RSD = 7.4%).

As for **walking**, the differences are relatively small: the skilled personnel and professionals walk for 18–19 minutes on an average daily ('A'), and graduates and people with primary education walk for 21 minutes and 24 minutes, respectively (RSD = 12.1%).

Comparing time-use bias between explanatory variables in 2009/10

Standard deviations (SD values)

If we want to see the extent to which the time-use of the people deviated at the different transport modes, then we have to use the total standard deviation values (SD).

After reviewing the time-use characteristics by two demographic (gender and age), two spatial (settlement status and county), and two social (activity and qualification) variables, we **compared the standard deviation (SD) values** (in Table 8 and Figure 36) that was produced by different variables in time use of transport as a whole and in time-use of each transport mode. The presented values relate to 'A' series, i.e. to the per-capita average daily social time-use values.

Table 8

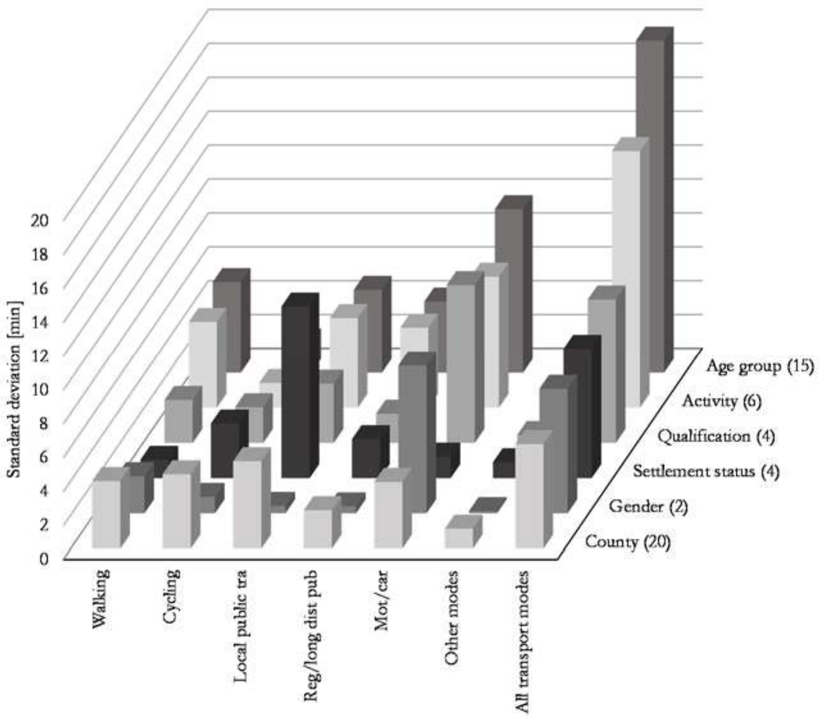
Comparing standard deviations [minute] caused by social, demographic, or territorial differences (based on average per capita daily transport time for the whole society [A] in 2009/10)

A	Walking	Cycling	Local public tra	Reg/long dist pub	Mot/car	Other modes	All transport modes
County (20)	3.98	4.38	5.15	2.26	3.93	1.17	6.16
Gender (2)	2.19	0.97	0.44	0.41	8.73	0.13	7.35
Settlement status (4)	1,08	3.22	10.15	2.30	1.27	0.94	7.65
Qualification (4)	2.54	2.11	3.50	1.73	9.34	0.47	8.49
Activity (6)	5,09	1.47	5.32	4.75	7.76	0.88	15.17
Age group (15)	5.37	1.81	4.91	4.22	9.67	1.11	19.61

The last row on the right side of Figure 36 relates to the time-use of **all transport modes**. We can see here that the biggest deviation values of the total time-use are produced by age and the activity. Almost all modes produce relatively big deviations in these two categories. Out of that, specific big deviations were produced in the case of motor/car transport by the qualification and the gender; in the case of local transport by the settlement status (and also by county); and in the case of cycling by county and settlement status.

Figure 36

Comparing standard deviations (SD) [minute] caused by social, demographic, or territorial differences (based on average per capita daily transport time for the whole society [A] 2009/10)



Important Annex: Multidimensional variance analysis: ANOVA

Concerning the recommendations of this study, we were able to explain a significant portion of the potential heterogeneity in the transport time-budget with selected independent variables. However, we studied two-dimension section planes of a multidimensional system and did not answer the main question about the percentage of the total transport time-use heterogeneity that was explained by the selected variables.

There are problems with the multidimensional handling of the selected system. The three most important difficulties are as follows: (1) the six explanatory variables we used are not independent, i.e. the interactions between the variables is significant (the activity or education or settlement positions influence the results seen in the gender case, etc.); (2) the variables are mainly quality and category variables (activity, counties, genders) demanding special treatment; and (3) as presented in Figure 5, the independent variable does not follow a normal distribution, but a characteristically asymmetric one. The first two problems were handled by choosing different ANOVA processes, while for the third one, we ran both the original values and the logarithmic transformed values (where the distribution was normal) and obtained very similar results. As a technical step, the ages were reduced to four groups and the counties to five geographic units.

Studying the whole transport time-use together ('A') and filtering the cross-effects of the variables, the E^2 variance showed that **only 9.7% of the heterogeneity was explained by the six explanatory variables**.

Looking at the effects of the explanatory variables separately, Table 9 presents the results.

Table 9

Factor summary to present the effect of the explanatory variables on the variations in transport time use (per capita daily social time-use in 2009/10)

Factor Summary			
		Eta	Beta Adjusted for Factors
Transport	age4	0.24	0.11
	qualification4	0.119	0,064
	activity6	**0.274**	**0.18**
	gender2	0.086	0.055
	settlement status4	0.092	0.046
	county5	0.099	0.057

Beta shows the partial effect of an explanatory variable: the percentage to which the given explanatory variable is able to explain from the heterogeneity of the dependent variable, if the other explanatory variables are fixed. By that calculation, the *'activity'* group shows the biggest effect, where the Beta-square is 3%, more than the sum of all the Beta-squares belonging to the five other variables.

The findings on the single modes support the proportions presented in Figure 36. In the case of the bikers, the geographical situation (*county*) shows the biggest effect (Beta = 0.195). For the motor and car transport *activity* (Beta = 0.143) and *gender* (Beta = 0.135); for the local transport the *settlement status* (Beta = 0.216), etc. The difference is that here it is also visible that even these 'high' values explain only several percentages from the total heterogeneity. Here, we must refer back to Figure 5 where it was presented that a decisive part of the deviations within the time-use patterns is regulated by a strict connection between the expected time-distance and the choice of the destinations.

Conclusions

Based on the last three Hungarian time-budget survey data and on the international and domestic experiences, the authors analysed the Hungarian transport time-use trends and distributions. An important starting point was that—opposite to the traditional general idea—developed transport technologies do not seem to save time for the society. In this regards, the time-budget used for transport is quite stable. A first and simple control supported this view: the average Hungarian citizen in the age group 15–74 has spent 61.8, 59.4 and 65.2 minutes on transport on an average day in 1986/87, in 1999/00, and in 2009/10, respectively, while the similar data for 1977 was 63 minutes (for population aged 15–69 years). Otherwise, one-third of the hour-long average daily transport time is covered by walking, 30% by motor/car transport, 16% by long distance public transport, and 8–9% by cycling and by local public transport each. The share of the motor/car transport almost doubled in the last 25 years, while the share of other modes proportionally decreased. It is also interesting that, at each mode, those using the given mode spend more time with that transport than they did earlier—the decrease comes from the fact that less people use the given modes.

Unlike the relative trend-stability of the averages, the personal time-use values show huge variance even in a single survey. The biggest diversity originates from the fact that the daily transport activity of the people consists of numerous short motions and much less longer trips. The daily travel durations show a very asymmetric distribution, resembling a scale-free function. To study this function and its principles is a great challenge, but it needs a wider, transport destination based survey than just the statistical background of the time-budget surveys. We did not examine this wider subject further at the moment.

The authors focused on two secondary but important relations. The first was the effects of different demographic, social, and spatial backgrounds on the transport time-use variability. The second was the specialities in the time-use character of the different transport modes.

Considering the social background, it was important to learn that the six explanatory variables together (age, gender, activity, qualification, county, and settlement-category) could explain only 10% of the heterogeneity of the time-use values. It was the *age* and the *activity* status that influenced the time-use pattern best. As for the age groups: between age 20 and 80 years the average daily transport time-use decreased from 90 minutes to 30 minutes. Looking at the activity, 60% of the time-use of the population is produced by the 51.7% employed people.

Concerning the modal specialities, *motor/car users* of the age group 25–44 years (33.2% of the people) produced half of the total traffic time; the men in 25–44 age group (16.5% of the people) produced 33% of the related traffic time. Looking at the activity, 77% of the total motor/car traffic time-use was produced by the employed group.

In *bicycle use,* there are huge time-differences. An average resident of the plain counties uses the bicycle daily for 6–17 minute, which is 10–25 times more than the cycle use by an average Budapest resident. The bikers use the bicycle for 40–50 minutes a day, when they use it. The difference in average cycle use comes mainly from the share of the users. While 15% of the people use bicycle in villages, only one-tenth of that share uses cycle in the capital. In addition, this almost 1.5% share is still three times more than a similar share 23 years earlier.

As for public transport, apparently, *local public transport* is used first of all by those living in cities, where there is access to local transport. An average Budapest resident spends 27 minutes on this mode daily, 4–5 times more than the capacity of a village or small town resident to avail of this mode. Contrarily, in case of *long distance public transport*, village residents spend almost three times more time on this mode than the residents of Budapest or of the county centres do.

The final part of the paper use an ANOVA model to filter out the cross dependencies among the explanatory variables. The beta factor shows that, within the 10% explanatory effect of all variables, *activity* is the only variable that is worth mentioning due to its own partial effect on the transport time-use heterogeneity.

Acknowledgements

The authors hereby thank the Central Statistical Office (KSH) for providing an opportunity to use the database, and the Centre of Economic and Regional Studies of the Hungarian Academy of Sciences (CERS HAS) for allowing access to the databank, which also contained elder datasets of the TÁRKI. We are personally grateful for the help of Tibor Czeglédi in the collection of a specific demographic dataset.

REFERENCES

ANDORKA, R.−FALUSSY, B.−HARCSA, I. (1990): Időfelhasználás és életmód In: ANDORKA, R.–KOLOSI, T.−VUKOVICH, GY. (eds.) *Társadalmi riport 1990* pp. 192−207. TÁRKI, Budapest: http://www.tarsadalomkutatas.hu/kkk.php?TPUBL-A-288/publikaciok/tpubl_a_288.pdf (downloaded in March 2016)

EUROSTAT (2009): *Harmonised European Time Use Surveys 2008 Guidelines* Office for Official Publications of the European Communities, Luxembourg. http://ec.europa.eu/eurostat/documents/3859598/5909673/KS-RA-08-014-EN.PDF/a745ca2e-7dc6-48a9-a36c-000ad120380e?version=1.0 (downloaded in March 2016)

GERIKE, R.–GEHLERT, T.–LEISCH, F. (2015): Time use in travel surveys and time use surveys – Two sides of the same coin? *Transportation Research Part A: Policy and Practice* 76: 4–24.

HÄGERSTRAND, T. (1970): What about people in regional science *Papers in Regional Science* 24 (1): 7–21.

HÄGERSTRAND, T. (1985): Time-geography: focus on the corporeality of man, society, and environment In: SHŪHEI, A. (ed.) *The science and praxis of complexity: contributions to the symposium held at Montpellier, France, 9–11 May, 1984*. pp. 193–216. United Nations University Press, Tokyo.

ILLICH, I. (1974): *Energy and Equity. Harper & Row* Calder & Boyars Ltd., GB https://www.uvm.edu/~asnider/Ivan_Illich/Ivan%20Illich_Energy%20and%20Equity.pdf (downloaded in March 2016)

JARA-DÍAZ, S.–ROSALES-SALAS, J. (2015): Understanding time use: Daily or weekly data? *Transportation Research Part A: Policy and Practice* 76: 38–57.

KSH (2010): *Időmérleg-módszertan* Központi Statisztikai Hivatal, Budapest. http://www.ksh.hu/docs/hun/xftp/idoszaki/pdf/idomerleg_mod.pdf (downloaded in March 2016)

KSH (2012): *Időmérleg 2009/10: Összefoglaló adattár* Központi Statisztikai Hivatal, Budapest. http://www.ksh.hu/docs/hun/xftp/idoszaki/idomerleg/idomerleg0910.pdf (downloaded in March 2016)

LAKATOS, M. (2013): *A foglalkoztatottak időfelhasználása az ingázás és a munkába járás idejének tükrében* Műhelytanulmányok 3. Központi Statisztikai Hivatal. Budapest. http://www.ksh.hu/docs/hun/xftp/idoszaki/pdf/muhelytanulmanyok3.pdf (downloaded in March 2016)

JONES, P.–BOUJENKO, N. (2009): *'Link' and 'Place': a New Approach to Street Planning and Design* Australian Transport Research Forum, Sydney. http://atrf.info/papers/2009/2009_Jones_Boujenko.pdf (downloaded in March 2016)

SZALAI, A. (ed.). (1972): *The use of time: daily activities of urban and suburban populations in twelve countries* Publication of the European Co-ordination Centre for Research and Documentation in the Social Sciences Mouton, The Hague-Paris.

WHITELEGG, J. (1993): Time Pollution *Resurgence & Ecologist* 23 (4): 131–134.

Migration settlement networks in the Carpathian Basin, 2001–2011

Áron Kincses
Hungarian Central Statistical Office,
Budapest, Hungary
E-mail: aron.kincses@ksh.hu

Lajos Bálint
Hungarian Demographic Research Institute
Budapest, Hungary
E-mail: balint@demografia.hu

Looking at the relationship between the place of birth and current residential locations of foreign citizens arriving in Hungary from the neighbouring countries, in general, we establish that smaller migration distance involves migrants with a lower level of education, while preference for longer distances is determined by higher qualifications of migrants. The potential impact area of migrants grows in line with the education attainments of migrants.

A scale-free settlement topology can be seen from the neighbouring countries of immigration to Hungary. This means that most of the settlements of Hungary have just a few links to settlements of neighbouring countries, from a migration point of view, while few Hungarian settlements have many connections. This finding also means that, instead of the national migration strategy, the subsidiary and the regional strategies can play a decisive role in the management of the international migration process.

Keywords:
international migration,
Carpathian Basin,
network analysis

Introduction

The key to migration settlement networks is linking the source countries to the destination areas.

The analysis is based on the data on foreign nationals that is collected from the census and the Office of Immigration and Nationality. Establishing the linkage between the two databases has created an opportunity to connect the study on migration settlements with the detailed information material of the census on the subject.

In this study, we strive to show in detail the characteristics of the international migrants' settlement network in relation to Hungary and its neighbouring countries and connect them to the attributes of migrants.

Analysing the relationship between the Carpathian Basin migration source and destination areas, our aim is not only to explore the regional features of the flows between a given emigration country and Hungary but also to depict a general territorial network of contacts in an integrated way, simultaneously taking into account all the neighbouring countries. In this network, which is considered as the specific sub-network of the entire Carpathian Basin network, we will examine the settlement networks of the migrants, mostly of Hungarian ethnicity, arriving to Hungary. The question is whether it is characteristic that some parts of the aforementioned territorial network are increasingly 'utilised' by the foreign citizens living in Hungary and having specific demographic, labour market properties, or given territorial connections are independent of the characteristics of migrants.

Data, relations of source, and destination areas

Census data represent a detailed demographic, labour market, and sociological data set on the population of migrant origin living in Hungary. However, information on the places of emigration and birth only country is available. Relevant Hungarian migration databases (database of Office of Immigration and Nationality and the HCSO data files) contain less information on the characteristics of migrants, but extend to their places of birth. The establishment of link between the two databases enables to connect the analysis on the emigration areas with the detailed census information material. The method is based on the use of a complex conversion key between databases, which assigns the municipalities in foreign places of birth[1] to the census data files according to the common variables (nationality, residence in Hungary, date of birth, gender, and marital status). Thus, although data on foreign settlements underlying the specific analyses was available, a separate classification became necessary. It is because such classification often contained the denominations of settlements or parts thereof in different languages.

Hereafter, we conduct a detailed examination of the relationship between the place of birth and current residence of foreign citizens coming to Hungary from neighbouring countries at Nomenclature of Territorial Units for Statistics (NUTS3) level[2], using the 2001 and 2011 census data (Tóth 2013). As for Ukraine, instead of

[1] As a first step, the two databases have been connected by all the five variables. With this method, most records could be corresponded to each other. In the second step, the non-connected rows have been unified with the help of four-element keys (citizenship, the settlement of the Hungarian residence, and two-element subsets of all three-element variable sets). In some remaining (not connecting) cases, we have used the less differentiated level of district instead of the settlement of the Hungarian residence, thus elaborating five and four-element keys with this method. The application of this method resulted in formation of the new, interconnecting database.

[2] This territorial classification is available for most countries, with the exception of Serbia and Ukraine. In case of Serbia, the most probable NUTS3 division is applied (For the territorial division of the Serbian Republic, the classification published in the 2011 Statistical Yearbook has been used. For detailed information, see: http://www.media.srbija.gov.rs/medsrp/dokumenti/SGS2011_cyr.pdf), while in Ukraine such a classification does not exist. The 'oblaszt' level is more integrated and the 'rayon' is more detailed than this classification. A more defined division will be applied to *Transcarpathia as it has a distinguished role within Ukraine (the vast majority of migrants come from here).*

the whole country – *due to its large size* – only Transcarpathia has been involved in the study since nearly 90% of Ukrainian migrants arrive from this region.

From the 161 regions established, in this manner significant concentrations can be observed in the matrix of migrations to the 19 Hungarian counties and Budapest. Extracting those region pairs, which contribute to the total turnover and have more than 0.5% of the total migration, we get a much tighter group than the previous one. Thus, 43% of migrations are concentrated in 1% of all the matrix cells (42x20=840) in 2011. This proportion increased by nine percentage points until 2011.

Central Hungary was the most attractive destination for those arriving from the counties of Transylvania in 2011. Nearly, 4% of all immigrations from Romania to Hungary took place between Harghita county and Budapest, while the share of moves between Mures County and Budapest was 3%. The border areas were of considerable importance as well, which can be explained partly by the phenomenon of circulation migration (Fercsik R 2008, Elijah S. et al, 2009, Illés S. et al, 2009) and partly by the ease that borders offer to migrants for maintaining a connect with family members staying at home (Rédei M 2007). Intensive flows (Anderson et al, 1999, Baranyi, B. et al. 2004, N. Hansen 1977; Van Geenhuizen, M. et al, 2001) and transnational areas were formed (Melegh, 2011) between the bordering counties. Among them, the most significant movements included North Backa, North Banat-Csongrád (2,37% and 1.16%) as well as Beregovo and Uzsgorod rayon-Szabolcs-Szatmár-Bereg county (2.42% and 1.35%). During the same period, a population of mostly Hungarian ethnicity came from the large emitting areas to Hungary.

The number of pairs of regions, which are affected by more than 0,5% by the migrations from neighbouring countries to Hungary showed an increase by 2011. The spaces of migratory relations increased in Hungary, with remoter parts of the neighbouring countries joining the source areas, thereby increasing the regional role of Trnava, Bratislava, Košice, Nitra districts, and Vienna'. (Estélyi K. et al. 2006). The significance of Budapest and Pest County further strengthened, while the weight of migration of Szabolcs-Szatmár-Bereg county and, to a certain extent, of Csongrád weakened. Generally, in these counties the number of migrants also increased, but not as much as in Central Hungary. By 2011, the migration ratios between Harghita, Mures, and Budapest continued to increase (3.8% and 4.5%, respectively). Generally, the cross-border relationship levels retreat. Often, those arriving from the proximity of the borders do not migrate to the neighbouring region of Hungary but to Central Hungary, which provides better structural (labour market and income) facilities.

Table 1

Proportions of major migration flows*, 2001

(%)

Foreign/ Hungarian counties	Buda-pest	Bács-Kiskun	Békés	Csong-rád	Fejér	Hajdú-Bihar	Pest	Sza-bolcs-Szatmár-Bereg
Bihor	1.63	0.18	0.39	0.10	0.18	0.70	2.16	0.20
Brasov	1.25	0.08	0.05	0.11	0.09	0.03	0.21	0.03
Cluj	2.55	0.14	0.19	0.18	0.50	0.35	0.74	0.20
Covasna	1.32	0.12	0.18	0.07	0.16	0.33	0.46	0.08
Harghita	3.75	0.46	0.51	0.65	0.41	0.40	2.61	0.30
Maramures	0.92	0.06	0.05	0.06	0.12	0.05	0.43	0.09
Mures	2.77	0.44	0.47	0.65	0.29	0.32	0.91	0.22
Salaj	1.00	0.06	0.04	0.04	0.16	0.23	0.25	0.06
Satu Mare	1.92	0.11	0.15	0.16	0.34	0.72	0.37	0.84
Timis	0.07	0.02	0.06	0.11	0.03	0.03	0.67	0.02
Trnava district	0.51	0.00	0.02	0.02	0.00	0.02	0.12	0.03
Berehove rayon	0.85	0.08	0.05	0.06	0.13	0.34	0.28	2.42
Vynohradiv rayon	0.56	0.10	0.06	0.06	0.01	0.15	0.21	0.90
Uzshorod rayon	0.67	0.06	0.11	0.03	0.10	0.30	0.42	1.35
Juznobacki district	0.79	0.36	0.10	0.85	0.08	0.02	0.08	0.02
Severnobacki district	0.71	0.50	0.30	2.37	0.11	0.05	0.12	0.01
Severnobanatski district	0.68	0.48	0.14	1.16	0.45	0.13	0.32	0.01

* Total number of those coming from the regions of all neighbouring countries to all Hungarian counties is 100%.

Table 2

Proportions of major migration flows, 2011

(%)

Foreign/ Hungarian counties	Buda-pest	Bács-Kiskun	Borsod-Abaúj-Zemplén	Csong-rád	Győr-Moson-Sopron	Hajdú-Bihar	Komá-rom-Eszter-gom	Pest	Sza-bolcs-Szat-már-Bereg
Bacau	0.67	0.04	0.00	0.16	0.03	0.02	0.14	0.50	0.00
Bihor	0.60	0.09	0.06	0.47	0.19	1.54	0.26	0.52	0.41
Brasov	0.19	0.05	0.00	0.01	0.05	0.03	0.02	0.76	0.16
Cluj	0.91	0.35	0.02	0.03	0.02	0.10	0.03	0.43	0.03
Covasna	2.97	0.31	0.05	0.32	0.28	0.05	0.19	1.81	0.02
Harghita	3.81	0.81	0.09	0.25	0.54	0.15	0.21	3.53	0.16
Hunedoara	0.90	0.39	0.01	0.11	0.01	0.00	0.04	0.21	0.01
Mures	4.52	0.70	0.09	0.73	0.29	0.17	0.14	1.57	0.04
Salaj	1.43	0.07	0.03	0.02	0.04	0.38	0.01	0.62	0.06
Satu Mare	1.77	0.19	0.06	0.11	0.10	0.40	0.08	0.85	0.86
Banska Bystrica district	0.55	0.01	0.12	0.04	0.31	0.04	0.06	0.14	0.01
Kosice district	0.39	0.03	0.74	0.03	0.26	0.07	0.04	0.08	0.08
Trnava district	0.65	0.01	0.02	0.01	0.46	0.06	0.07	0.14	0.03
Nitra district	1.30	0.07	0.03	0.04	0.96	0.02	0.93	0.44	0.03
Bratislava district	0.09	0.02	0.03	0.00	1.09	0.00	0.06	0.03	0.00
Berehove rayon	2.68	0.13	0.12	0.09	0.11	0.25	0.10	0.60	0.94
Vynohradiv rayon	0.94	0.09	0.09	0.01	0.07	0.21	0.05	0.49	0.94
Uzshorod rayon	0.83	0.09	0.10	0.02	0.01	0.24	0.07	0.22	0.52
Severnobacki district	1.02	0.75	0.02	0.75	0.09	0.01	0.00	0.26	0.00
Severnobanatski district	0.66	0.29	0.04	1.93	0.07	0.04	0.03	0.23	0.02
Wien	0.04	0.02	0.00	0.01	0.56	0.01	0.03	0.08	0.00

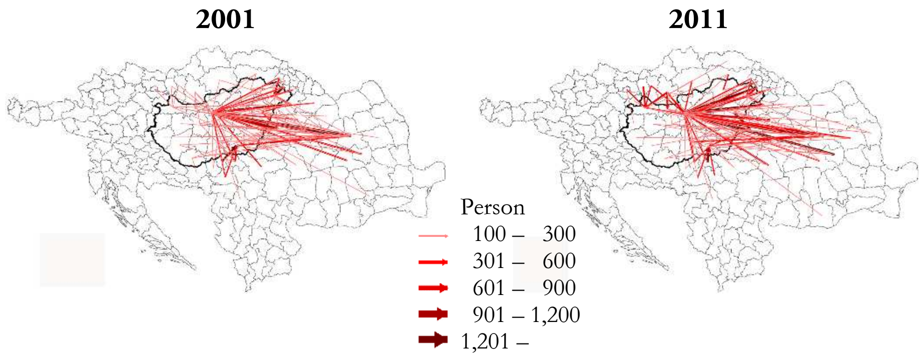
The increasing attraction of the capital area is perceptible not only in the case of major emitting regions but also almost in the entire Carpathian Basin (Rédei M. 2009). This Hungarian area is *unequivocally* a destination for the migrants, even those coming from remote geographical destinations (Soltész B. et al. 2014). This statement is especially valid for working age migrants, those having higher education, working in leadership positions, and living in households without children. The major attraction of Central Hungary is a partial consequence of the magnitude of population and partially, in the most important way, due to its economic attractiveness: advantageous labour market and income positions. The focussed nature of flows underlines the economic motivations behind a significant portion of the migratory movements; however, the cultural and educational opportunities might have also contributed to the attractiveness of the destination. The border areas are rather considered as local destinations. In case of small geographic distances and movements near the border, the rate of those moving with their children is higher. The distance factor displays the cost and partly the risks involved in migration. In this light, it is not surprising that people that are more educated undertake or are capable of visiting distant destinations (e.g. Budapest and Pest county, which are the same in terms of labour market). However, shorter distance is less costly (shorter distance provides greater access to local knowledge), less risky, and provides easier mobility. Most likely, these reasons might contribute towards the fact that families with children prefer the short distance option. For example, this option enables them to reserve the possibility of non-monetary transfers from the family (e.g. family assistance, contribution towards caring, and making provisions for children), which cannot be utilised due to the cost factors involved in case of greater distances. In case of near-border movements, the occupations and levels of education of the migrants are more diversified; however, the differences between the economic activity of short-distance and long-distance migrants are not significant. Therefore, the differences are not primarily in terms of employment, but in the nature of occupations. Migrants who worked in occupations that did not require educational qualifications could be found almost everywhere in the country by 2011, and such migrants gained prominence in studies focussing on cross-border relationships. In our opinion, in the case of migration among the Carpathian regions, the settlement of less skilled labour and families with children clearly underlines the special consideration given to the cost factor involved in the migration distance. The role of distance as a cost factor is markedly present in the theoretical tradition of migration. Among classical authors, Sjaastad (1962) assumed that the cost of moving is proportional to the distance of migration. Distant moves are accompanied by higher risk or significant monetary and psychic costs, while the shorter ones have the opposite effect. Therefore, it is real that families with children have preferred short-haul destinations while reserving their previous connections and networks (for example, seeking assistance from parents located across the border).

Depiction of migration relations

The most common way to display migration links is through flow charts (Tobler 2003; Daróczi–Bálint, 2015). The charts presented here—due to the one-way connections—correspond to the so-called 'Desire line' maps that indicate the connections of given points using straight lines. Counties across the border and those within Hungary have been identified with their geometrical centres. The lines point to the destinations of those arriving in Hungary and the connection structure of the counties across the border. For the 'readability' (transparency) of the map, the representation under 100 migrations was disregarded. The thickness of the lines and the differences in hue, similar to that applied in thematic maps, express the difference in the number of migrations.

Figure 1

Area relations between source and destination regions*



* The maps illustrating the networks have been prepared with the help of QGIS software FlowMapper plugin (Güllüoğlu 2013).

Figure 2

Relationship between regions of the place of birth and current Hungarian residence of foreign citizens above 24 years, by education

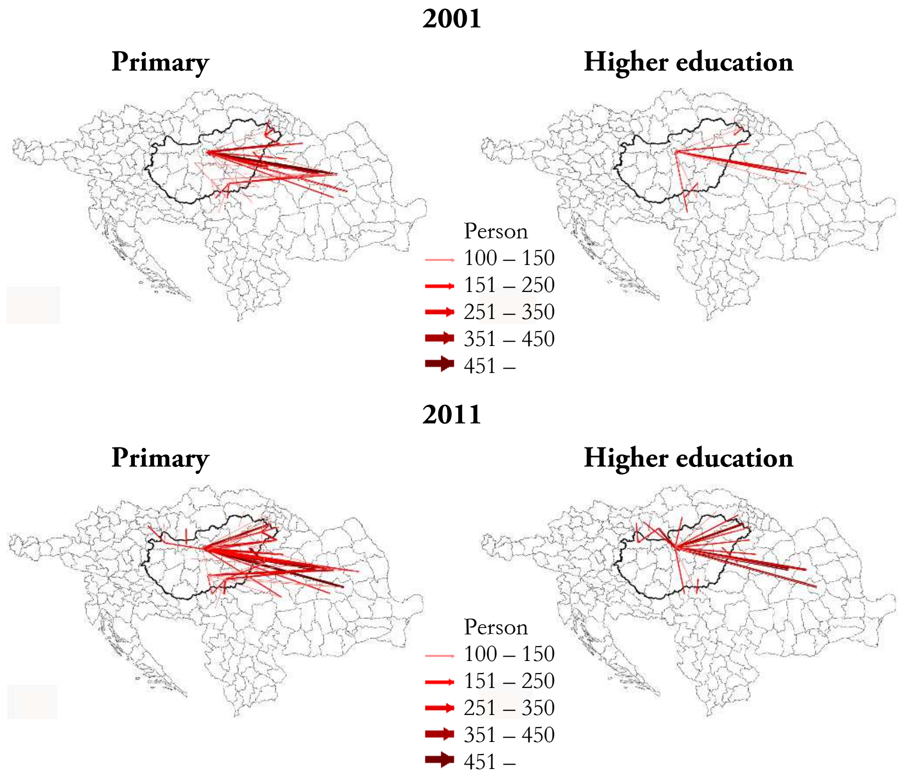
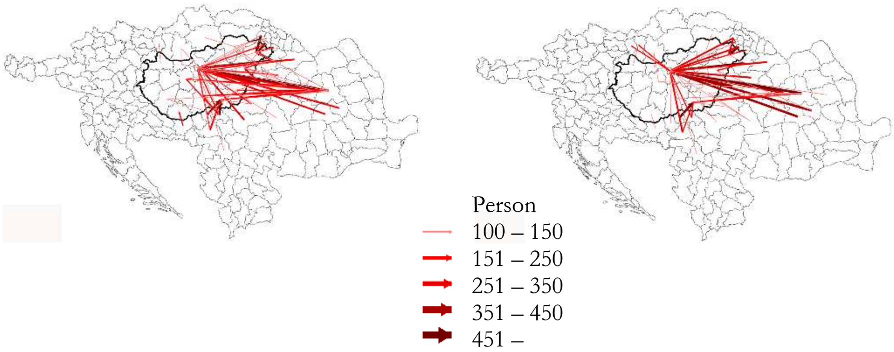


Figure 3

Relationship between regions of the place of birth and current Hungarian residence of foreign citizens above 24 years, by type of household

2001

Households with children **Households without children**



2011

Households with children **Households without children**

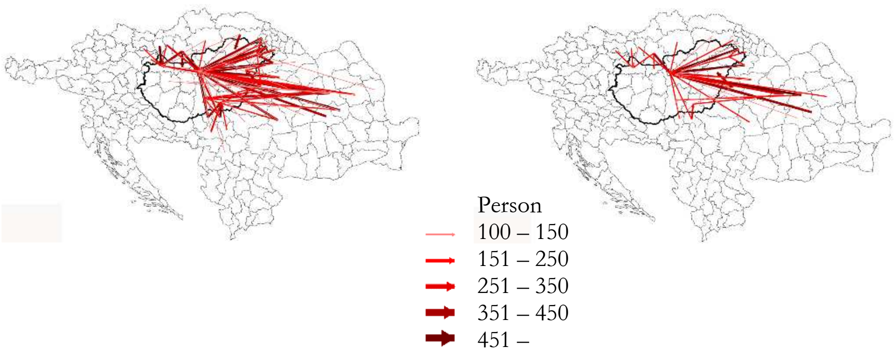


Figure 4

Relationship between regions of the place of birth and current Hungarian residence of foreign citizens aged 25–64 years, by economic activity

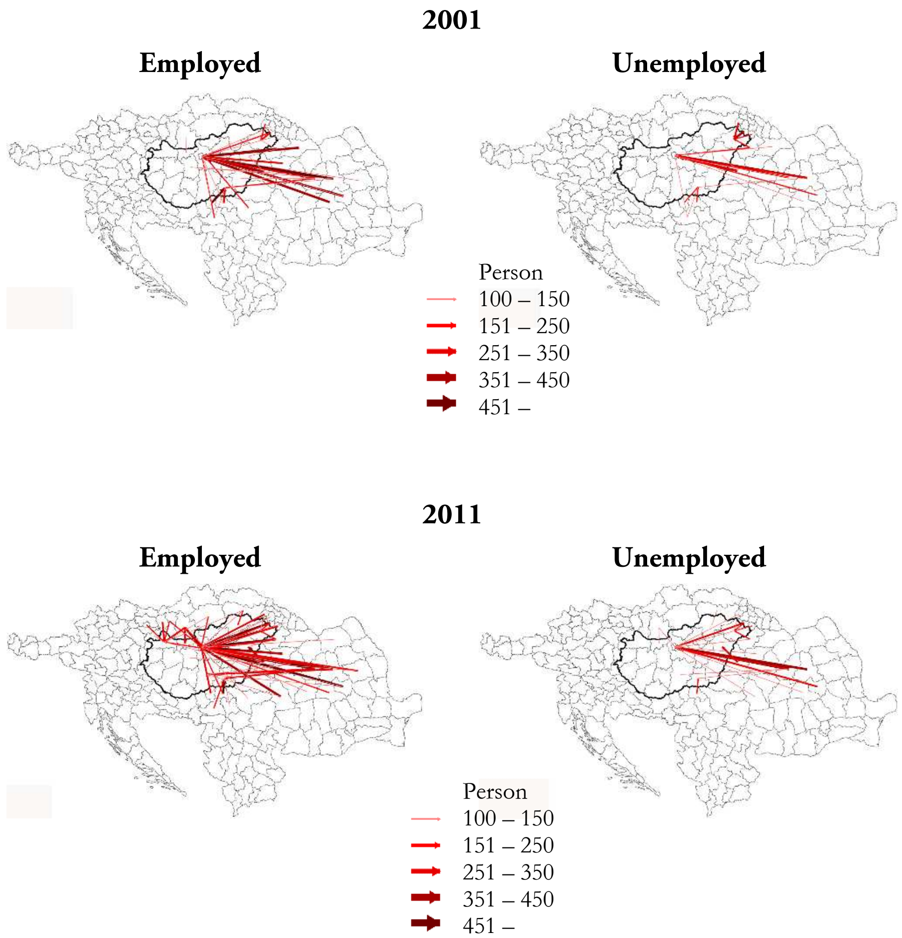
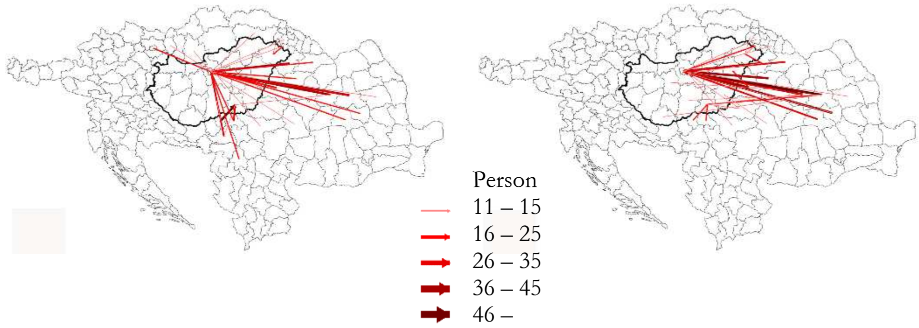


Figure 5

Relationship between regions of the place of birth and current Hungarian residence of foreign citizens aged 25–64 years, by occupation

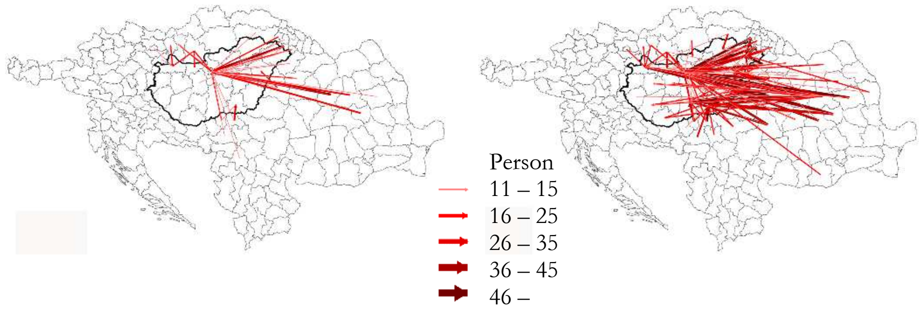
2001

Economic, administrative managers, advocacy leaders | **Simple occupations (not requiring professional skills)**



2011

Economic, administrative managers, advocacy leaders | **Simple occupations (not requiring professional skills)**



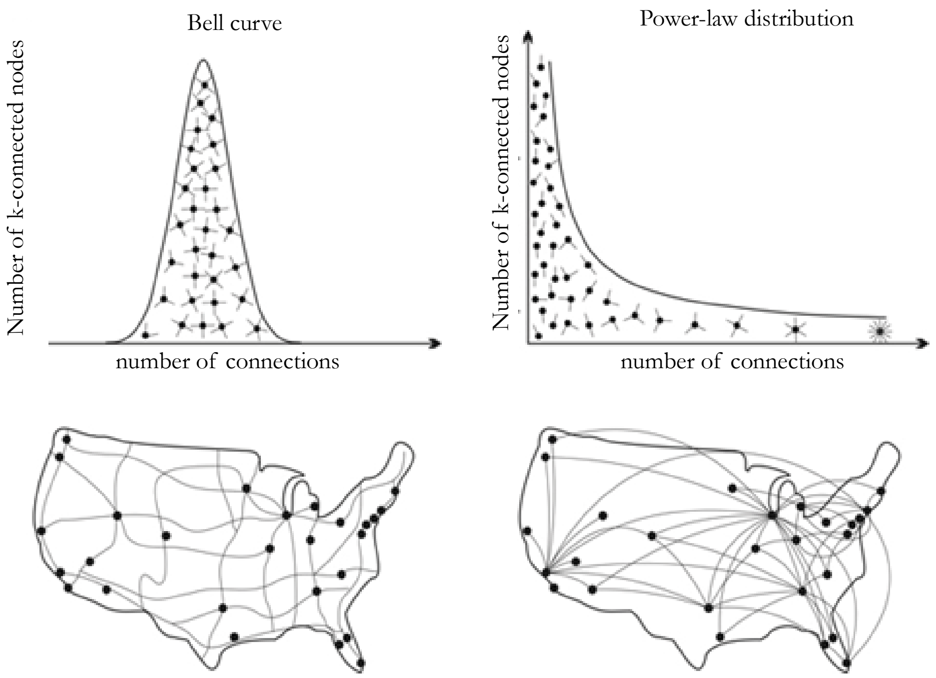
Networks of migration settlements

In the previous section, we have strived to establish the relationship between regions of the place of birth and current Hungarian residence of foreign citizens, at the NUTS3 level, arriving from neighbouring countries in Hungary and the demographic and labour market variables of migrants. Currently, we have tried to analyse in detail the internal features of the settlement networks. Network analysis started in the second half of the twentieth century (Erdős P. et al, 1959, 1960; Bollobás B. et al, 1976). One of the most interesting theses of this era representing a paradigm changing (M. Buchanan, 2003) is that any two people on earth are connected by six degrees of separation (acquaintances connection). Subsequent to these initial theses of graph theory, the network theory gained recognition as a new discipline due to its new abstractions. This was based on the findings, according to which every network, living or dead, existing in nature or artificial, comes into being partly according to identical organising principles. That means that the Internet, human relations, and neural network of the brain are very similar in their properties (Barabási A. L., 2008).

Degree distribution (links distribution) of random systems follows a bell curve. In other words, the majority of points have about the same number of connections and the probability of the existence of those points having large and a small number of connections is insignificant (distribution of people according to height produces this curve, and the probability of the existence of extremely tall or extremely short people is small). In general, a road network of a country—where the connections of the settlements refer to specific points and the interconnection of roads refer to the links—resembles of a random network. Precisely, the scale-free network (according to a power law), like in most networks, describes the relationships between people. Here, the degree distribution according to a power law predicts that the highest number of points (people) have only a small number of different social connections (the majority of people do not know ten thousand or hundred thousand other people in person), which is held together by some hubs facilitating large interconnections. A good example for this is the map in the bottom right of the following figure that shows the air traffic system: many small airfields can be connected by a few main hubs (Barabási A., 2008.)

Figure 6

Random ad scale-free networks



Source: Barabási Albert-László (2008): Behálózva.

For us, the network theory is important from the point of view of migration because of the links between the settlements, which is connected by international migration that is affecting Hungary. It means that the settlements are the vertices of the graph.

Two settlements are connected by an edge (they are linked) if international migration has taken place between the two settlements, i.e. someone has migrated from a settlement by moving to another (Hungarian) settlement, regardless of the number of migrants (probably personal relationship among migrants shows similar attributes, but data compilation on such personal contacts of migrants is unavailable). We consider it important to analyse the connection network of Hungarian host settlements in the Carpathian Basin because it shows the extent to which a migration is diversified and the degree of 'embeddedness' in the settlement.

On October 1, 2001 (referring to the date of the census), Budapest had the most Romanian migration connections. Romanian citizens arrived in the capital from 378 different settlements of Romania. Debrecen had the second most links (96), followed by Szeged (87), Békéscsaba (64), and Nyíregyháza (60). By 2011, Budapest further increased the number of its links (417), while Debrecen slightly relegated (69). However, e Érd (71), Szeged (87), Budaörs (48), Kecskemét (55), and Szigetszentmiklós (55) significantly strengthened the settlement links, thereby enhancing the attractiveness of Central Hungary. Degree of the nodes Békéscsaba – Gyula and Debrecen – Nyíregyháza is decreasing, and their place is taken over by the villages surrounding Szeged and Kecskemét. The network's centre of gravity has shifted to the West in the period between the two censuses, but Eastern Hungary, Budapest, and Pest County remained the two key areas.

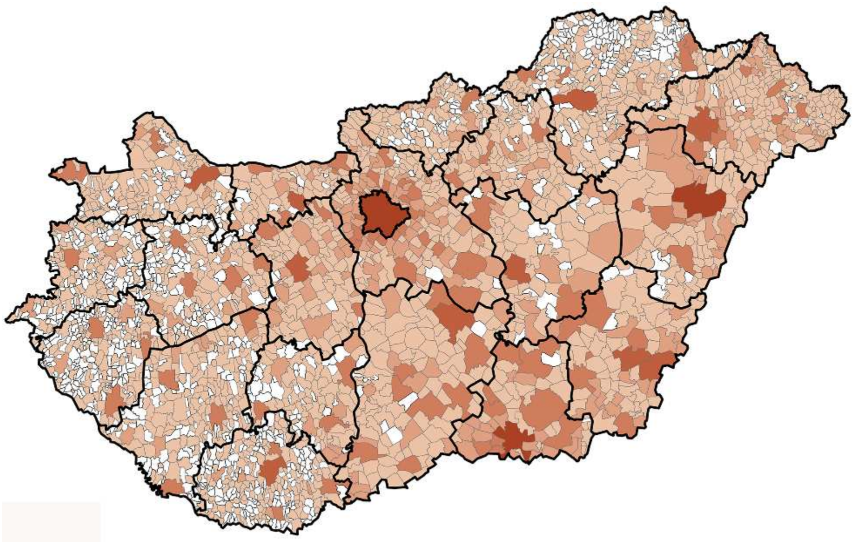
In case of Serbia, it is also true that the capital has the settlement links of the highest degree (in 2001: 94, in 2011:109). Connectedness of Szeged (in 2001: 37, in 2011: 97) is the second largest, but more Serbian citizens (2047 people) live here than in the capital (1801 people). This means that more people arrive in Szeged from fewer Serbian settlements near the border (on an average in greater numbers by settlement), while migrants come to the capital from many places but on an average in smaller numbers. In the regional relations, a slight increase can be observed between the two censuses for Pécs (33 to 38), Kecskemét (27 to 37), and Baja (25 to 28), while Kiskunhalas (25 to 22), Hódmezővásárhely (from 22 to 14), and Röszke (from 22 to 8) show significant decreases.

In case of Ukraine, the number of links of the major cities along the Hungarian border does not increase significantly, while Budapest and several settlements of Pest County are characterised by growth. However, the order among the most connected settlements remained mostly unchanged. Thus, the following ranking is given to the settlement networks: Budapest (145 to 221), Debrecen (45 to 76), Nyíregyháza (46 to 65), Kisvárda (40 to 38).

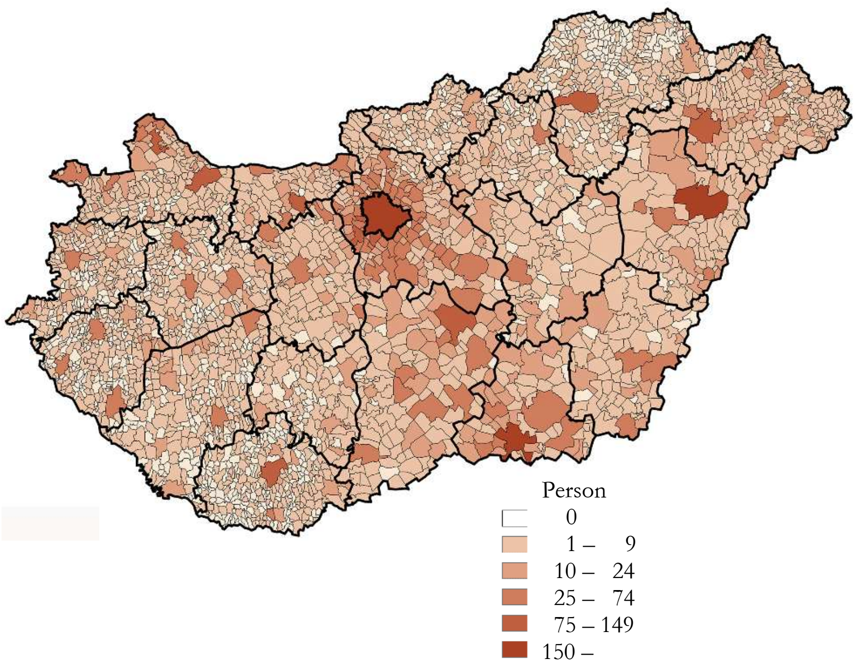
Figure 7

Number of relationships between migration settlement networks in Carpathian Basin

2001



2011



Other neighbouring countries do not facilitate much connectedness among the settlements (and migrants) in Hungary. Here, the capital's growing dominance can be observed. In case of Slovakia, the links with Érd (4 to 71) and Budapest (64 to 106) developed in the most dynamic way in 2011. Besides Győr (13 to 37), the connectedness of Miskolc (from 16 to 28), Esztergom (13 to 22), and Komárom (16 to 26) registered an increase. As for the Austrian settlements, they have the most pronounced connections with Budapest (growth: 31 to 164). In addition, the dynamics of Mosonmagyaróvár (7 to 43), Kaposvár (1 to 23), Lenti (3 to 18), Győr (5 to 13), Hegyeshalom (3 to 21), and Fertőd (1 to 21) worth mentioning. Croatia's migration settlement links with Budapest (28 to 34), Győr (1 to 17), and Harkány (9 to 16) have been intensified, while those of Pécs (24 to 14), Mohács (20 to 2), Baja (12 to 2) and Siklós (22 to 9) decreased, i.e. the nearby settlements have lost their network strength. Although Slovenian citizens in Budapest came from eight different settlements, the number of Slovenian citizens in Hungary is minimal. The examination of Hungarian migratory connections in relation to the settlements of all the neighbouring countries reveals that the central role of Budapest and Pest county in strengthening the connections is indisputable.

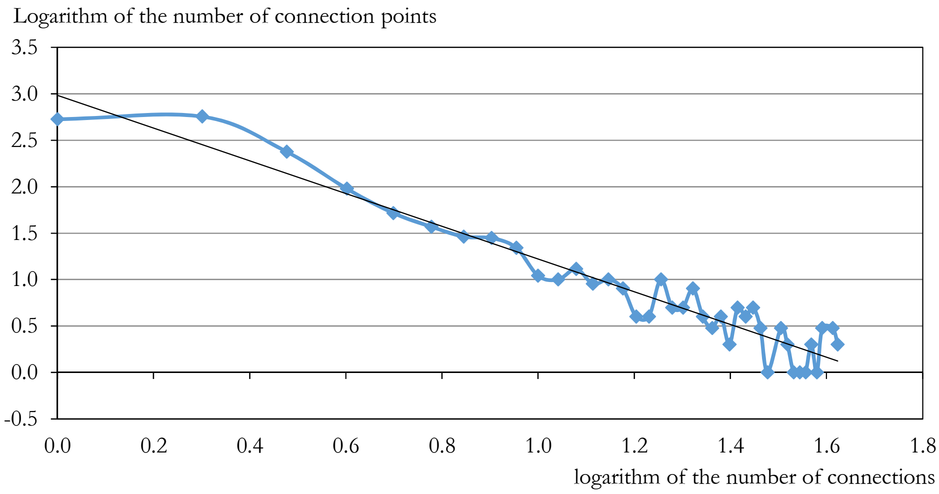
Overall, the regions of Central Hungary dynamically increase the number of migration settlements connections, while major border cities lose connectedness. People arrived in Budapest from 744 different settlements of neighbouring countries on 1 October 2001, whose number increased to 1059 by 2011. Dynamically developing settlements migration networks are characteristic along the axes of Dunakeszi, Fót, Göd, Vác, Szentendre, Pomáz, Budakalász, and Solymár; and Pécel, Maglód, Kerepes, and Gödöllő, with both set of regions starting from Budapest. Linear developments can be witnessed in the major sending countries, while on the western side of the capital a block-like growing structure can be seen with regard to two separate sets of regions: Üllő, Vecsés, Gyál, Monor, Pilis, Cegléd and Érd, Tárnok, Biatorbágy, Budaörs, Törökbálint, Budakeszi, and Szigetszentmiklós.

Examining the degrees (connections) of migration settlement networks, besides Budapest, the migratory links of settlements with neighbouring countries comprising Szeged, Debrecen, Pécs, Kecskemét, Győr, Nyíregyháza, Mosonmagyaróvár, Érd, Miskolc, Szigetszentmiklós, Budaörs, Tatabánya, Rajka, Cegléd, Székesfehérvár, Vecsés, Szentendre, Dunakeszi, Fót, and Vác, the large cities and the settlements close to Budapest, are significant. Connections of the agglomeration of Budapest and those of Szeged, Győr, and Kecskemét increased during the sampled ten years, while the stagnation or decline in the number of connections is typical for the other major towns. *Settlement links and dynamics imply the regional changes in the volume of future migrations. The degree is likely to decrease in cases when the sender areas have exhausted or the host areas have saturated and the previous migration flows have declined. In addition, the attractiveness of other areas might influence the choice of new migrants.*

After determining all Hungarian settlements by the degrees, i.e. the number of different settlements of neighbouring countries they had through the international migration, we examined the number of Hungarian settlements disposed of the given degree. The question is whether a random, a scale-free, or another topology will be explored. The results are shown in the case of Romania as of 1 October 2011. Having the given link according to the logarithm of the number of connections, the logarithm of the number of settlements was depicted. The result shows that the closer the graph is to the line, the truer it is that the migration settlement network has power law distribution.

Figure 8

Degree distribution of settlements affected in Romanian-Hungarian migration, 2011



Most Hungarian settlements have a few connections with Romanian ones through the migration (there are many points of small degrees), while there are a few settlements that have several links. Amount of Hungarian settlements having the given link by the number of connections according to a power law is ($R^2 \approx 0{,}86$). *It can be concluded that the Hungarian migration settlement connections from Romania towards Hungary show a scale-free topology. It is not only met in the case of Romania, but also for all the neighbouring countries, individually and together as well.* The R^2 values that measure the accuracy of fit are mentioned in the following table.

Table 3

Fit of degree distributions of migration settlements to the scale-free topology by citizenship (R2)

Citizenship	2001	2011
Romania	0.862	0.856
Serbia	0.846	0.687
Ukraine	0.827	0.788
Slovakia	0.722	0.905
Austria	0.714	0.700
Croatia	0.777	0.726
Slovenia	0.985	0.900
Total	*0.868*	*0.815*

Our results confirm that the network of Hungarian settlements affected by the migration clearly shows a scale-free topology (Kincses Á., 2012). Scale-free topology is the direct consequence of the expanding nature of real networks (Barabási A., 2008). A network exhibits scale-free topology because of its number of connections—settlements with more connections will be far more attractive for the migrants than those having a lesser degree *(1)*. According to the theory of migration networks (Sandu D., 2000; Kis, T., 2007), integration into a new environment is successfully carried out when it is aided by former family and friends. Following the 'beaten path' of emigration involves establishing contacts with previous immigrants, which also has a significant impact on the subsequent migration decisions (Rédei M., 2005). This is underpinned by the fact that one of the main 'tracks' of access to a certain country involves family reunification in most countries even today. Several newcomers also settle down with their kin and *acquaintances.*

Therefore, migration is more 'embedded' into a settlement with more links. In addition, greater potential migrant populations can be absorbed into a network by tapping the social and family circle of the migrants. A migrant is likely to choose a popular settlement with many connections, about which more information is available, rather than an unfamiliar settlement. Thus, the direction and magnitude of migration are affected the most by the appearance of migration networks in addition to income inequalities and migration distances. However, it is also visible that the settlements of greater degrees are generally larger in size, which may also be related to labour market positions. The size of target areas is cardinal because of the gravitational nature of migration. Relationship by distance and education indicates clearly that network mechanisms do not necessarily prevail.

Summary

In our study, the connections between the place of residence of birth and current dwelling of foreign citizens coming to Hungary from neighbouring countries were examined in detail. An examination of the connection system indicates the regions that have emerged stronger and those with weaker functional connections (Bell et al., 2003). In addition, depiction of migrations by socio-demographic groups offers more opportunities for understanding differentiated motivations. In general, we can state that the proportion of less-qualified, migrants with children, and unqualified migrants preferring short-distance migration is high, while for longer distances those with higher education become dominant. The potential impact area of migration grows in line with the education and qualification of migrants.

The Hungarian settlement connections of the migration from the neighbouring countries towards Hungary show a scale-free topology. A network exhibits scale-free topology because of its number of connections—settlements with more connections will be far more attractive for the migrants than those having a lesser degree. Greater potential migrant populations can be absorbed into a network by tapping the social and family circle of the migrants. A migrant is likely to choose a popular settlement with many connections, about which more information is available, rather than an unfamiliar settlement. The reasons why the Hungarian settlements are linked more strongly to the network are not random. In development, the build-up of networks causality is an important explanatory factor. If immigrants can find work, higher income, and better living conditions in a settlement, compared to other settlements, the higher standard of living creates a certain attraction; this establishes a pull-effect, and the established networks continue to reinforce this attraction. This implies that most Hungarian settlements have few links from neighbouring countries by migration (there are many points of small degree), while few settlements have many connections. The finding also means that, subsidiarity and regional strategies can play a decisive role, instead of the national migration strategy, in the management of the international migration process.

Acknowledgement

The Bolyai Research Scholarship has supported this study.

REFERENCES

ANDERSON, J. – O'DOWN, L. (1999): Borders, Border Regions and Territoriality: Contradictory Meaning, Changing Significance *Regional Studies* 33 (7): 593–604.

BARANYI, B. – BALCSÓK, I. (2004): *Határ menti együttműködés és foglalkoztatás – Kelet-magyarországi helyzetkép* In: Műhelytanulmányok 2004/20, MTA Közgazdaságtudományi Intézet, Budapest.

BOLLOBÁS, B. – ERDŐS, P. (1976): Cliques in Random Graphs *Math. Proc. Cambridge Phil. Soc.* 80 (3): 419–427.

BUCHANAN, M. (2003): *Nexus, avagy kicsi a világ* Typotex, Budapest.

BARABÁSI-ALBERT, L. (2008): *Behálózva* Helikon Kiadó, Budapest.

BELL, M. – BLAKE, M. – BOYLE, P. – DUKE-WILLIAMS, O. – REES, P. – STILLWELL, J. – 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): 435–465.

GÜLLÜOĞLU, C. (2013) FlowMapper v0.4 https://plugins.qgis.org/plugins/FlowMapper/version/0.4/

DARÓCZI, G. – BÁLINT, L. (2015): Áramlástérkép R-ben *Területi Statisztika* 55 (4): 339–355.

ERDŐS, P. – RÉNYI, A. (1959): On Random Graphs. I. *Publicationes Mathematicae* 6: 290–297.

ERDŐS, P. – RÉNYI, A. (1960): On The Evolution of Random Graphs *Magyar Tudományos Akadémia Mat. Kutató Int. Közl.* 5 (1-2): 17–61.

ESTÉLYI, K. – KESZEGH, B. – KOVÁCS, P. – MIKÓCZY, I. (2006): Munkaerőmozgás a szlovák–magyar határ mentén *Fórum Társadalomtudományi Kutató Szemle* 8 (1): 25–48.

FERCSIK, R. (2008): Szülőföldről a hazába – és vissza? In: Szarka, L. – Kötél, E. (szerk.): *Határhelyzetek. Külhoni magyar egyetemisták peregrinus stratégiái a 21. század elején* pp. 124–138., Balassi Intézet Márton Áron Szakkollégium, Budapest.

HANSEN, N. (1977): Border Regions: a Critique of Spatial Theory and an European Case Studies *Annals of Regional Science* 11 (1): 1–12.

ILLÉS, S. – KINCSES, Á. (2009): Cirkuláció és migráció Magyarország nemzetközi vándormozgalmában *Statisztikai Szemle* 87 (7–8): 729–747.

KINCSES, Á. (2012): *A Kárpát-medence ezredforduló utáni migrációs hálózatának vizsgálata* Központi Statisztikai Hivatal, Budapest.

KIS, T. (2007): Demográfiai modellek és a migráció *Régió* 18 (2): 160–189.

MELEGH, A. L. (2011): Globalizáció és migráció Magyarországon *Educatio* 20 (2): 206–219.

RÉDEI, M. (2005): A nemzetközi vándorlás folyamatának irányítása *Statisztikai Szemle* 83 (7): 662–680.

RÉDEI, M. (2007): *Mozgásban a világ, a nemzetközi migráció földrajza* Eötvös Kiadó, Budapest.

RÉDEI, M. (2009): Foreigners in Budapest *Hungarian Statistical Review* 13 (87): 31–49.

SOLTÉSZ, B. – ERŐSS, Á. – KARÁCSONYI, D. – KINCSES, Á. (2014): Hungary: cross border migration in a fragmented ethnic space In: ERŐSS, Á. – KARÁCSONYI, D. (eds): Discovering migration between Visegrad countries and Eastern Partners pp. 72–87., Geographical Institute, Research Centre for Astronomy and Earth Sciences of the Hungarian Academy of Sciences, Budapest.

SANDU, D. (2000): Migraţia transnaţională a românilor din perspectiva unui recensământ comunitar *Sociologia Romaneasca* 3–4: 5–52.

SJAASTAD, L. A. (1962): The Costs and returns of Human Migration *The Journak of Political Economy* 70 (5): 80–93.

TOBLER, W. (2003): Movement Mapping (unpublished paper) http://csiss.ncgia.ucsb.edu/clearinghouse/FlowMapper/MovementMapping.pdf (letöltve: 2015. január)

TOBLER, W. (2012): Looking at Some Data from Isard's "Methods of Regional Analysis" *Regional Statistics* 2: 3–12.

TÓTH, G. (2013): *Bevezetés a területi elemzések módszertanába* Miskolci Egyetemi Kiadó, Miskolc.

VAN GEENHUIZEN, M. – RATTI, R (2001): *Gaining Advantage from Open Borders. An active Space for Regional Development* Ashgate, Aldershot.

Estimates of Net Capital Stock and Consumption of Fixed Capital for Australian States and Territories, 1990–2013*

Serguei Mikhailitchenko
Holmes Institute,
Brisbane, Australia
E-mail:
smikhailitchenko@holmes.edu.au

Being an important input for many economic models and being widely used in economic decision making by the federal and regional government, capital stock data are not readily available at the sub-national level for most countries, including Australia. The study closely follows the methodology of the Australian Bureau of Statistics and presents a complete set of capital stock data for the states and territories of Australia for the period 1990–2013. The robustness of the method is assessed by comparing the aggregated data by type of asset estimates for the states, with the capital stock data by type of asset for Australia as a whole published by the Australian Bureau of Statistics.

Keywords:
regional capital stock,
perpetual inventory method
Australia

Introduction

Capital stock data is highly important for researchers in the areas of productivity and economic growth analysis. It allows a deeper insight into labour productivity and economic growth because, with capital stock data available, one can conduct multifactor productivity analysis to account for technological and institutional innovations as well as perform a detailed labour productivity analysis by decomposing output per unit of labour into output per unit of capital and capital-to-labour ratios. In the absence of official capital stock data for Australian states, researchers have to rely on the series developed by experimental studies. However, none of the previous studies has fully utilised the process used by the Australian Bureau of Statistics (ABS) for deriving the capital stock series and none of them has derived consumption of fixed capital data (COFC), which might be more useful for conducting productivity analysis because COFC, to some extent, accounts for the stock composition (OECD 2009: 61).

* A version of this paper was presented at the 2015 Australian Conference of Economists, Brisbane, 7–10 July 2015.

The study reports the methodology employed for the estimation of stocks and flows of capital and presents a complete set of capital stock data for the six states and two territories of Australia for the period 1990–2013. The author departs from the simultaneous exit assumptions implicitly used by the previous studies, such as Levtchenkova and Petchey (2000), Louca (2003), and Mikhailitchenko, Nguyen and Smith (2005), and closely follows the Perpetual Inventory Model (PIM) process (ABS 5216.0, 2012). One of the advantages of the approach used by the author is that, before estimating the series at a sub-national level, he looks at how the model works on highly aggregated investment data for Australia as a whole and compares his estimates with the NCS and COFC data published by the ABS. This permits the author, at least in the first approximation, to assess the quality of the estimates[1].

Another advantage of this study is that it does not require any initial allocation of the national capital stock of each type among the states and territories at the beginning of the estimation period or between the public and the private sectors in each state. This is because the PIM process is applied directly to the data published by ABS Gross Fixed Capital Formation (GFCF). Without the need for two rounds of stock allocation, the method effectively removes one major source of statistical errors, which might be considered as an advantage over the methodology adopted by Mikhailitchenko, Nguyen, and Smith (2005).

Further, the author used age-and-type specific depreciation rates for all the series. In Mikhailitchenko et al. (2005), for example, the assets were depreciated using the national asset-specific depreciation rates, calculated by dividing COFC by NCS in relevant periods. This implies that all assets of a particular type, accumulated in the previous periods, are discounted at the same rate irrespective of the actual age of the existing assets in the stocks, which represents yet another source of error, the size of which positively correlates with the length of the service life of an asset. In the absence of other alternatives, utilising research conducted by Walters and Duppelsman (1985), ABS adopted a similar approach to discount pre-1948 NCS of Dwellings (DW). However, the errors in the latter can be expected to be of much smaller size than in Mikhailitchenko et al. (2005) because the relative weight of the pre-1949 items in the stock is very small due to the age of these assets and the small relative value of pre-1946 GFCF when compared to later years.

Finally, unlike previous research in the regional capital stock, this study applies not only the Age-efficiency and Age-price functions to the investment inflows but also the Survival function to account for stock retiring during the accounting period. In all the previous studies on NCS for Australian regions known to the author, the estimates were based on an assumption that all assets retire immediately after they reach the age equal to the mean useful service life, although the authors have not

[1] The estimates of NCS and COFC yielded satisfactory results. The results of the method pre-testing are not presented due to space and word count limits, but are available on request.

stated this explicitly. For most stocks, the author used the Winfrey's **S3** function, except for alternations and additions, where the **S0** survival function was applied as specified in Winfrey (1938) and ABS (5216.0, 2012).

The study is the first to the author's knowledge that presents the estimates of COFC for Australian states and territories, which makes the whole set of capital stock data available for researchers. COFC might have an advantage over NCS in some research areas, such as the productivity analysis, because, being a part of the capital services flow, it accounts better for this variable in any given year and can be used to impute capital services series at user cost. Certainly, there is a significant difference in the flow of capital services between the value of a dollar invested in DW and the value of a dollar invested in Computer Software (CS) due to significant differences in the PIM parameters between these two types of assets.

Although the study closely follows the ABS PIM process, the aggregate estimates of total stocks and flows might differ from the data published by ABS due to the differences in the GFCF data available for the states and data available for Australia as a whole, in terms of the level of aggregation at which the PIM process is applied and the length of the available series. For example, ABS applied PIM separately for detached houses, residential units, and alterations and additions. In this study, the author was able to disaggregate alternations and additions from new dwellings, but not new dwellings by type. For the same reason, several stocks, such as cultivated biological products (CBR), intellectual property products (INTEL), and public (PUB) were estimated using aggregate GFCF data, despite the significant differences in PIM parameters between the stocks.

The rest of the article is organised as follows. Section II discusses the models applied for estimating capital stock series, while the discussion of data used for estimation is presented in Section III. Section IV gives the derived estimates of NCS and COFC for Australian states and territories and compares them with the data published by the Australian Bureau of Statistics. Section V provides the concluding comments.

The model

The study utilises the Perpetual Inventory Model (PIM) to estimate the NCS series, as Hulten (1990) describes it and ABS 5216.0 (2000).

$$K_t^{ij} = K_{t-1}^{ij} + I_t^{ij} - D_t^{ij} , \quad (1)$$

where, K_t^{ij} is NCS of asset type i in region j in period t; K_{t-1}^{ij} is NCS in period t−1; it accounts for accumulated NCS less retirement during the period; I_t^{ij} is Gross Fixed Capital Formation (GFCF) (or investment) of asset type i in region j in period t; D_t^{ij} is Consumption of Fixed Capital (CFC) of asset type **i** in region **j** in period **t**.

In this model, the NCS in the current period is estimated by adding GFCF to the NCS in the previous period less retirement and less depreciation in the current period.

The investment made in the current period is added to the capital stock accumulated in the previous periods with zero depreciation. Depreciation applies only to the assets added in previous periods and at a vintage-specific rate. If equation **(1)** is used in default, as was the case, for example, by Mikhailitchenko, Nguyen and Smith (2005), we would need to initiate the series and derive $\boldsymbol{K_{t-1}^{ij}}$ by going through at least two rounds of allocation of NCS of each type; first, between the states and territories, and, second, between the public and the private sectors for each state and for each type of asset. This inevitably results in an allocation error that does not diminish rapidly, especially for assets with long service lives, such as dwellings and non-dwelling construction. As an alternative, we can re-write $\boldsymbol{K_t^{ij}}$ as a sum of GFCF discounted by the age-specific depreciation rates, $\boldsymbol{\delta_t}$ in each year of service and by the age-specific retirement rates as per Winfrey's (1938) survival function.

$$K_t^{ij} = \sum\nolimits_{\tau=0}^{T} I_{t-\tau}^{ij}(1-\delta_\tau^{ij})(1-\theta_\tau^{ij}), \quad (2)$$

where, $\boldsymbol{I_{t-\tau}^{ij}}$ is the real value of investment of type **i** in a region **j** with the remaining service life of **t-τ**; $\boldsymbol{\delta_\tau^{ij}}$ is the accumulated loss of efficiency of an asset when it reaches the age τ; $\boldsymbol{\theta_\tau^{ij}}$ is the accumulated value of loss of assets due to their retirement from the stock.

As can be seen from the equation **(2)**, PIM cumulates past investment flows, discounting them for loss in efficiency and outflows of assets from the stock due to retirement. A similar approach has been described in ABS (1351.0055.004, 2005: 22), with the only difference being, they applied price changes to the assets to obtain the net capital stock estimates in real terms. As these price deflators have already been applied to the investment data to produce GFCF series in chain volume measures, there is no need for the price factor in this place.

Apart from the investment data, the PIM, in the form of equation **(2),** requires parameters of the age-efficiency function for deriving vintage- and asset-specific depreciation rates and the survival function to account for the retired assets. The hyperbolic age-efficiency function is used to derive depreciation rates due to a decline in efficiency of an item, as shown in equation **(3)** below (ABS 5216.0, 2000).

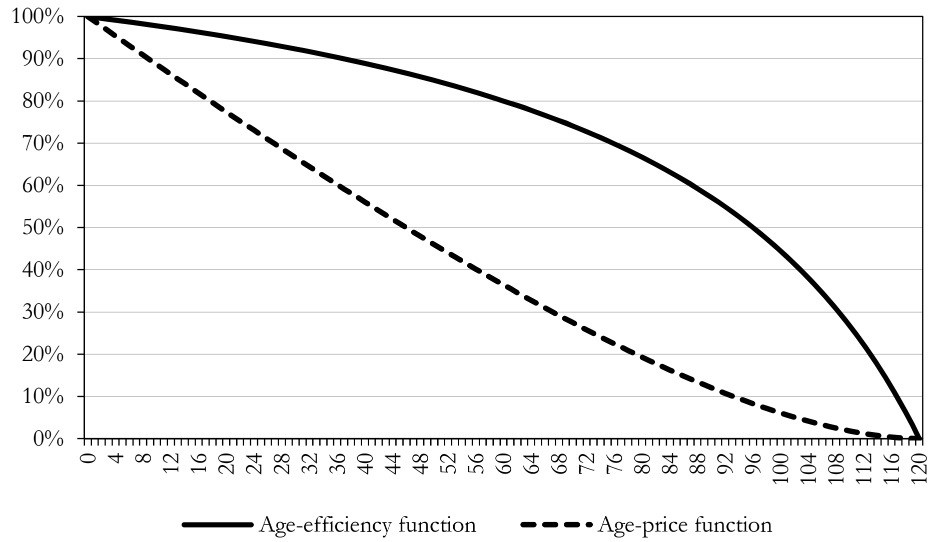
$$E_t = \frac{M - A_t}{M - bA_t}, \quad (3)$$

where, $\boldsymbol{E_t}$ is the efficiency of a capital asset at a point of time t, $\boldsymbol{M}$ is the service life of an asset according to the Winfrey (1938) distribution, $\boldsymbol{A_t}$ is the current age of an asset (equals t – τ in equation (**2**)), **b** is the asset-specific efficiency reduction parameter.

Figure 1 below presents derived age-efficiency and age-price functions for dwellings as an example of the PIM processes used in this study. The calculated efficiency of an asset at each year between 0 and **T** gives the proportion of the initial potential of an asset to produce capital services, which is then used to construct the productive capacity function of an asset over its service life. The time-variant depreciation rate is then calculated as a percentage change in productive capacity in consecutive years.

Figure 1

Age-price and age-efficiency functions, dwellings



Source: Author's calculations.

Winfrey's survival functions are described as equation **(4)** below.

$$F_t = F_0(1 - \frac{T^2}{a^2})m, \qquad (4)^2$$

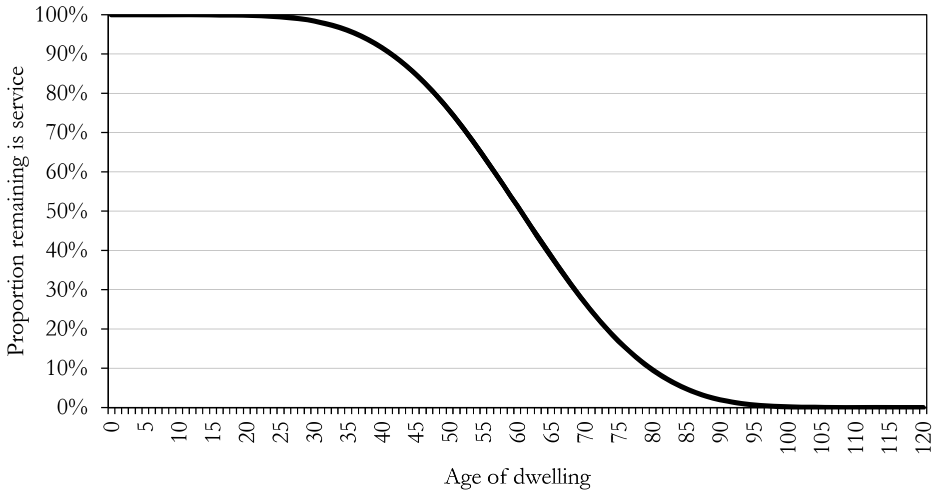
where, $\boldsymbol{F_T}$ is the proportion of the cohort retiring at the time T, $\boldsymbol{F_0}$ is the proportion of the cohort retiring at the mean asset service life, **a** the parameter that determines the time periods (e.g. a = 10 for deciles), **m** is the parameter that determines the flatness of the distribution.

Figure 2 presents the Survival function for new DW used in this study with the age of dwelling appearing on horizontal axis and the proportion of stock remaining in service on the vertical axis.

[2] For the sake of practicality, we used Winfrey's survival functions in a cumulative form.

Figure 2

Winfrey's S3 survival function, dwellings



Source: Author's calculations.

Estimates of Consumption of Fixed Capital come as a by-product of the model. Equation **(5)** illustrates the process of the estimation of COFC.

$$D_t^{ij} = \sum_{A_i=0}^{T} I_{t-\tau}^{ij}(\delta_\tau^{ij} - \delta_{\tau-1}^{ij}), \quad (5)$$

This process specified by equation **(5)** uses the same parameters as per equation **(2)** and estimates COFC per asset as the change in the value of stock between two latest consecutive years.

Section III presents a discussion of GFCF data available as an input into the PIM process to derive NCS and COFC series at sub-national level.

Data description and availability

In most developed countries, data for physical capital are available at a national level. The ABS has been publishing annual NCS, GFCF and COFC data in chain-volume measures (CVM) and current prices for Australia as a whole since 1960 (ABS 5204.0), with the data disaggregated by type of asset (Table 56), by institutional sector (Table 57), and by industry (Table 58). However, with only a few exceptions, the same level of coverage does not apply to the data at sub-national level. In Australia, only GFCF by type of asset series are available for the states and territories. Official NCS and COFC estimates at sub-national level exist in only a few countries and are not available for Australian states. To the author's knowledge, only Canada has a complete set of data for the capital stock of its provinces (Statistics Canada 2012); in addition, India has capital stock estimates for some, but not for all its states.

The published capital stock data by type of asset for Australia as a whole is of particular interest for this study because these data are used for preliminary testing and calibrating the models that are subsequently applied for deriving the corresponding regional series. GFCF data (ABS 5204.0) for nine types of assets, Dwellings (DW), Non-dwelling construction (NDC), Machinery and equipment (ME), Weapons (W), Cultivated biological resources (CBR), Research and development (R&D), Mineral and petroleum exploration (PE), Computer software (CS), and Artistic originals (AO)), will allow for testing if equation **(2)** yields results that are comparable with the NCS official data from the same ABS publication.

It is challenging to test the model for the Ownership Transfer Costs (OTC) due to some discrepancies in the ABS manual (ABS 5216.0, page 376). While GFCF for this stock is supposed to depreciate over the period, as per the owner's plan to hold the asset, ABS gives the mean service life for OTC as equal to zero, which is mathematically impossible. In addition, the age-efficiency reduction parameter **b** is not specified in the ABS National Accounts Manual (5216.0, 2012), which makes equation (**2**) unidentified. Therefore, the maximum life of assets in OTC will be selected by trial and error for the length of the service life and the age-efficiency parameter under an assumption that most OTCs are for dwellings and non-dwelling construction assets.

The ABS has been publishing annual private GFCF data for the Australian states and territories in chain volume measures and current prices (ABS 5220.0) since 1990 and quarterly GFCF data (ABS 5206.0) since September 1985. The private sector data are disaggregated by type of asset, while the public sector series are disaggregated by the level of government (national, state, and local) and public corporations (commonwealth, state, and local). The same sources produce similar series for Australia as a whole with annual data starting from 1960 and quarterly data starting from September 1959.

Despite the fact that the annualised GFCF data now cover the periods of up to 26 years, the series still are too short for deriving NCS for assets with long service lives. For example, roads have a mean life span of up to 50 years, water and sewerage infrastructure can be used for up to 71 years, and private brick homes depreciated over 88 years, on average. This implies that some of the series must be over 150 years long and, for constructing net capital stocks series by using equation **(2),** will inevitably require some proxy inputs for the years where data are not available.

This study extended the work of Mikhailitchenko et al. (2005) and Mikhailitchenko (2008) in utilising proxy data by using other available series for DW. Using proxies should not be considered a significant deviation from the ABS methodology. For example, according to ABS 5216.0 (2012), an application of the normal S3 curve for survival function (Winfrey, 1938) for estimating NCS of houses implies that we need construction series longer than 150 previous years, while the first national account statistics, including building activity, became available only in the late 1940s (ABS 5216.0, 2000). This point has not been made explicitly clear in the manual; ABS refers to a range of sources (5216.0, page 370, clause 14.55), which implies that the value of stock added in early years is less reliable. Table 1 below summarises the GFCF data available for Australia as a whole and for its states and territories.

Table 1

Availability of GFCF data

Type of asset	Australia as a whole	Australian states and territories	Proxies available
Dwellings	ABS 5204.0 Table 2: annual, private, subdivided into dwellings, and alternations and additions, 1960–2013 ABS 5204.0 Table 56: annual, total of all sectors, total subtypes, 1960–2013 ABS 5220.0: annual, total all sub types, total all sectors, 1960–2013	ABS 5220, annual, private, dwellings, and alterations and additions, 1990–2013 ABS 5206, quarterly, private, dwellings, and alterations and additions, Sep 1985–current	ABS 8752, Number and value of dwellings completed by state, quarterly, private sector, and total, Sep 1969–current ABS 1300 Year books for the Australian states and territories, annual, number of new dwellings constructed, total all sectors, 1947–current (except SA, ACT, and NT) ABS 3105, Table 1.11: population, annual, 1860–1994
Ownership transfer costs	ABS 5204.0, annual, total, total, 1960–2013	ABS 5220, annual, private, total, 1990–2013 ABS 5206, quarterly, private, total all types, Sep 1985–current	ABS 3105.0, Table 1.11: population, annual, 1860–1994
Non-dwelling construction	ABS 5204.0 Table 2: annual, private, sub-divided into new buildings (1977–2013), new engineering construction (1977–2013), and purchase of second hand assets (1960–2013) and total all sub types (1960–2013) ABS 5204.0 Table 56: annual, total all sectors, total sub types, 1960–2013	ABS 5206, annual, private, sub-divided into new buildings, new engineering construction, and purchase of second hand assets Sep 1985–current ABS 5220, annual, private, sub-divided into new buildings, new engineering construction, and purchase of second hand assets 1990–2013	ABS 3105, Table 1.11: population, annual, 1860–1994

(Table countinues on the next page.)

(Continued.)

Type of asset	Australia as a whole	Australian states and territories	Proxies available
Machinery and equipment	ABS 5204.0 Table 2: annual, private, sub-divided into new assets (1960–2013), net purchase of second hand items (1971–2013), and total (1960–2013) ABS 5204.0 Table 56: annual, total all sectors, total sub types, 1960–2013	ABS 5206, quarterly, private, subdivided into new assets, net purchase of second hand assets, and total, Sep 1985–current ABS 5220, annual, private, sub-divided into new assets, net purchase of second hand assets, and total, 1990–2013	ABS 3105, Table 1.11: population, annual, 1860–1994
Weapons systems	ABS 5204.0, annual, public, total all types, 1960–2013	Not available	Not available
Cultivated biological resources	ABS 5204.0, annual, total, total, 1960–2013	ABS 5220, annual, private, total, 1990–2013 ABS 5206, quarterly, private, total all types, Sep 1985–current	No need for proxies
Intellectual property Products	ABS 5204.0, Table 2: annual, private, disaggregated by RD, MPE, CS, and AO, 1986–2013; total private by type, 1960–2013 ABS 5204.0, Table 56, annual, total of all sectors, disaggregated by RD (1960–2013), MPE (1960–2013), CS (1964–2013) and AO (1972–2013)	ABS 5220, annual, private, total all types, 1990–2013 ABS 5206, quarterly, private, total all types, Sep 1985–current	ABS 3105, Table 1.11: population, annual, 1860–1994
Public	ABS 5204.0, Table 2: annual, public, disaggregated by public corporations (Commonwealth; state and local; total public corporations); General government (national (defence; non-defence; and total) and state and local; total general government), total public, 1960–2013	ABS 5220, annual, public, disaggregated into Public Corporations and General Government, 1990–2013 ABS 5206, quarterly, public, disaggregated into Public Corporations and General Government (National and State and Local; total General Government), and total public, 1990–2013	ABS 3105, Table 1.11: population, annual, 1860–1994

As can be seen from Table 1, GFCF data are available for Australia, which allows for pre-testing the models as per equations **(2)** and **(5)** before applying them for estimation of NCS and COFC at the sub-national level. The issues associated with the level of aggregation of GFCF data for the states are also visible from Table 1. At the state level, GFCF appears either as private investment flows disaggregated by asset type or as data series for public GFCF, disaggregated into public corporations and general government without providing details of the stock composition.

For DW, the author is using the value of residential work done (ABS 8752.0, 2013 and ABS 1300, various years) as a proxy for GFCF for the period 1948–1985. The author finds that GFCF data exhibit higher volatility than population growth numbers. In addition, given the slow depreciation of DW, the size of the errors will be larger if the population is used as a proxy for GFCF in years when the residential work data are available. The percentage changes in the value of building work done has been selected on the basis that the old series for the value of residential work done in current prices have a significantly steeper slope that chained GFCF series. However, the annual percentage changes produce a very good fit for the proxies into the real GFCF data. For years prior to 1948, the author chose to use population as a proxy.[3]

Further, difficulties are related to the Intellectual Property Products (Intel) series. At the state level, these data are published as a total for all type of assets within a stock, while Australian data are disaggregated into the following four subcomponents: Research and Development (R&D), Mineral and Petroleum Exploration (MPE), Computer Software (CS), and Artistic Originals (AO). This represents a challenge due to the significant differences in the PIM parameters for the sub-components in Intel. For example, the mean service lives in MPE is 35 years, while AO are depreciated within the first 2 years, on average. In this study, the author used the national weighted average of each sub-type in Intel to determine the PIM parameters.

Given the issues discussed above, the estimates of the derived series should be treated with caution, especially the INTEL series. Now, the published and proxy data will be used in Section 4 to evaluate the performance of the model for Australia as a whole.

Estimating Net Capital Stock and CFC for Australian states and territories

In this section, the author applies equations (2) and (5) to derive NCS and COFC series by type of asset for Australian states and territories and reports the estimates of NCS and COFC by type of asset for the Australian states and territories. The PIM processes adopted in this study applies the Age-Efficiency and the Winfrey's survival

[3] Mikhailitchenko (2008) used a similar approach; graphs showing the proxy series are available on request.

functions, simultaneously, as per equations (2) and (5). Table 2 presents the PIM parameters used for estimating capital stock data in this study.

Table 2

Parameters of Perpetual Inventory Model (PIM)

	Winfrey Survival function	Mean service life	Efficiency reduction parameter
Dwellings	S3	60	0.750
Alternations and additions	S0	35	0.500
Ownership transfer costs	S3	13	0.500
Non-dwelling construction	S3	35	0.750
Machinery and equipment	S3	12	0.500
Cultivated biological products	S3	6	0.500
Intellectual property products	S3	13	0.135
Public assets	S3	40	0.500

Source: Author's estimates.

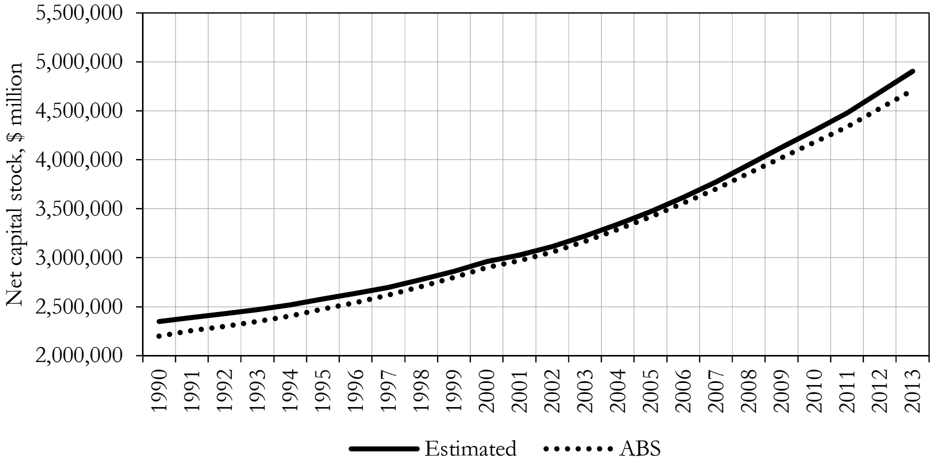
Similarly, the public investment data is given by the ABS at a highly aggregated level, which implies the same process of guessing the PIM parameters. We assume that the efficiency reduction parameter and the mean service life for the Public assets are 0.75 and 30, respectively. The New Dwellings were assumed to have an average service life of 60 years, while the Alternations and Additions and Non-Dwelling Construction assets remain in service for 35 years, on average. The efficiency reduction parameters for these assets are assumed to be equal to 0.75 for dwellings and non-dwelling construction and 0.5 for the Alterations and Additions.

For Machinery and Equipment, the mean service life and efficiency reduction parameters were 12 and 0.5, respectively. The Cultivated Biological Resources serve for 6 years, on average, and follow the age-efficiency function with b equal to 0.5. For Intellectual Property Products, the author used 13 years and 0.135 for the mean service life and the efficiency reduction parameters, respectively.

The estimated capital stock data for Australian states and territories by asset type are provided in Tables A.1 to A.8 in Appendix. Figure 3 provides the comparison between the aggregate estimates of NCS and COFC by ABS and this study. The author presents the data 'as is' without adjusting the data at the base year of 2012 to illustrate the degree of reliability. As can be seen from this figure, the error term is largest in 1990 (6.7%), which gradually diminishes until 2008 (2.2%) and then increases again until 2013 (4.1%) (The estimated capital stock data for Australian states and territories by asset type are provided in Tables A.1 to A.8 in Appendix on the Regional Statistics's website.)

Figure 3

Net Capital Stock, total, Australia 1990–2013

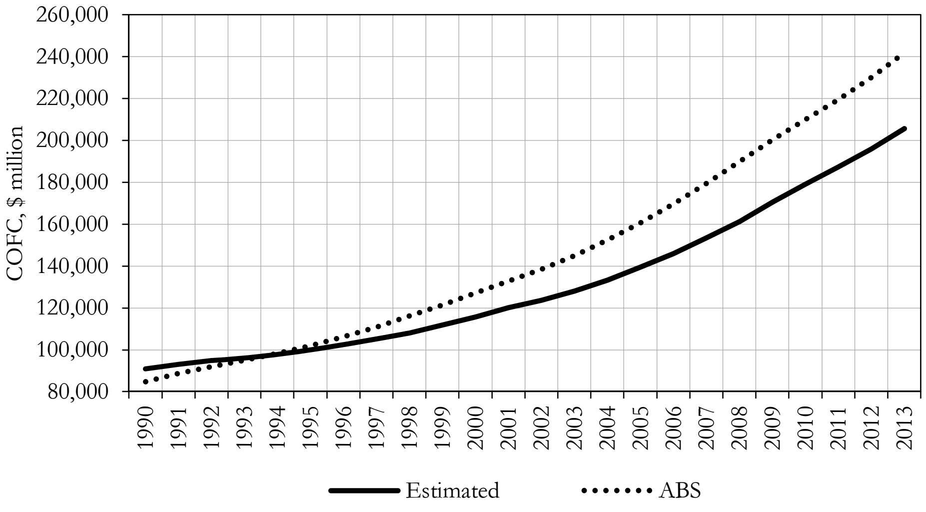


Source: Author's estimates.

Figure 4 presents the estimates of total COFC for Australia as a sum of COFC of all types for all states. As can be seen from the figure, the estimates of COFC are not as accurate as in the case of NCS. This study overestimates the ABS series by 13% in 1990 and underestimates by 17% in 2013, which indicates to the author that some further research effort is required to correct the COFC series for errors wherever it is possible.

Figure 4

Consumption of Fixed Capital, total, Australia 1990–2013

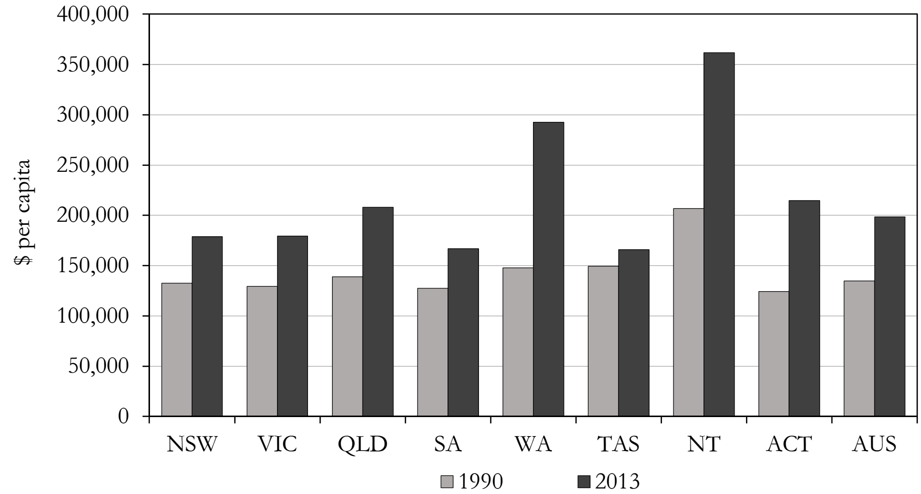


Source: Author's estimates.

Figure 5 presents regional Net Capital Stock data in 1990 and 2013 in terms of NCS per capita. As can be seen from the graph, the amount of physical capital per head of the population has grown significantly between 1990 and 2013 for all states and Australia as a whole.

Figure 5

Net Capital Stock per capita, Australia and states, 1990 and 2013



Source: Author's estimates.

The highest level of capital per capita was in the mining states of Queensland, Western Australia, and Northern Territory in 1990 and remains such in 2013. The lowest per capita NCS is in traditionally underperforming states of South Australia and Tasmania. Further research into the regional economic performance in Australia can be conducted by utilising the developed series by type of asset. For example, the value of public NCS per capita in Tasmania has not changed between 1990 and 2013, which might indicate that a lack of public infrastructure is hindering the faster economic growth of this state.

Although there are differences between the NCS values in this study and the corresponding estimates by ABS, it should be kept in mind that chained series do not necessarily sum up, with the exception of the base year, which is 2012 in this case[4]. The largest percentage error for NCS is −3.6%, which might be considered as non-significant given the issues with the lack of investment data at the state level.

[4] ABS 5216.0 (2012), page 70.

Conclusion

The study provided the series for the Net Capital Stock and CFC for Australian states and territories for the period 1990–2013. The series were derived through the PIM process that closely followed procedures used by the ABS for deriving NCS and COFC series for Australia as a whole. There are several advantages of this study compared to the previous attempts, with the first one being the removal of initial allocation of the capital stock among the states and the sectors within the states. In addition, this study extends the previous work by applying the survival function to the assets and effectively removing another error term present in all previous estimates at sub-national level for Australia.

There are several obvious caveats present in this study. Despite following the ABS methodology as closely as possible, a number of PIM parameters have been selected based on educated guess for some of the GFCF series, such as Dwellings, Intellectual Property Products, and the Public sector, because the series are highly aggregated compared to the data used by ABS for Australia as a whole. The significant differences in the PIM parameters for different types of assets might cause an error that is impossible to assess in this study in the absence of micro data.

Although this type of error is also related to the aggregation level of the data, in the absence of micro GFCF data for Machinery and Equipment, it is impossible to determine the PIM parameters in any reliable way. This is because of the changing composition of this stock due to the growth in proportional holdings of assets with relatively short service lives, such as computers and peripherals, and a gradual decrease in the proportion of assets, such as cars and machinery.

The NCS and COFC estimates for Australian states and territories provided by this study facilitates spatial analysis of the capital availability in Australia as well as utilisation of this data for other types of analyses, such as the multifactor productivity analysis. The results should be treated with caution due to the sensitivity of the estimates to the selected parameters of the Perpetual Inventory Model.

REFERENCES

AUSTRALIAN BUREAU OF STATISTICS (various years): *Australian System of National Accounts* Cat. no. 5204.0, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (2013): *Australian National Account: National Income, Expenditure and Product*, Cat. no. 5206.0, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (2012): *Australian System of National Accounts: Concepts, Sources and Methods* 3rd ed., Cat. no. 5216.0, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (various years): *Australian Year Book* Cat. no. 1301.0, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (various years): *New South Wales Year Book* Cat. no. 1300.1, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (various years): *Victorian Year Book*, Cat. no. 1300.2, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (various years): *Queensland Year Book* Cat. no. 1301.3, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (various years): *South Australian Year Book* Cat. no. 1301.4, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (various years): *Western Australian Year Book* Cat. no. 1300.5, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (various years): *Tasmanian Year Book* Cat. no. 1301.6, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (various years): *Northern Territory at a Glance* Cat. no. 1304.7, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (various years): *Australian Capital Territory Business Indicators* Cat. no. 1303.8, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (2005): *Estimating Industry-Level Multifactor Productivity for the Market-Sector Industries in Australia: Methods and Experimental Results* Australia, Cat. No. 1351.0055.004, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (2006): *Population Projections*, Australia, Cat. no. 3202.0, ABS, Canberra.

AUSTRALIAN BUREAU OF STATISTICS (various years): *Building Activity*, Australia, Cat. no. 8752.0, ABS, Canberra.

Australian Bureau of Statistics (various years), *Australian National Accounts: State Accounts* Cat. no. 5220.0, ABS, Canberra.

HULTEN, C. (1990): The Measurement of Capital In: BERNDT, E. R.–TRIPLETT, J. (eds.) *Fifty Years of Economic Measurement* pp. 119–152., NBER, Washington, D.C.

LEVTCHENKOVA, S.–PETCHEY, J. (2000): Regional Capital Stock Data for Australia *Australian Economic Review* 33 (2): 193–197.

LOUCA, J. (2003): Multifactor Productivity and Innovation in Australia and Its States In: WILLIAMS, C.–SMITH, C.–DRACA, M. (eds) *Productivity and Regional Economic Performance in Australia* pp. 99–138., Queensland Government, Brisbane.

MIKHAILITCHENKO, S.–NGUYEN, D. T.–SMITH, C. (2005): Estimates of Capital Stocks for the States and Territories of Australia, 1985-2004 *Economic Analysis and Policy* 35 (1/2): 1–22.

MIKHAILITCHENKO, S. (2008): *Australian Housing Market: Price Dynamics and Capital Stock Growth* PhD thesis, Griffith University, Brisbane.

OECD (2009): *Measuring Capital. OECD Manual.* 2nd edn., OECD, Brussels.

STATISTICS CANADA (2012): *Investment, capital stock and capital services of physical assets, by North American Industry Classification System (NAICS), annual (Dollars)* (downloaded: 23 August 2012)
http://www5.statcan.gc.ca/cansim/a34? lang=eng&id=3830025&search TypeByValue=1&mode=tableSummary&p2=35

WALTERS, R.–DIPPELSMAN, R. (1985): *Estimates of Depreciation and Capital Stock* Australian Bureau of Statistics, Occasional Paper No. 1985/3, ABS, Canberra.

WINFREY R. (1938): Statistical *Analysis of Industrial Property Retirements* Iowa State College of Agricultural and Mechanic Arts, Ames, Iowa.

Characteristics of transit tourism in Hungary with a focus on expenditure

Áron Kincses
Hungarian Central Statistical Office,
Budapest, Hungary
E-mail: aron.kincses@ksh.hu

Géza Tóth
Hungarian Central Statistical Office,
Budapest, Hungary
E-mail: geza.toth@ksh.hu

Mihály Tömöri
University of Nyíregyháza,
Nyíregyháza, Hungary
E-mail: tomori.mihaly@nye.hu

Gábor Michalkó
HAS RCAES,
Corvinus University of Budapest, Hungary
E-mail: michalko.gabor@csfk.mta.hu

Keywords:
transit tourism,
tourism expenditure,
Hungary

This study introduces the behaviour of participants in transit tourism in Hungary with a focus on their expenditure. With the help of multivariable mathematical-statistical methods, the motivational background and the spending characteristics of foreigners visiting Hungary between 2009 and 2013 are explored; in addition, the factors influencing expenditure, according to nationality, are investigated.

According to our investigations, people in transit, whose spending is continuously increasing, make up a significant share of the expenditure of foreigners arriving in Hungary. Typical types of spending during transit are fuel purchases and dining at restaurants. Among transit visitors to Hungary, Romanian, Serbian (including Montenegro and Kosovo), and Bulgarian nationals have the highest share. While the number and expenditure of transit visitors slightly increased during the examined period, the per capita spending decreased. The results of the study show that this is due to the changes in the composition of the countries involved.

Changes in transit tourism expenditures are largely determined by nationality. The most important conclusion of our research is that the most significant characteristics of transit depend on general European trends (labour market conditions, tourism supply, etc.) and conditions (visa requirements, transport infrastructure, accommodation along transit routes, among others) provided for transit visitors by Hungary.

Regional Statistics, Vol 6, No 2. 2016: 129–148; DOI: 10.15196/RS06207

Introduction

Tourism is one of the driving forces of world economy. The 4% annual average growth rate of international tourist arrivals and the similar growth rate of transit visitors' expenditure positively influence the GDP, employment, and export figures of the affected countries (UNWTO 2015). In parallel with the increase in the number of visitors and their expenditure, the spectrum of the interpretation of international tourism is becoming wider. As a result, nowadays, cross-border trips lasting for less than 24 hours are also regarded as international tourism (Cooper–Hall 2016). Due to the spatiotemporal differentiation of leisure activities, the availability of cross-border services and the loosening spatial ties of working, numerous new motivations behind travelling have emerged. A considerable proportion of these motivations are realized during short or day trips (Wynen 2013, Yousuf–Backer 2015). The European Union also recognized the importance of the phenomenon, and therefore the European Parliament and the Council introduced a regulation[1] in 2011 concerning the statistical measurement of trips that were not included in the previous (1995) directive. In Hungary, the Geographical Research Institute of the Hungarian Academy of Sciences (HAS), Research Centre for Astronomy and Earth Sciences (RCAES), has been conducting research since 2012 in the framework of OTKA[2] on the economic, social, and environmental characteristics of trips labelled 'invisible tourism' (Michalkó–Rátz 2013). The examination of Hungary's transit tourism began within the framework of this project (in cooperation with the Hungarian Central Statistical Office); the first results contributed towards the identification of the affected destinations (Kincses et al. 2014).

Theoretical background

Despite the fact that transit has been a noteworthy phenomenon for a long time in many Central-European countries, including Austria, the Czech Republic, Hungary, Slovakia, and Slovenia (at least a quarter of incoming visitors enter and exit the given country from two different sections of the border), still, the international literature hardly discusses this issue (Bakic 1988, Johnson 1995, Hall 1998). Among the causes behind the neglect of this topic by researchers, the nature of transit should be mentioned first, since transit is a rapid form of mobility in a relatively closed system, whose observation causes numerous difficulties. The main aim of transit is to shorten idle time and reach destinations as fast as possible, that is, bridge the section (which is often an individual entity, such as a state or microstate) between the place of origin

[1] Regulation (EU) No 692/2011 of the European Parliament and of the Council of 6 July 2011 concerning European statistics on tourism and repealing Council Directive 95/57/EC.

[2] Hungarian Scientific Research Fund.

and the destination with minimal stops and/or with the shortest possible length of stay. Therefore, data gathering concerning transit visitors, who are continuously under time pressure, is extremely difficult, as they are not willing to fill in questionnaires or answer interview questions; in addition, the service providers who contact them provide only limited information. Another possible reason behind the unrecognized potential in researching problems related to transit lies in the underestimation of the role and effects of the phenomenon. Within the group of foreigners arriving in a country, transit visitors (in relation to social, economic, and environmental effects of their one-sided activity) besides tourists and excursionists, have been hardly examined compared to their real share.

The international literature directly discussing transit primarily focuses on traffic occurring at the airports. In addition to the examination of various services available at international airports, with special regard to duty free shops (Graham 2009), various security issues, including border crossing, customs (Prager et al. 2015), and epidemiological risks (Malone et al. 2009) have been researched. It has been recognized that hub airports with favourable geographical location and significant numbers of transit passengers profit not only from the special demand of transit passengers staying at airport hotels but also from dynamic tourism development built upon the attractiveness of cities (Lohmann et al. 2009).

Studies dealing with transit indirectly are concerned with transportation and environmental issues, among other things. The assessment of the role of transportation companies in contributing to successful transit management (Schiefelbusch et al. 2007) and the contribution of the private sector (Gopalan–Narayan 2010) are among these research topics. In the latter case, neither the role of the informal (Michaud 1991) nor the sharing economy (Böhler et al. 2006) can be neglected in the process of understanding transit. The support services provided by local communities are crucial for reaching the inaccessible destinations of the third world and crossing areas with poor transport infrastructure. Environmental pollution associated with transit and ensuring the implementation of the sustainability principle along the affected routes are important (Filimonau et al. 2014) research areas.

Exploration of the travelling habits and tourism behaviour of transit visitors is seldom discussed in the literature. Within the topic of transit, hitchhiking (Laviolette 2016) represents a special field, which is, on the one hand, a lifestyle, and on the other hand, a phenomenon on the periphery of the sharing economy and the manifestation of social responsibility and solidarity at the same time. While the effects of hitchhiking in transit are indeed marginal, the behaviour of visitors arriving by caravans or motorhomes can be examined in the context of 'traditional' tourism as well. Due to the freedom and flexibility provided by caravans and/or motorhomes, it is not unusual for the motivation of transit to be coupled with an overnight stay at the destination (Green 1978). This can be planned (calculated), spontaneous (passengers get attracted to the place where they stopped), or can occur due to compulsion and other circumstances (unexpected events, traffic jam, etc.). The use of the caravan and/or motorhome (toll, parking fee, fuel), and its potential repair, passengers' consumption (retailing, catering), leisure activities, or overnight stay can already generate substantial spending.

Transit is not registered in the framework of accommodation statistics, since in most cases it is not coupled with an overnight stay in the given country (Parroco et al. 2012). In most EU countries belonging to the Schengen Area, the flow of EU citizens is not registered, thus this phenomenon is assessed by estimations, derived data, or data coming from statistical sampling (Volo–Giambalvo 2008). The international literature uses the terms unobserved or unmeasured tourism concerning mobility related to transit passenger traffic (De Cantis et al. 2015), while the Hungarian literature describes trips lasting for less than 24 hours and/or without overnight stay at registered accommodation as invisible tourism (Michalkó–Rátz 2013).

Spending is the most vital element of tourist behaviour (Rátz 2004). Spending makes a traveller become a tourist, and the use of paid services involves the visitor in tourism. Expenditure correlates with many factors determining travel, and the strongest correlation of expenditure is observed with motivation (Suh–McAvoy 2005). Within the structure of an average tourist's spending, the costs of accommodation, catering, events, and shopping have roughly the same share (Michalkó 2012).

Method

The Hungarian Central Statistical Office has been measuring transit visitors to Hungary since 2007 as part of the questionnaire called 'Foreigners' tourism-related and other expenditures (quarterly)'. The survey covers foreigners leaving and Hungarians entering Hungary by any vehicle, except for lorry. The study uses non-probabilistic and structured sample. Since 2008, the yearly sample size has been approximately 60,000, of which, the sample of transit visitors has been around 14,000 or 15,000. The selection of days is systematic and random. Data are collected through personal interviews.

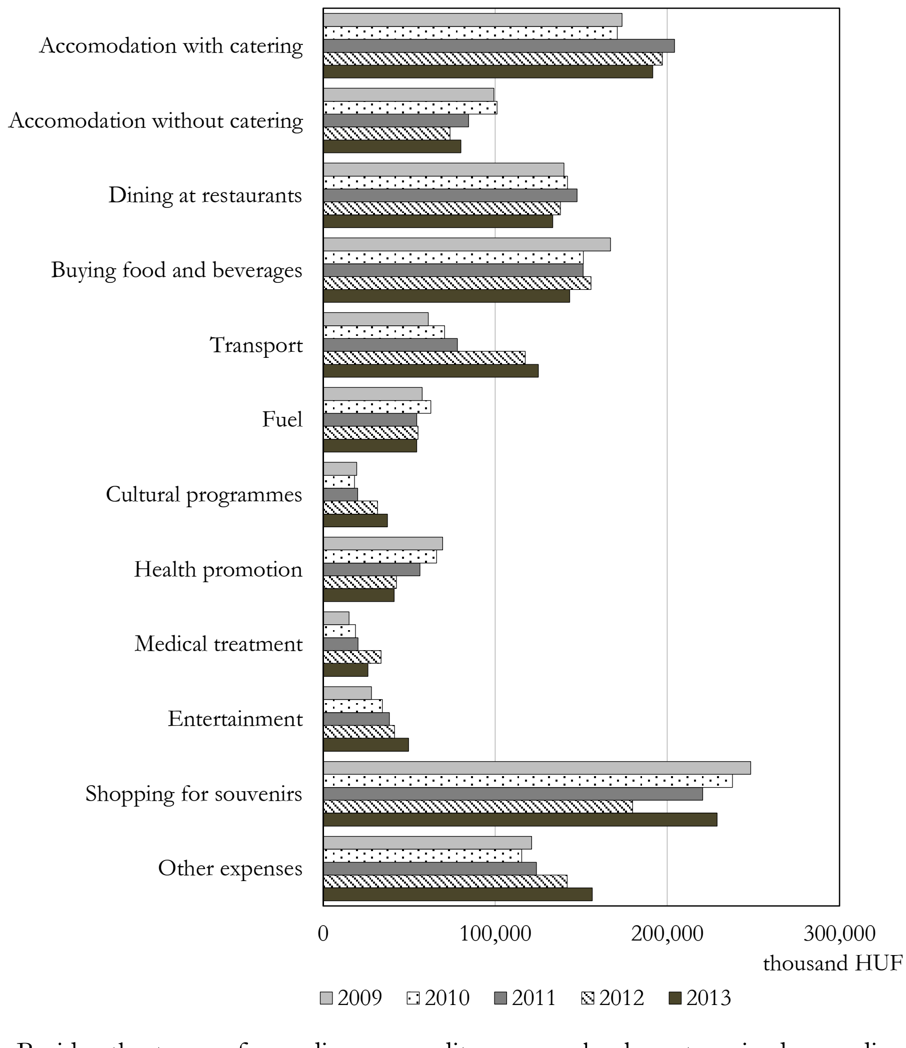
In our research, five years' data were examined. The basic reason behind the selection of the period between 2009 and 2013 is that data of the period after the economic crisis clearly differs from the preceding period. From 2009, however, it can be detected how data on transit reflects economic recovery.

Characteristics of foreign tourists

Foreign tourists' spending in Hungary can be categorized according to the types of spending motivations and nationalities. Tourists' spending in Hungary was increasing slowly, but continuously. During the examined five-year period, tourists' spending rose by 6%. The highest tourists' expenditure categories include 'shopping for souvenirs' and 'accommodation with catering'. From 2009 to 2013, the most dynamic increase in expenditures was observed in the following categories: transportation (205%), cultural programmes (192%), entertainment (177%), and medical treatment (173%).

Figure 1

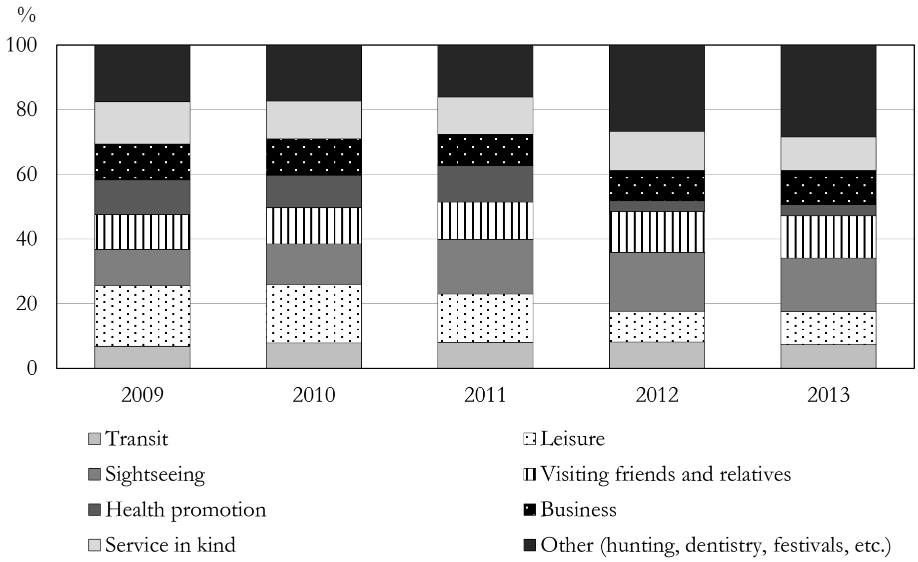
Total spending of tourists by type of spending



Besides the types of spending, expenditures can also be categorized according to travel motivations.

Figure 2

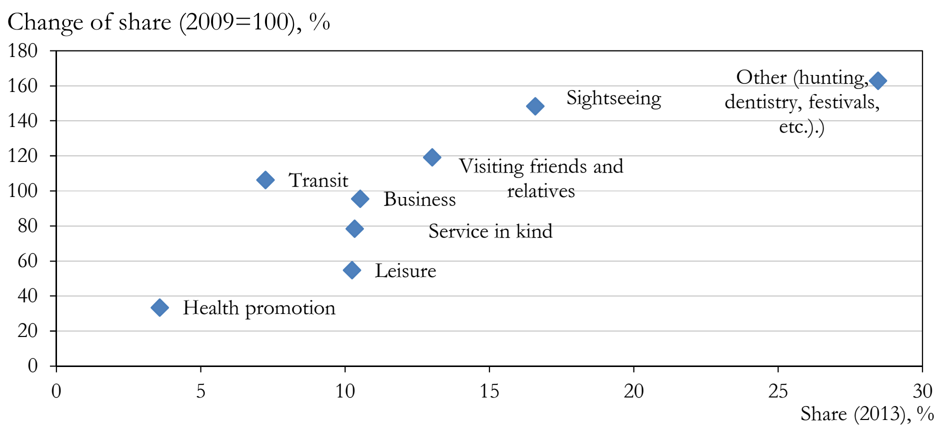
The distribution of spending of all tourists, according to their motivations



City tours (sightseeing) and visiting friends and relatives are the most important motivations of foreign tourists in Hungary. During the examined five-year period, the share of motivations of foreign tourists' spending (denominated in Hungarian forints) changed significantly. Traditional motivations (such as health promotion and leisure) lost ground, while the proportion of and the sum spent on hunting, dentistry, and sightseeing increased continually and dynamically.

Figure 3

The share of tourism motivations of total spending and its changes



Motivation and the type of spending are interconnected. The strength indicators measure the strength of the correlation between qualitative variables. The basic concept of association analysis is that the expected frequency is calculated assuming the independence of variables in every pair. The contingency table displays cross-correlations and the frequency distribution of the variables.

Table 1

Cross-correlations

	x_1	x_2	Σ
y_1	p_{11}	p_{12}	$p_{1.}$
y_2	p_{21}	p_{22}	$p_{2.}$
Σ	$p_{.1}$	$p_{.2}$	1

Let us consider a 2x2 table (non-symmetric), where $p_{ij}=P(X=x_i, Y=y_j)$, i,j=1,2. In this case, the independence of the two variables is measured by:

$$\Phi^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(p_{ij} - p_{i.}\cdot p_{j.})^2}{p_{i.}\cdot p_{j.}}$$

- By the normalization of $\mathbf{\Phi^2}$, we get Pearson's correlation:

$$P = \sqrt{\frac{\Phi^2}{\Phi^2 + 1}}$$

- P=0 independence
- P=1 functional
- $0 \le P \le 1$ stochastic relationship

Table 2

Strength of relationships between foreigners' spending and travel motivations

Year	ϕ^2	P	Proportion of transit of ϕ^2,%
2009	0.755	0.656	17.177
2010	0.734	0.651	17.065
2011	0.696	0.641	19.690
2012	0.611	0.616	21.363
2013	0.571	0.603	24.335

There is a moderate, although continuously weakening, relationship between foreigners' spending and travel motivations. While the value of Pearson's correlation coefficient decreased by 5% during the examined five-year period, the dependence of the types of spending and transit increased further. This means that typical types of spending, while this is not true of other travel motivations, characterize transit.

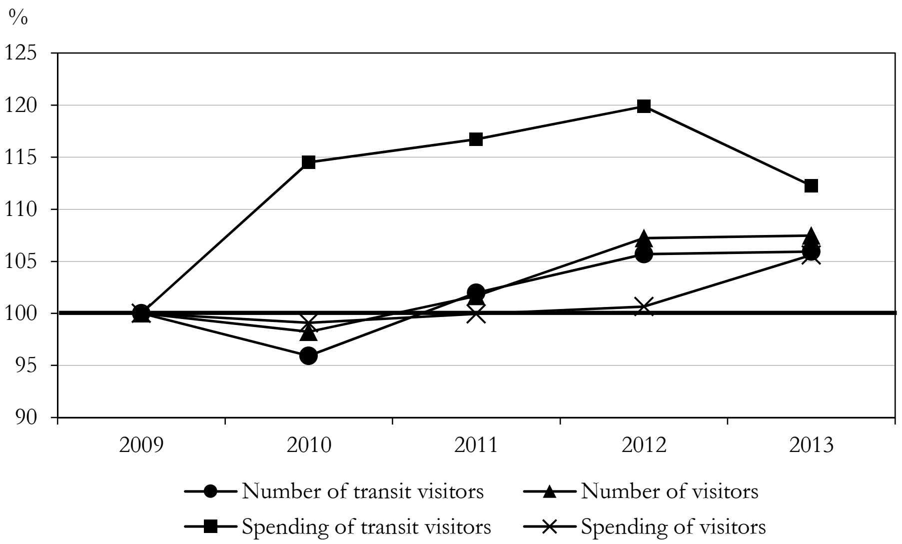
The share of transit visitors' spending accounts for 49% and 14% of foreigners' total expenditure on fuel and dining at restaurants, respectively, while they do not spend on health promotion, cultural programmes, or medical treatment.

The role of transit in Hungary's inbound tourism

In 2013, approximately 15,563 thousand transit visitors arrived in Hungary, and their expenditure was 91,914 million forints. From 2009 to 2013, the number of visitors increased by 6%, while their expenditure grew by 12% at current prices. As compared to the base year and the number of all visitors, the number of transit visitors rose one percentage point slower, while their expenditure surged six percentage points faster.

Figure 4

The number of foreign transit visitors to Hungary and their total expenditure (2009 = 100%)



As shown in tables 3 and 4, transit visitors to Hungary between 2009 and 2013 accounted for more than a third of foreigners arriving in Hungary. The share of the length of stay and expenditure of these transit visitors was 15–16% and 7–8%, respectively, which showed only minimal fluctuations from year to year. The per capita spending of transit visitors reached 18–20% of the total expenditure of foreign visitors during the period. The average length of stay, due to the nature of the activity, was 'only' one day, although there was a slight increase.

Table 3

Share of transit visitors of foreign visitors' total number, length of stay, and expenditure in Hungary

(%)

Year	Visitors	Length of stay	Expenditure
2009	36.2	15.6	6.8
2010	35.3	15.4	7.9
2011	36.3	16.0	8.0
2012	35.6	16.6	8.1
2013	35.6	16.2	7.2

Table 4

Comparison of transit and foreign visitors' total expenditure and length of stay

Year	Per capita expenditure, thousand forints		Average length of stay, days	
	Transit	total	Transit	Total
2009	5.6	29.6	1.03	2.39
2010	6.7	29.8	1.05	2.40
2011	6.4	29.1	1.06	2.40
2012	6.3	27.7	1.07	2.29
2013	5.9	29.0	1.06	2.33

Transit tourism according to nationality

Transit visitors' country of origin was analysed according to 2013 data. Our analysis took into account only the countries whose sample size is above 100, therefore related results are statistically relevant.

The impact of transit tourism, due to its nature, can only be identified within a limited geographical distance. Therefore, it is primarily the expenditure of tourists from Europe and especially from countries near Hungary that can be detected.

Concerning the frequency, length of stay, and spending of transit visitors to Hungary, the nationals of Romania, Serbia (together with Montenegro and Kosovo), and Bulgaria stand out. These three countries account for more than half the number, length of stay, and expenditure of transit visitors to Hungary. The average length of stay is the longest in the case of Italian, German, and Bulgarian nationals. There is an increasing tendency among these nationals (apparently because of the length of the journey) to break their journey and resume it after an overnight stay at an accommodation establishment.

The per capita expenditure is the highest among the Ukrainian nationals, followed by the Italians and Germans.

Table 5

Major characteristics of transit visitors to Hungary according to nationality, 2013

Country	Share (%)			Average length of stay, days	Per capita expenditure, thousand forints
	visitors	length of stay	expenditure		
Total	*100.0*	*100.0*	*100.0*	*1.06*	*5.91*
of which:					
Austria	4.7	4.6	3.7	1.02	4.67
Bulgaria	7.2	7.4	8.8	1.09	7.25
Czech Republic	3.3	3.4	3.2	1.07	5.57
Netherlands	0.6	0.6	0.7	1.08	7.04
Croatia	2.6	2.4	1.5	1.00	3.33
Poland	7.3	7.2	6.4	1.04	5.24
Germany	5.5	5.8	7.3	1.12	7.85
Italy	1.5	1.6	2.3	1.13	9.23
Romania	32.5	33.0	32.6	1.07	5.92
Serbia[a)]	13.6	13.3	11.5	1.03	4.99
Slovakia	5.2	4.9	3.3	1.01	3.78
Slovenia	1.1	1.1	0.5	1.00	2.44
Ukraine	5.1	5.2	8.3	1.08	9.61

a) Together with Montenegro and Kosovo.

Naturally, the most important part of transit visitors' expenditure is fuel costs, which accounts for 30% of their spending. It is followed by the category 'dining at restaurants' that makes up approximately 20% of the expenditure, while 'buying food and beverages' represents a slightly smaller proportion. Additional important expenditure categories include 'other expenses' and 'shopping for souvenirs'.

Table 6

Shares of the main spending types of transit visitors' total spending, 2013

(%)

Country	Accommodation with catering	Dining at restaurants	Buying food and beverages	Fuel	Shopping for Souvenirs	Other expenses
Austria	2.7	20.1	12.6	30.8	21.9	10.1
Bulgaria	3.4	18.5	12.7	39.6	8.5	14.7
Czech Republic	6.0	21.9	11.7	36.9	5.1	15.6
Netherlands	5.3	20.8	14.2	33.7	20.2	5.5
Croatia	0.0	20.8	19.6	23.2	6.0	28.2
Serbia[a)]	1.8	20.4	19.9	26.1	15.8	11.8
Poland	3.6	21.9	12.3	38.9	6.0	14.0
Germany	6.1	22.7	12.0	31.7	14.1	10.9
Italy	6.1	23.0	11.3	31.2	16.6	11.5
Romania	3.3	18.5	20.5	23.7	13.5	19.2
Slovakia	1.1	18.0	11.3	40.4	6.0	21.6
Slovenia	0.0	25.5	14.3	26.2	7.9	23.0
Ukraine	2.0	24.4	22.5	21.9	12.0	12.7
Average	*3.2*	*20.2*	*17.3*	*28.8*	*12.4*	*15.6*

a) Together with Montenegro and Kosovo.

Table 7

Number and expenditure of transit visitors according to the countries of origin

Country	2009			2013		
	visitors (B0)	expenditure, thousand forints (A0)	per capita expenditure, thousand forints (V0)	visitors (B1)	expenditure, thousand forints (A1)	per capita expenditure, thousand forints (V1)
Austria	610,150	4,722,363	7.74	737,167	3,442,160	4.67
Bulgaria	–	–	–	1,118,328	8,107,816	7.25
Czech Republic	656,343	2,397,038	3.65	521,254	2,904,563	5.57
Netherlands	115,411	773,133	6.70	90,672	638,031	7.04
Croatia	328,784	2,247,046	6.83	400,621	1,335,655	3.33
Poland	914,063	4,580,746	5.01	1,130,210	5,926,767	5.24
Germany	970,348	6,623,624	6.83	853,517	6,698,591	7.85
Italy	252,584	1,590,964	6.30	233,682	2,155,759	9.23
Romania	5,424,806	25,303,358	4.66	5,065,355	29,988,409	5.92
Serbia[a)]	1,140,318	8,620,381	7.56	2,122,500	10,591,261	4.99
Slovakia	1,050,741	2,988,138	2.84	802,685	3,034,746	3.78
Slovenia	146,487	381,116	2.60	176,038	429,404	2.44
Ukraine	618,018	5,435,836	8.80	791,861	7,609,932	9.61
Other	2,463,683	16,207,534	6.58	1,519,312	9,050,732	5.96
Total	*14,691,736*	*81,871,277*	*5.57*	*15,563,202*	*91,913,827*	*5.91*

a) Together with Montenegro.

By 2013, the number and expenditure of transit visitors had slightly increased as compared to 2009. Meanwhile, the per capita spending had increased from 5.57 thousand forints to 5.91 thousand forints.

The structure of Hungary's transit tourism is further analysed using the grand mean, part average, and composition effect indices.

The following marks are used:

$$I' = \frac{\sum B_{j1} V_{j1}}{\sum B_{j1}} : \frac{\sum B_{j1} V_{j0}}{\sum B_{j1}} = \frac{\sum B_{j1} V_{j1}}{\sum B_{j1} V_{j0}}$$

$$I'' = \frac{\sum B_{j1} V_{j0}}{\sum B_{j1}} : \frac{\sum B_{j0} V_{j0}}{\sum B_{j0}}$$

$$I = \overline{V}_1 : \overline{V}_0 = \frac{\sum A_{j1}}{\sum B_{j1}} : \frac{\sum A_{j0}}{\sum B_{j0}} = \quad = \frac{\sum A_{j1}}{\sum A_{j0}} : \frac{\sum B_{j1}}{\sum B_{j0}} = \frac{\sum B_{j1} V_{j1}}{\sum B_{j1}} : \frac{\sum B_{j0} V_{j0}}{\sum B_{j0}}$$

According to the analysis, change in part ratios has a positive effect on expenditures, while a change in the composition, change in the composition of the subpopulation (changes in the composition of countries), has a negative effect on the grand mean.

I = 105.98

I' = 110.36

I" = 96.04

The changes in the structure of spending were further examined by shift-share analysis. There is a large body of literature on this method, wherein numerous examples of applications in the field of tourism can be found (Houston 1967, Stevens–Moore 1980, Selting–Loveridge 1992, Andrikopolous–Carvalho 1990, Fuchs et al. 2000, Sirakaya et al. 2002, Toh et al. 2004, Yasin et al. 2004). Our research was based on these works, and hence the detailed description of the method and its various application possibilities are disregarded.

In the following section, a shift-share analysis is applied to examine the extent to which the structure of spending of all transit visitors explains the changes in expenditures from 2010 to 2013 and the extent to which other factors, which are only the characteristics of a given country, play a role.

From 2010 to 2013, the average spending of transit visitors from the examined countries grew by 1.03 times at current prices. In the 'total change' column of Table 8, 100 can be seen in the case of those countries where the change was bigger than average, while −100 is displayed where the change was smaller than average. This change can be split into two elements. The spatial (geographical) component shows the role of individual processes, concerning particular countries, in the total change. The structural (related to spending) component describes the importance of types of spending within the changes in total expenditures, with respect to the all the examined countries.

It can be concluded that the component related to particular countries plays a greater role than the structural element, since in absolute values the figures for the former are higher.

Table 8

Components of the changes in spending, 2010–2013

Country	Total change	Spatial (geographical)	Structural (spending)
Austria	–100.0	–11,696.3	11,596.3
Czech Republic	100.0	100.5	–-0.5
Croatia	–100.0	–101.5	1.5
Poland	100.0	101.5	–1.5
Germany	–100.0	–152.9	52.9
Italy	100.0	24.9	75.1
Romania	–100.0	–78.3	–21.7
Serbia[a)]	100.0	91.5	8.5
Slovakia	–100.0	1,119.9	–1,219.9
Slovenia	–100.0	–97.8	–2.2
Ukraine	100.0	106.4	–6.4

a) Together with Montenegro and Kosovo.

Table 9 shows that the changes in expenditures between 2010 and 2013 primarily affected transit visitors from Serbia (along with Montenegro and Kosovo), followed by visitors from Ukraine and Poland. The nationals of Romania, Germany, and Croatia experienced a smaller change than the average. The positive and negative role of the spatial component can be highlighted in the case of countries that have already been mentioned with respect to total change. The structure of spending changed greater than the average, especially in the case of visitors from Germany, Serbia (along with Montenegro and Kosovo), and Italy, while the change was smaller in the case of Romania, Slovakia, and Ukraine.

Table 9

Components of the changes in spending, 2010–2013

(%)

Country	Total+	Total–	Spatial+	Spatial–	Structural+	Structural–
Austria	0.0	0.0	0.0	1.2	6.7	0.0
Czech Republic	12.0	0.0	12.4	0.0	0.0	0.3
Croatia	0.0	17.0	0.0	17.7	1.5	0.0
Poland	16.7	0.0	17.5	0.0	0.0	1.5
Germany	0.0	18.1	0.0	28.5	54.6	0.0
Italy	3.7	0.0	0.9	0.0	15.6	0.0
Romania	0.0	62.2	0.0	50.0	0.0	76.9
Serbia[a)]	44.7	0.0	42.0	0.0	21.6	0.0
Slovakia	0.0	0.2	2.1	0.0	0.0	12.6
Slovenia	0.0	2.5	0.0	2.6	0.0	0.3
Ukraine	23.0	0.0	25.1	0.0	0.0	8.4
Total	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*

a) Together with Montenegro and Kosovo.

Table 10

Grouping of countries according to the components of changes in spending, 2010–2013

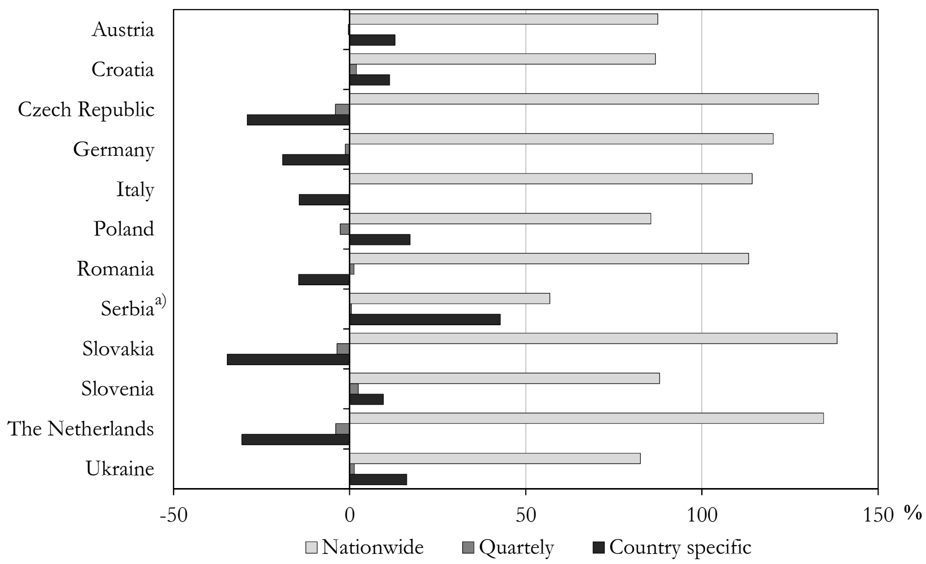
Spatial+ and structural+	Spatial+ and structural–	Spatial– and structural+	Spatial– and structural–
Italy	Czech Republic	Austria	Romania
Serbia[a)]	Poland	Croatia	Slovenia
	Slovakia	Germany	
	Ukraine		

a) Together with Montenegro and Kosovo.

Following the analysis of the changes in expenditures, the role of nationwide processes, seasonality, quarterly data, and country specific processes in the changes of the number of visitors and their expenditure form 2009 to 2013 was examined. It was concluded that, with respect to the number of transit visitors and their spending, the most important factor during the examined period was the *general tendency* characterizing Hungary's average transit traffic. It implies that the most important processes of transit do not depend on the quarterly structure of transit visitors or on a country specific factor; instead, they are closely related to general European tendencies (such as labour market conditions and tourism supply), and conditions (such as visa requirements, transport infrastructure, and accommodation along transit routes) provided for them by Hungary.

Figure 5

Factors in the changes in the number of transit visitors to Hungary, 2009–2013

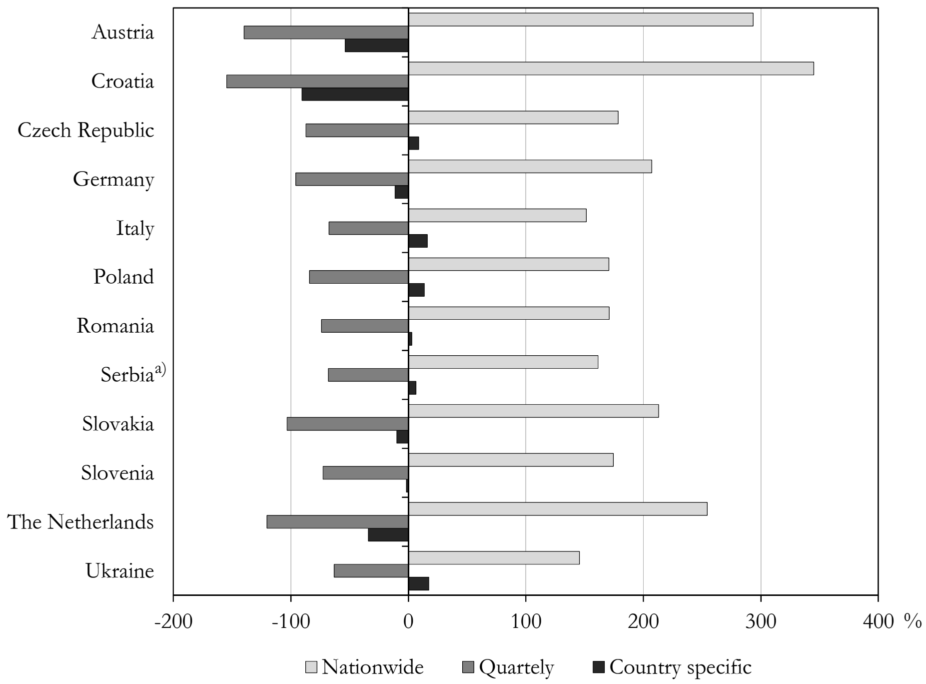


a) Together with Montenegro and Kosovo.

The effect of the quarterly structure, the quarterly distributions, on the change in the number of visitors is marginal, while its impact on the change in expenditures is more significant. While this factor has positive values in some countries in case of the number of visitors, the value is negative in all the cases related to expenditure. This shows that the rate of increase from quarter to quarter was frequently below the yearly average, between 2009 and 2013. The country specific factors influenced the number of visitors and their expenditures both positively and negatively, even though tendencies were determined by nationwide processes. The role of country specific factors in influencing the number of visitors is somewhat greater than in the case of expenditures.

Figure 6

Factors in the changes in the expenditure of transit visitors to Hungary, 2009–2013



a) Together with Montenegro and Kosovo.

The extent to which per capita spending depends on the time (quarterly distribution) of transit and on nationality was also examined. In the revenue surplus/deficit column of Table 11, a value of 100 is assigned to countries whose per capita spending is higher than the average of the examined countries in 2013, whereas −100 is assigned to countries whose per capita income is below the average of the countries examined during the period. It can be concluded that per capita spending is primarily determined by the visitor's country of origin. This is supported by the fact that the spatial/geographical factor is more important in absolute value than the quarterly structure.

Table 11

Revenue surplus/deficit and its components among transit visitors, 2013

Country	Revenue-surplus/deficit	Spatial	Quarterly
Austria	–100.0	–101.1	1.1
Bulgaria	100.0	90.7	9.3
Czech Republic	–100.0	–18.5	–81.5
Netherlands	100.0	125.0	–25.0
Croatia	–100.0	–103.8	3.8
Poland	–100.0	–60.1	–39.9
Germany	100.0	103.9	–3.9
Italy	100.0	99.4	0.6
Romania	100.0	–263.1	363.1
Serbia[a)]	–100.0	–104.5	4.5
Slovakia	–100.0	–90.0	–10.0
Slovenia	–100.0	–100.3	0.3
Ukraine	100.0	98.5	1.5

a) Together with Montenegro and Kosovo.

The per capita spending is the highest in cases of Ukrainian, German, and Bulgarian nationals. In contrast, Serbia (together with Montenegro and Kosovo), Slovakia, and Croatia are characterised by lowest per capita spending. The country specific factors explain the positive effect on per capita expenditure of transit visitors in cases of nationals of Ukraine, Germany, and Bulgaria, while this effect is negative in cases of Serbia (together with Montenegro and Kosovo), Slovakia, and Croatia. The quarterly structure, the over-representation of quarter(s), with high per capita spending is characteristic of nationals of Romania, Bulgaria, and Serbia (together with Montenegro and Kosovo). An opposite tendency is observed in cases of transit visitors to Hungary from Poland, Slovakia, and the Czech Republic.

Table 12

Revenue surplus/deficit and its components among transit visitors, 2013

Country	The role of revenue surplus/deficit		The role of geographical distribution		The role of the quarterly structure	
	positive	negative	positive	negative	positive	negative
Austria	0.0	12.8	0.0	13.3	1.5	0.0
Bulgaria	21.3	0.0	19.9	0.0	20.1	0.0
Czech Republic	0.0	2.4	0.0	0.5	0.0	20.0
Netherlands	1.5	0.0	1.9	0.0	0.0	3.7
Croatia	0.0	14.5	0.0	15.5	5.5	0.0
Poland	0.0	10.5	0.0	6.5	0.0	42.6
Germany	23.4	0.0	25.1	0.0	0.0	9.3
Italy	11.0	0.0	11.2	0.0	0.6	0.0
Romania	1.4	0.0	0.0	3.9	52.9	0.0
Serbia[a)]	0.0	27.2	0.0	29.3	12.6	0.0
Slovakia	0.0	24.0	0.0	22.2	0.0	24.4
Slovenia	0.0	8.6	0.0	8.9	0.2	0.0
Ukraine	41.4	0.0	42.0	0.0	6.5	0.0
Total	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*

a) Together with Montenegro and Kosovo.

The seasonality of transit visitors to Hungary was also examined in this study. Seasonality was quantified by the calculation of the ratio between the highest and lowest quarters. It was concluded that the degree of seasonality decreased between 2009 and 2013 in the examined countries. Concerning transit visitors, the Polish, Slovak, and Dutch nationals showed the greatest seasonality, while Bulgarian, Romanian, and Croatian nationals displayed the lowest seasonality.

Table 13

Ratios of the quarters with the highest and lowest number of visitors, 2009–2013

Country	2009	2010	2011	2012	2013
Austria	4.6	3.1	4.0	2.3	2.3
Bulgaria	–	–	1.7	1.7	1.4
Czech Republic	11.4	12.5	14.1	13.4	7.2
Netherlands	9.7	12.9	11.1	16.0	8.5
Croatia	1.5	1.5	1.9	1.2	1.8
Poland	11.3	8.7	11.1	8.0	9.7
Germany	3.6	4.7	3.8	3.7	3.4
Italy	2.1	2.3	2.3	2.7	2.4
Romania	1.7	1.7	1.7	1.7	1.8
Serbia[a)]	1.3	1.8	2.1	1.9	2.4
Slovakia	10.5	11.1	21.5	9.0	7.4
Slovenia	1.2	1.7	133.4	1.7	2.9
Ukraine	1.7	1.6	2.6	1.7	2.0

a) Together with Montenegro and Kosovo.

Conclusions

The most important objective of our research, based on the major research results of our previous study (Kincses et al. 2014), was to highlight the role of transit tourism and justify its significance by statistical examination. Our investigations undoubtedly prove that, although this research topic has been relatively neglected, the topic requires a scientific analysis. This is proved by the fact that the share of spending by transit visitors comprises a significant portion of foreigners' total spending in Hungary; in addition, the spending of these visitors is registering a continuous rise. Between 2009 and 2013, the motivations of transit visitors showed a slight change, as traditional motivations were replaced with new ones. Tourism professionals and economic policy makers should also consider these changes.

Our research has also proved that transit visitors' most important types of spending are fuel costs and dining at restaurants. The nationals of Romania, Serbia (together with Montenegro and Kosovo), and Bulgaria have the highest shares of transit visitors to Hungary. During the examined period, the number and expenditure of transit visitors slightly increased and per capita spending also increased. It was shown that the reason for this is the change in the composition of the affected countries.

Our research results show that the country of origin is a more important factor than the structure of spending for determining the change in transit visitors' expenditure. It was concluded that the most important processes of transit do not depend on the quarterly structure (distribution) of transit visitors or country specific factors, but on general European tendencies (labour market conditions, tourism supply, among others) and conditions (visa requirements, transport infrastructure, accommodation along transit routes, etc.) provided for them by Hungary. In our opinion, in following these European trends, it is necessary that all actors in Hungary's tourism make the right decisions to exploit the economic benefits of transit tourism fully and avoid its drawbacks.

Acknowledgement

The study was based on the research supported by the Hungarian Scientific Research Fund under the OTKA K 100953 project (Invisible tourism in Hungary).

REFERENCES

ANDRIKOPOLOUS, A. B. J.–CARVALHO, E. (1990). Shift-share analysis and the potential for predicting regional growth patterns: some evidence for the region of Quebec, Canada *Growth and Change* 21 (1): 1–10.

BAKIC, O. (1988): Yugoslavian tourism – strategic planning of destination product *Tourism Management* 9 (2): 145–151.

BÖHLER, S.–GRISCHKAT, S.–HAUSTEIN, S.–HUNECKE, M. (2006): Encouraging environmentally sustainable holiday travel *Transportation Research Part A: Policy and Practice* 40 (8): 652–670.

COOPER, CH–HALL, M. (2016): *Contemporary tourism – an international approach* Goodfellow Publishers Limited, Oxford.

DE CANTIS, S.–PARROCO, A.–FERRANTE, M.–VACCINA, F. (2015): Unobserved tourism *Annals of Tourism Research* 50: 1–18.

FILIMONAU, V.–DICKINSON, J.–ROBBINS, D. (2014): The carbon impact of short-haul tourism: a case study of UK travel to Southern France using life cycle analysis *Journal of Cleaner Production* 64 (1): 628–638.

FUCHS, M.–RIJKEN, L.–PETERS, M.–WEIERMAIR, K. (2000): Modelling Asian incoming tourism: a shift-share approach *Asia Pacific Journal of Tourism Research* 5 (2): 1–10.

GOPALAN, R.–NARAYAN, B. (2010): Improving customer experience in tourism: A framework for stakeholder collaboration *Socio-Economic Planning Sciences* 44 (2): 100–112.

GRAHAM, A. (2009): How important are commercial revenues to today's airports? *Journal of Air Transport Management* 15 (3): 106–111.

GREEN, F. (1978): Recreation vehicles – a perspective *Annals of Tourism Research* 5 (4): 429–439.

HALL, D. (1998): Tourism development and sustainability issues in Central and South-eastern Europe *Tourism Management* 19 (5): 423–431.

HOUSTON, D. B. (1967): The shift and share analysis of regional growth: a critique *Southern Economic Journal* 33 (4): 577–581.

JOHNSON, M. (1995): Czech and Slovak tourism – patterns, problems and prospects *Tourism Management* 16 (1): 21–28.

KINCSES, Á.–MICHALKÓ, G.–TÓTH, G. (2014): A Magyarország átutazó forgalmában érintett települések lehatárolása *Területi Statisztika* 54 (3): 237–252.

LAVIOLETTE, P. (2016): Why did the Anthropologist Cross the Road? Hitch-Hiking as a Stochastic Modality of Travel *Ethnos* 81 (3): 379–401

LOHMANN, G.–ALBERS, S.–KOCH, B.–PAVLOVICH, K. (2009): From hub to tourist destination – An explorative study of Singapore and Dubai's aviation-based transformation *Journal of Air Transport Management* 15 (5): 205–211.

MALONE, J.–BRIGANTIC, R.–MULLER, G.–GADGIL, A.–DELP, W.–MCMAHON, B.–LEE, R. –KULESZ, J.–MIHELIC, M. (2009): U.S. airport entry screening in response to pandemic influenza: Modeling and analysis *Travel Medicine and Infectious Disease* 7 (4): 181–191.

MICHALKÓ, G. (2012): *Turizmológia* Akadémiai Kiadó, Budapest.

MICHALKÓ, G.–RÁTZ, T. (2013): Rejtett dimenziók a Kárpát-medence turizmusában In: FRISNYÁK, S.–GÁL, A. (szerk.): *Kárpát-medence: természet, társadalom, gazdaság* pp. 463–476, Bocskai István Gimnázium – Nyíregyházi Főiskola Turizmus és Földrajztudományi Intézet, Szerencs – Nyíregyháza.

MICHAUD, A. (1991): A social anthropology of tourism in Ladakh, India *Annals of Tourism Research* 18 (4): 605–621.

PARROCO, A.–VACCINA, F.–DE CANTIS, S.– FERRANTE, M. (2012): Multi-destination trips and tourism statistics: empirical evidences in Sicily. Economics: The Open-Access *Open-Assessment E-Journal* 6 (2012-44): 1–27.

PRAGER, F.–ROSE, A.–WEI, D.–ROBERTS, B.–BASCHNAGEL, CH. (2015): Economy-wide impacts of reduced wait times at U.S. international airports *Research in Transportation Business & Management* 16: 112–120.

RÁTZ, T. (2004): *European Tourism* Kodolányi János Főiskola, Székesfehérvár.

SCHIEFELBUSCH, M.–JAIN, A.–SCHAFER, T.–MÜLLER, D. (2007): Transport and tourism: roadmap to integrated planning developing and assessing integrated travel chains *Journal of Transport Geography* 15 (2): 94–103.

SELTING, A. C.–LOVERIDGE, S. (1992): *A summary of the literature on shift-share analysis* Department of Agricultural and Applied Economics, University of Minnesota College of Agriculture St. Paul, Minnesota USA.

SIRAKAYA, E.–CHOI, H. S.–VAR, T. (2002). Shift-share analysis in tourism: examination of tourism development change in a region *Tourism Economics* 8 (3): 303–324

STEVENS, B. H.–MOORE, C. L. (1980): A critical review of the literature on shift-share as a forecasting technique *Journal of Regional Science* 20 (4): 419–437.

SUH, Y.–MCAVOY, L. (2005): Preferences and trip expenditures – a conjoint analysis of visitors to Seoul, Korea *Tourism Management* 26 (3): 325–333.

TOH, R. S.–KHAN, H.–LIM, L. L. (2004). Two-stage shift-share analyses of tourism arrivals and arrivals by purpose of visit: the Singapore experience *Journal of Travel Research* 43 (1): 57–66.

UNWTO 2015: *Tourism Highlights 2015* UNWTO, Madrid.

VOLO, S.–GIAMBALVO, O. (2008): Tourism Statistics: Methodological Imperatives and Difficulties: The Case of Residential Tourism in Island Communities *Current Issues in Tourism* 11 (4): 369–380.

WYNEN, J. (2013): Explaining travel distance during same-day visits *Tourism Management* 36: 133–140.

YASIN, M.–ALAVI, J.–SOBRAL, F.–LISBOA, J. (2004). A shift-share analysis approach to understanding the dynamic of the Portuguese tourism market *Journal of Travel and Tourism Marketing* 17 (4): 11–22.

YOUSUF, M.–BACKER, E. (2015): A content analysis of Visiting Friends and Relatives (VFR) travel research *Journal of Hospitality and Tourism Management* 25: 1–10.

The uneven transformation of consumption spaces and the rise of new marginalities in Hungary

Erika Nagy
Institute for Regional Studies, Centre for Economic and Regional Studies, Hungarian Academy of Sciences, Békéscsaba, Hungary
E-mail: nagye@rkk.hu

Gábor Nagy
Institute for Regional Studies, Centre for Economic and Regional Studies, Hungarian Academy of Sciences, Békéscsaba, Hungary
E-mail: nagyg@rkk.hu

Gábor Dudás
Institute for Regional Studies, Centre for Economic and Regional Studies, Hungarian Academy of Sciences, Békéscsaba, Hungary
E-mail: dudasg@rkk.hu

The optimistic vision of a balanced socio-spatial development trajectory of the European economic space has been constantly challenged by the rise of new dimensions of socio-spatial inequalities and, recently, by the polarising effects of the global economic crisis. Inequalities emerged not only as income differentials but also as various forms of material deprivation, such as having limited access to basic goods, services, and housing and struggling with permanent financial difficulties. In this study, we focus on spatial processes produced by various institutional–corporate and administrative-bureaucratic–strategies that changed the 'material' conditions of consumption practices profoundly, before and after the crisis. The critical political economic concept that we employ allows us to consider consumption as a socio-spatial practice in which various (class, gender, family, etc.) social relations manifest, and, as a social act, are embedded into a 'material reality' of consumption spaces. Thus, we consider the on-going restructuring of consumption spaces and the related practices as manifestations of uneven development and the underlying (spatial) logic of capital. We seek to reveal the following: i) how localities were polarised by the influx of retail capital and the spread of new retail forms; (ii) if a shift towards a tighter control in retail and consumption enhanced or eased the unevenness of consumption landscapes; and (iii) whether the polarisation processes were accelerated by the crisis (the spatially uneven decline of retail investments and household incomes). To achieve these objectives, we rely on a combined methodology including the analysis of spatial statistics on retail restructuring and consumption in relation to corporate strategies and changing regulations. For the latter, we employ the content analysis of the relevant policy documents and media discourses as well as empirical results of fieldworks (interviews with executives) targeting various retail organisations. Our results reflect the role of retail capital–multiple territorial embedding–in driving uneven development during crisis through corporate strategies for protecting the spatially fixed assets of retailers. The retailers' spatial strategies should be understood as the manifestations of the responses to the crisis of the financialised Central and Eastern European (CEE) markets (declining demand, indebtedness, and flight of capital from peripheral markets) that differentiated consumers by income and by place of residence across Hungary. This process was also facilitated by the state through the parallel employment of an interventionist retail policy and a re-fashioned and socially selective redistributive system, which fuelled the emergence of 'deserving' and 'undeserving' consumer groups–marginalizing primarily the poor living in peripheral (mostly, rural) spaces within Hungary.

Keywords:
retail,
consumption,
uneven development,
socio-spatial polarisation,
marginality

Introduction

While the economic upturn of the early 2000s raised the optimistic vision of a balanced socio-spatial development trajectory in the European economic space, the in-depth analyses of the geographical distribution of wealth and risk across Europe suggested a more differentiated view (ESPON 2014). The enhanced flows of capital and the unfolding new spatial division of labour in the pre-crisis period produced new inequalities at the sub-national scale: household incomes and consumption became increasingly dependent on the labour market position as the welfare systems were replaced by the practices of the workfare state (TÁRKI 2009). Such processes produced a highly uneven socio-spatial 'landscape' within Europe that manifested at various scales (in metropolitan-non-metropolitan, urban-rural, regional-local inequalities) and differentiated the social spaces of the European periphery deeply and thoroughly (Ehrlich et al. 2012, Lang 2015, Stenning et al. 2011). Inequalities emerged not only as income differentials but also as various forms of material deprivation, such as having limited access to basic goods, services and housing, and struggling with permanent financial difficulties (utility bills and unexpected expenses) (OECD 2015).

The post-2008 downturn – that we interpret as a mechanism of neoliberal capitalism spreading the risks, and then the consequences of enhanced global capital flows to redistribute wealth in favour of capital at the expense of labour and poor (Hadjimichalis 2011, Harvey 2010) – polarised the highly uneven social spaces further. The crisis hit mostly the vulnerable social groups and peripheral spaces through capital extraction (debt crisis; downsizing productive capacities; devaluing household assets), enforced austerity schemes, that shrank collective consumption (Aalbers 2009, Harvey 2010, Hudson–Hadjimichalis 2014). The growing depth and scale of poverty were particularly shocking in the European periphery (in the East and the South), where an increasing pool of households was put at risk of getting even daily meals (TÁRKI 2014).

To understand how profound macroeconomic changes transformed the European periphery, in particular, how social and spatial marginality reinforced one another in context of neoliberal capitalism, we focussed on spatial processes produced by various institutional–corporate and administrative-bureaucratic–strategies that changed the 'material' conditions of consumption practices profoundly, before and after the crisis. The *critical political economic concept* that we employed allows us to consider consumption as follows: (i) a socio-spatial practice in which various (class, gender, family, etc.) social relations manifest (ii) a social act embedded into a 'material reality' of consumption spaces (Goss 2006, Miller 1999, Wrigely et al. 2005, Zukin–Maguire 2004). Thus, we consider the on-going restructuring of consumption spaces and the related practices as manifestations of uneven development and the underlying (spatial) logic of capital (Smith 1991) through which we can reveal the mechanisms of the reproduction of socio-spatial marginalities.

The above processes shall be discussed in the context of the Hungarian economy that we consider here as a laboratory of corporate strategies and regulative-bureaucratic experimentation in the European periphery (Hess 2004, Coe–Hess 2005). The profound changes in the organisation of production and distribution (capital and organisational concentration, enhanced networking, and spatial restructuring) stimulated by the liberalisation processes and the influx of retail capital (Nagy 2009b) were linked intimately to changing practices and perceptions of consumption. Purchasing goods and services was re-contextualised in the emerging capitalist society – connecting citizenship, social status/identity, and consumption – backed by the unfolding of the 'cultural industry' of global capitalism[1] and domestic discourses dominated by the principles of Neoliberalism (Szalai 2006). Consumption practices were researched intensively – particularly in sociology and marketing studies – with a strong focus on consumption as an indicator of social restructuring and a

[1] We refer here to Adorno's concept of cultural industry as a system functioning to hide the social relations of capitalist production and to control the consumers by producing and attaching meanings and values to goods (see Jameson 1974).

source of identity/self-reflection. These studies have highlighted the decisive role of social status (in our approach: class position) in consumption practices and the on-going differentiation of the society along class and non-class (age, urban/rural context) related characteristics in Hungary (Hetesi et al. 2007, TÁRKI 2008, 2014). However, the analyses were abstracted from their material (spatial) context, and thus failed to consider the wider social context (the 'institutional field'[2] defined by relationships of the state, retailers, NGOs, and consumers) in which consumption takes place. Meanwhile, recent geographical analyses of retail and consumption were limited mostly to the spatial organisation of consumption practices and were scarcely linked to theoretical discourses on uneven development and social relations of capitalism.

In this study, we provide an overview of the spatial restructuring of consumption spaces to shed some light on the role of spatial logic of retail capital and the related regulative frameworks in the (re)production of socio-spatial marginality in a crisis-hit peripheral economy. We seek to reveal the following: (i) how localities were polarised by the influx of retail capital and the spread of new retail forms; (ii) if a shift toward a tighter control in retail and consumption enhanced or eased the unevenness of consumption landscapes; (iii) whether the polarisation processes were accelerated by the crisis (the spatially uneven decline of retail investments and household incomes). In this way, we enrich consumption studies by focussing on the institutional-spatial contexts of purchasing goods, and contribute to the political-economic concept of spatial inequalities and marginality by embracing consumption in the analysis as a set of social relation(s) and as a spatial practice.

In the following sections, we have conceptualized retail capital and consumption as spatial processes in a political economic approach to understand how they contribute to uneven development in the specific context of a peripheral economy (highly contested/an emerging regulative system) (chapter 2). The restructuring of the retail sector/concentration and centralisation processes in relation to changing regulations and global capital flows shall be discussed in chapter 3 and the spatial manifestations of such processes analysed in chapter 4. The key findings on role and relations of various market agents are summarized in chapter 5.

Uneven development and the political economy of retail and consumption – Conceptual framework

Consumption and retailing have been considered mostly as separate realms in domestic academic discourses, and conceptualizations that relate to retail restructuring and changing consumption practices as spatial processes are particularly scarce–largely due to the belated 'spatial turn' in social sciences. In the following

[2] Bourdieu, 1984.

section, we consider retailing and consumption as sets of social actions that are tied intimately, and, going beyond the self-evident relationships (e.g. retail shops as scenes to consumption acts), we interpret them as momentums of the productive process in which capitalist social relationships are manifested. This critical political economic approach–rooted in Marx's widely discussed and contested concept of commodity fetishism–puts the underlying systems of exchange and the role of retail in the circuits of capital in focus (Goss 2006, Wrigley 1996). Accordingly, the strategies of retailers, which are driven by the logic of capital (efficient use of resources, increasing productivity, etc.), should be analysed in the context of the production process (Clarke 1996, Goss 2006): retailers exploit their role as a mediator to accelerate the circulation of capital to enhance the profit rates (Dicken 2007, Wrigley et al. 2005). Their strategies rest on the extension of control over production as well as on the re-definition of the role of consumption by creating new aesthetics and institutional contexts for producing demand (Zukin–Maguire 2004) and introducing organisational and technological innovation.

In this approach, space, various social interests, activities, and interpretations are mutually constitutive. Consumption spaces, 'where everyday life meets the machinations of capitalism' (Clarke 1996: 295), are not just the products of retailers' strategies but of a wide spectrum of social relationships (regulative environments, suppliers' networks, local politics, etc.) that corporate strategies are embedded into (Wrigley et al. 2005).

To understand the mechanisms that produced a highly uneven consumption landscape within Hungary, here we focus on two powerful groups of agents as drivers of changes: (i) retailers and (ii) the state.

(i) Retailers produce 'material' conditions of consumption by organising global flows of goods and commodity chains; in addition, they act as mediators of related ideologies by constructing spaces for shopping in a rapidly changing market. Although, it is a highly diverse assemblage of various agents, retail capital has its distinctive spatial logic. Retailers are strongly dependent on their territorialized relationships–on their local consumers, on the regulative-institutional context to which they have to adapt through a permanent learning process, and on their (increasingly transnational *and* regionalised) supply networks. Moreover, retailers' assets majorly include spatially fixed elements (shops and logistic infrastructure) that make market exit a costly step. Thus, highly selective location planning and bargaining on (national/local) regulations are key elements of spatial strategies of retailers.

(ii) The state is the primary agent that controls the conditions of (retail) capital accumulation by regulating the relationships of market agents. The global expansion of retail capital challenged state institutions that responded by introducing new and tighter regulations on FDI, new retail forms, land use, competition law, and international trade (Coe–Hess 2005). Meanwhile, regulation was also rescaled in supranational/national and central/local state nexus (the consumer protection and competition law at the EU-level and the control of retail restructuring through local land use regulation unfolded across CEE). It was a contested process shaped by a series of bargaining (state/various groups of retail capital) and conflicts of various interests (state/capital; local/central state; between various groups of retailers), which was complicated further by the inner logic and internal struggles of state bureaucracy (Jessop 2004). In CEE, weak intermediaries and paternalistic relationships in the state bureaucracy led to the gradual emergence of a contradictory regulative environment and institutional practices (Hess 2004) that contributed largely to the prevailing (global) centralisation trends in retailing and in the production of consumption spaces.

The spatial logic of retail capital and the deficiencies of the regulative system produced a highly differentiated consumption landscape in Hungary. The process was embedded within the uneven development of the European economy and its emerging peripheries, stimulated by the liberalization of trade and capital flows (Hadjimichalis 2011) and manifested as a consumption gap between the European core and the Eastern periphery[3], as a debt crisis that hit Hungary, particularly from 2009[4], and as being dependent on retail transnational corporations-led (TNCs) production networks (Keller 2008, Nagy 2012, OECD 2015). However, the spatial organisation of consumption added a new–yet scarcely discussed–dimension to the on-going socio-spatial differentiation at sub-national scale: the centralisation processes in retailing (enhanced by the recent crisis) marginalized rural spaces, particularly, the low-income and immobile social groups living in those spaces.

In the following section, we endeavour to give an overview of the changing socio-spatial context of consumption in the pre-/post-crisis times to understand how uneven development and marginality 'materialise' in the conditions of everyday life in various spaces. For this, we employed a multi-tiered methodology. We relied on the analysis of macro- and regional databases of the Central Statistical Office and of GfK Hungary on consumption and retail to explore the spatial and temporal changes in and the causal relationships of consumption patterns, retail restructuring, and social macro-processes. The analysis focussed on local scale to reveal the expansion of marginalised consumption spaces in line with the centralisation and 'thinning' of the

[3] The proportion of households reporting problems to get food and deal with unexpected expenses is extremely high in CEE and in Southern European countries before the crisis (Ward et al. 2009). The recent results of a recent research project (TÁMOP-4.2.2.A-11/1/KONV-2012-0069) project focussed on Hungary suggested that rural households in marginalized spaces are in even worse situation (40% struggle to get food and 85% cannot deal with unexpected expenses).

[4] Here, we interpret debt expansion as a mechanism through which consumption is stimulated to maintain the profit rate, while, in the time of crisis, capital is withdrawn from risky/underperforming sectors and regions hitting vulnerable social groups and peripheral regions heavily (Harvey 2011).

network of retail centres in the crisis period. The discussion of the uneven post-crisis recovery rests on the regional and municipal-scale analysis of incomes, purchase power, and retail sales. The data was processed in cartograms by using RegioGraph software of GfK GeoMarketing.

To understand power relations (capital/state; national/domestic capital; labour/capital) in the retail sector that shaped the material conditions of consumption, we analysed relevant acts and policy documents as well as related news available in the electronic media and the key agents' (OKSZ, Kisosz, major retailers') websites. We focussed on the way the power relations manifested in the content–using pre-defined categories based on the relationships of market agents–and the way they were interpreted and responded to by retailers, to understand discursive context of the regulations. Changing regulations and their impact on corporate strategies have explanatory power for the reproduction of socio-spatial inequalities in consumption and its material conditions. The contextual analysis also rested on a series of empirical studies, mostly semi-structured interviews with corporate executives[5] and local retailers[6], focussing on business strategies in CEE/Hungarian retail markets, with a major emphasis on location policy, changing regulations, and corporate responses to the crisis.

Capital flows, regulation, and restructuring in the retail sector

The emerging economies of Europe have been scenes to growing retail investments as well as to corporate experimentation since the mid-1990s. CEE markets–being transformed dynamically by the privatization and restitution processes as well as by the domestic enterprising rush–were targeted by retail TNCs that accelerated centralisation processes in terms of capital, organisation, and spatial processes. Transnational corporate strategies employed in the European periphery were embedded within a global restructuring of the sector, fuelled by a sluggish demand and increasingly competitive market conditions in the core economies, to which key market agents (e.g. Safeway, Ahold, Delhaize, Tesco, Carrefour) responded by improving efficiency (IT-based reorganisation of flows), innovation (e.g. introducing new retail forms), and transnational expansion (combining organic growth and M&As) from the late 1980s. They entered the rapidly liberalised and inconsistently regulated, yet risky, CEE markets to exploit retail growth potentials and property markets that produced high yields (over 10%) in the late 1990s (Colliers International

[5] The interviews were conducted in 2004-2006 and 2008-2011 within the Bolyai postdoctoral program (Hungarian Academy of Sciences).

[6] The interview series were part of the project supported by the National Research, Development and Innovation Office grant OTKA/109296 (Institutional and individual responses to state restructuring in different geographical contexts).

2007, 2014). The latter was a key move towards improving corporate performance spoiled by sunk costs[7] in core markets and inducing spatially selective strategies with a strong urban (primarily, metropolitan) focus. The expansion of retail capital was stimulated by the financialisation of the retail sector–its increasing dependence on financial investors and the growing pressure to get high returns[8]–and household consumption through extensive lending (2000–2008) (Baud–Durand 2011).

The retail market of Hungary was transformed largely by the above processes, such as the growth and spread of investments and series of innovations that made the domestic market increasingly contested and uneven from the early 2000s (Nagy 2009a). The 2000–2008 period should be considered as the peak of retail investment and restructuring that was fed largely by global capital flows (based on cheap lending) in retail and property development as well as by the growing domestic demand. The latter was boosted by raises in the public sector and and in the minimum wage level (2002, 2003) and by extensive lending that was exploited by households to maintain the consumption level after fiscal restrictions (2004, 2006) that caused a decline in social transfers and public sector wages (Bauer et al. 2013, Szívós–Tóth 2013).

The retail investment boom stimulated profound changes in the sector and beyond. Major retail corporations' strategies stimulated a rapid concentration in terms of capital and organisation through the following:

(i) Expansion across space and various segments of retailing and related services and the efficient use of new facilities (e.g. providing logistic and IT-services to domestic retailers through buying groups, franchise systems, etc.);

(ii) Gaining control over the production process by organising and controlling supplier networks at European/CEE scale, introducing own-label goods, and building own production bases by exploiting domestic capacities (Nagy 2014a);

(iii) Adapting to the changing social context of retailing, such as tighter regulations and changing consumption patterns through market repositioning, enhancing the selection of goods, and facilitating organisational innovations.

The expansion was reflected by the growing share in investments of retailers–particularly from 2001–that reached 42% (on average) in the 2001–2008 period. The growth took place in a period when the overall capital influx was growing rapidly in the sector. This process was a topping (a) in the pre-accession years that coincided with the major retailers' expansion targeting small- and medium-sized towns across Hungary (240–260 bn HUF/year), (b) in the 2007–2008 period (350–355 bn

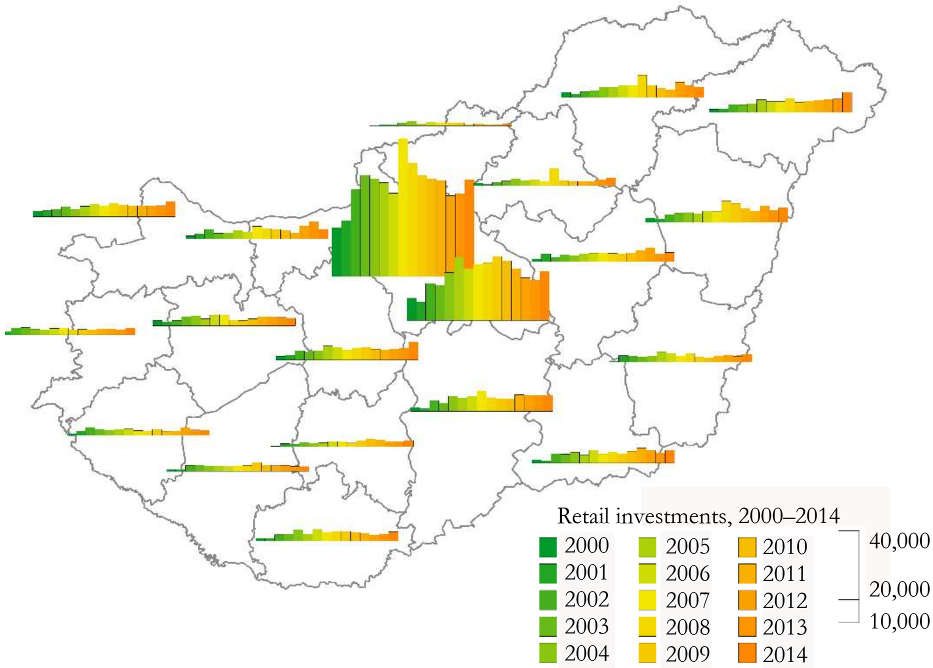
[7] Sunk costs are unrecoverable investments in case of market exit (Wrigley 1996). However, in our extended definition, crises-hit property market also produces sunk costs, as market processes are devalorising spatially fixed capital (e.g. the retail store). It is a spatially selective process.

[8] Since retail developments were linked to property market processes intimately due to its strong spatial embedding, financialisation induced aggressive expansion and raised a major concern with location choices contributing spatial (urban) concentration of investments—as it was suggested by property market reports and our interviews with retail executives.

HUF/year), when the expansion of new entrants, such as the discounters (Aldi, Lidl) in the food sector, pressed 'embedded' agents to respond through new investments and reorganisations[9] (Figure 1). By 2008, foreign capital accounted for a share of 49.4% of the total retail investments. International investors' activity followed a hierarchical and spatially centralized spatial pattern, with a persisting dominance of Budapest and Pest county, and new consumption spaces (hypermarkets, shopping malls) within non-metropolitan areas (KSH 2008).

Figure 1

The spatial distribution of investments in retail, wholesale, and repair (sector 'G')



Source: Central Statistical Office, Hungary, edited by G. Dudás.

Domestic retailers–running more than 90% of the domestic store network in 2008–were challenged and selected by market restructuring and series of innovation that transformed the ways of sourcing and selling. Despite the growing demand, the number of domestic enterprises declined by one-third (over 56,000) in the 1997–2008 period.[10] The trend hit mostly individual entrepreneurs: the number of stores run by

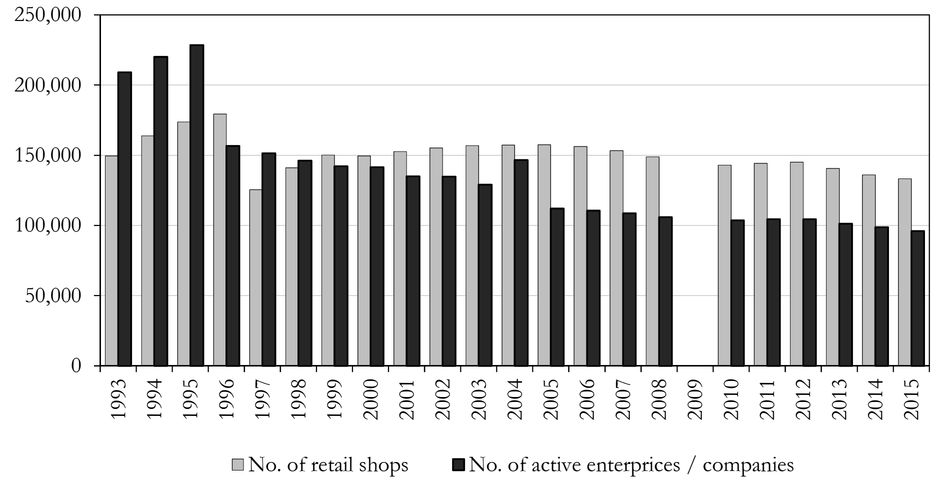
[9] For e.g. Tesco 'small hypers' and the Express shops; taking over the Tengelmann Plus network by Spar; launching the CBA 'Cent' and 'Prima' network; mushrooming of web shops, and related services in the whole sector.

[10] The year 1995 marked the peak of retail enterprising in Hungary, with over 230.000 registered enterprises in domestic market.

this group shrank by more than 18,000. Meanwhile, small- and medium-size companies that were still standing on the contested market were able to grow and expand by organising small store networks, joining buying groups, and investing in IT-systems. Thus, the centralization process that dominated the boom period manifested in terms of capital, organisation, and space, such as in the control taken by a shrinking number of corporate and partnership companies over the store network, in the constantly growing retail floor space. It also manifested in the rise and expansion of corporative organisations that aimed to gain control over market processes–while the first major market exits of TNCs occurred (e.g. Julius Meinl, Tengelmann/Kaiser's, Rewe/Billa) and the number of small family businesses was declining rapidly (Figure 2).

Figure 2

Number of shops and enterprises in retail sector (1993–2015)



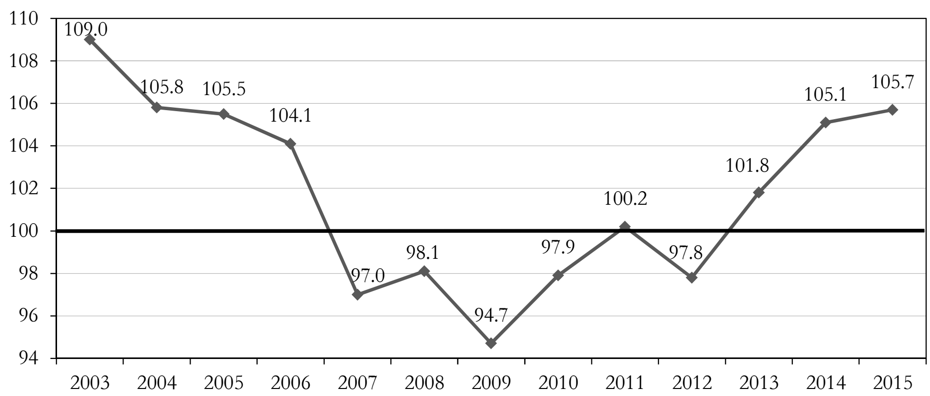
Source: Hungarian Central Statistical Office, Hungarian Statistical Yearbook, 1993–2015.
Note: Data from 2009 is missing.

Although, the above processes were perceived and problematised in professional discourses, the regulative framework remained inconsistent until 2005. In the discussed period, Hungarian governments acted as facilitators of the flows and operation of retail capital in a neoliberal manner, supporting foreign investments, liberalising trade, and delegating regulative issues to the local state. The regulation was overly bureaucratic and fragmented, and meanwhile, deficient on key issues, such as consumers' protection, market dominance, and suppliers' status. The new Act CLXIV of 2005 on Retail regulated consumer-retailer as well as supplier-retailer relations, defined the scope of state interventions, and set up a more or less consistent institutional framework for sectoral policy. Such steps reflected the priorities of the common (EU) economic policy particularly in product safety and fair market conditions. The rescaling of retail policies in the pre-crisis period–the EU safeguarding on market transactions and consumers' protection in new member states–and the national regulation produced complex and cost-raising conditions for daily operations, while the unfolding liberalisation of flows made the market increasingly contested. Thus, re-regulation contributed to the on-going centralisation processes in the domestic market[11]–favouring agents with substantial capital assets to exploit opportunities of European (global) sourcing and to adapt to changing regulations–that could not be counterbalanced by governmental (EU/national) funding that was provided for SMEs for small scale developments in the discussed period. Since regulations were focused on macro-level, the centralisation processes manifested in the rise of local and regional oligopolies/monopolies in peripheral spaces in the early 2000s.

The 2008/2013 crisis has changed the macro-context of retailing and consumption profoundly. The mechanisms of the financialised global economy drained household assets–the net financial assets of Hungarian households fell by 15% (1800 bn HUF) in 2008/2009 due to the revalorization of foreign currency-based debts (Gfk Hungary 2012)–and threw them into debt, while labour incomes were slightly declining/stagnating and the access to bank loans was strictly limited by tighter regulations and risk-reducing strategies of foreign shareholders of the banks across CEE (Bauer et al. 2013). The households responded to the crisis by paying debts, reducing consumption, and delaying investments, and it was evident in the 14% decline in retail turnover (2006–2013) and a slow consumption recovery from 2014 (Figure 3).

Figure 3

Changing level of consumption before, under, and during the crisis (2003–2015)(previous year=100)



Source: Hungarian Central Statistical Office, Hungarian Statistical Yearbook, 2003–2015.

[11] For example, in the food sector, market exits of retail TNCs rather reinforced centralization trends through chain/store takeovers, even though domestic capital benefited from the changes as the 'winners' (CBA, Coop) held oligopolistic position in the national market.

The scope of governmental policies in crisis management was limited by state debt and the austerity pressures–the consequences of earlier policy rescaling processes and the mechanisms of financialisation–the dependence of the national economy on global capital flows, and the unfolding social conflicts stemming from the debt crisis and economic decline. The enhanced regulative activity of the Hungarian central government from 2010 reflected such limits, and the strategy of the political elite to manage state debt, recover domestic consumption while getting control over market relationships–between various groups of cap forms of capital (bank/retail; national/domestic retail), consumers and retailers, and suppliers and retailers–and recast the regulatory competences within the state itself (local vs. central state). Economic growth and domestic capital accumulation was supported by the campaigns for enhancing the consumption of domestic products through bargaining processes (state/retailers) and the introduction of a complex administrative system for controlling the origin of goods; in addition, it was supported by stricter regulation of retailer-supplier relationships[12]. Such steps along with redefining market domination (2014), limiting the construction of large-scale retail facilities through strict (and non-transparent) central government control (2011), introduction of taxes (2011–2013), and new and expensive elements of administrative control[13] primarily hit the retail TNCs that–in government rhetoric–were supposed to improve the market positions of domestic retailers. Nevertheless, the enhanced state control and its costs (e.g. setting up the online cash register system), the introduction of new monopolies (e.g. in the tobacco trade), the sectoral ('crisis') tax, and the series of institutional reforms (e.g. in the cafeteria system) made the domestic market non-transparent and enhanced the costs of all retailers.

Although, the central state extended its control over flows of goods and raised tax revenues, re-regulation was a never-ending bargaining process between various groups of retailers and the state. The latter had to refashion regulations occasionally (in several cases, due to EU pressure) and make compromises–favouring capital interests at the expense of labour, as per the renewed Act on Labour in 2012, the poor (reducing state debt through cutting state transfers back), or other capital forms (producers, banks). Consequently, re-regulation selected domestic retailers rather than improving the conditions of capital accumulation for them, and this selectivity was spatially uneven due to the increased costs and polarising effects of re-regulation on business conditions in marginalised spaces.

Tighter regulation and declining demand challenged retailers in various ways. Powerful market agents responded by lobbying and bargaining for tax reduction and reducing administrative costs, and mobilising relational capital on the 'particularistic'

[12] Supporting domestic producers during the crisis period was not alien to core economies either; moreover, 'fair' retailer-supplier relations have also been in the focus of EU legislation since the last 10 years.

[13] For example, the costs of controlling the origin and trading of food; or the electronic system of following the flows of goods along with the HuGo that hit major retailers with centralised sourcing and logistics.

Hungarian market[14]; in addition, operating professional bodies facilitated transparent business conditions (OKSZ, Áfeosz). In the crisis period, corporate strategies were also re-fashioned by focusing on:

– Exploiting economies of scale by extending store networks, e.g. by franchising (Áfész, Spar);
– Economies of scope by diversifying into new retail segments (Tesco Express, Spar City, CBA Cent/Príma, etc.), combining traditional forms with enhanced e-tailing (omni-channel retailing), and introducing a highly structured offer targeting various consumer groups;
– Stronger spatial embedding–while bargaining with suppliers to share rising regulation-related costs (e.g. in quality control);
– Carefully designated development focusing on prime locations; in addition, less investment in construction and more in organisation, IT, and innovation.

Retail restructuring was also shaped by market exits involving powerful agents in all segments (Delhaize, Tengelmann, Bricostore, Electroworld, etc.), and a high number of small-scale businesses (over 10.000), mainly, in the daily goods' (FMCG) market (Figure 2). This shrinkage was associated with declining investments in the sector (−20%) and a falling share of FDI (from 49.4 to 34.6%) until 2013 (Figure 1). The performance of the sector improved from 2014 due to a slight improvement in demand and the adaptive strategies of retailers in the re-regulated market. However, likely, the growth in turnover was also generated by enhanced information on retail transactions (the introduction of the online cash register systems), which should be researched in-depth to understand the drivers of retail dynamics.

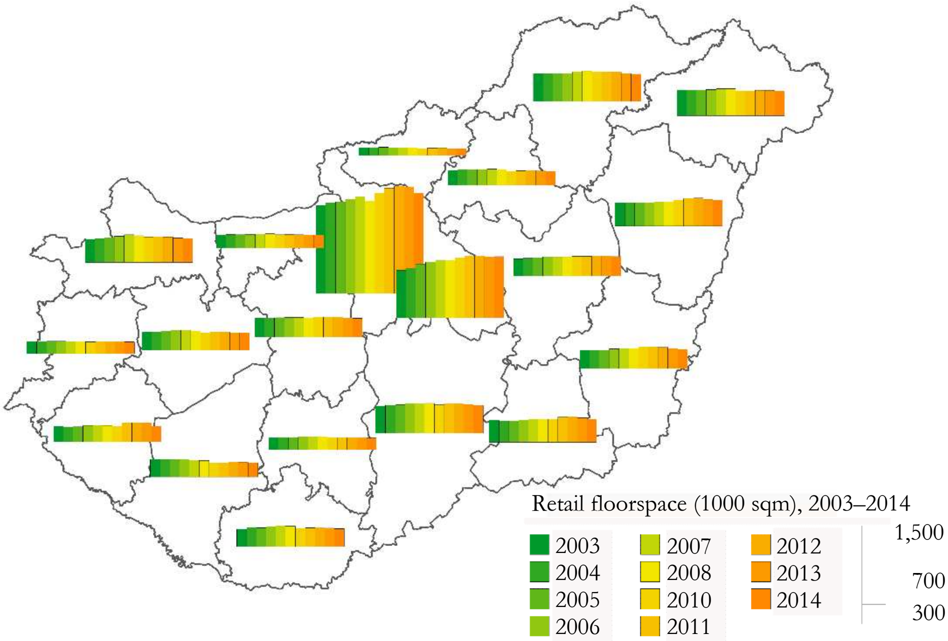
The crisis and regulative changes accelerated centralization processes and the polarisation of consumption spaces. The store network had been shrinking by 25% between 2006 and 2015, and this stock was run by a shrinking number of enterprises; the decline was fuelled not only by the failures of (over 13.000) small family businesses but also by closing down of stores run by small companies (over 2.600 shops) including local and regional store chains (mostly the ones running 2–5 shops). Capital and organisational concentration resulted in a centralisation of consumption spaces: the gross retail floor space fell to the 2005-level by 2015, while the average store space rose from 103.6m^2 to 124.4m^2 (Figure 4). New openings and takeovers that followed market exits by major retailers were spatially selective and concentrated mostly in urban spaces of high value, following the spatial logic of retail capital to minimise potential sunk costs. The process was reflected by an increasing focus on metropolitan areas, the revival of main shopping streets, and growing concentration of foreign retailers in malls and hypermarkets (Colliers International 2014, KSH

[14] Particularistic markets are characterised by weak state mechanisms to regulate market processes in a consistent and transparent way as various groups' interests, e.g. in CEE and Latin America (Hess 2004), overwrite the conditions constantly.

2008). It was also reflected through the abandonment of stores in peripheral rural spaces, including their urban centres (Picture 1), and the growing urban-rural differences in the densities of retail and related services across Hungary, as it is discussed in the following chapter in detail.

Figure 4

The spatial distribution of retail floor space at NUTS3 level



Source: Central Statistical Office, Hungary, edited by G. Dudás.

Picture 1

Closed supermarket in a peripheral region (Mezőkovácsházai járás, 2014)

Source: Erika Nagy's photography.

Socio-spatial polarisation and marginalisation in consumption:

Changing regional patterns, networks, and hierarchies

The pre-crisis period was characterised by a strong metropolitan focus (Budapest and its suburbs) of retailers' location strategies and capital flows (Nagy 2005). This spatial pattern recovered in the post-crisis time reflecting the security-seeking policy of investors for the crisis-hit CEE markets (Nagy 2009b). Retail capital targeted non-metropolitan areas between the years 2004 and 2008 in the increasingly relaxed institutional-regulative context–a contested market reflected by the declining yields of the commercial property markets in CEE metropolises (Colliers International 2007). The landscape of retail investments was highly uneven outside the Budapest region and it was shaped largely by major developments (malls, hypermarkets, DIY and discount stores, logistic centres, etc.). Northwest Hungary received major capital influx earlier, while developments in East and South Hungary took place a few years later, ending in the initial years of the crisis. It was only the more well off north-west region that recovered its status in capital flows soon after the crisis in non-metropolitan Hungary. The spatial distribution of investments had no strong correlation either with purchase power or with other socio-economic indicators at NUTS3 scale. The investments' spatiality (Figure 1.) reflected corporate location strategies that followed rather hierarchical than clear regional patterns and were shaped by local as well as by intra-firm factors (Nagy 2010).

The investment activity at county and local scale depended largely on shopping mall and hypermarket developments. Hungary entered the first phase of mall projects during the 1995–2000 period (40 new units). The primary targets were the capital city, emerging edge-cities around Budapest, and the largest regional centres. In the second phase (2001–2009), 73 new projects were completed, focussing on primary and secondary urban centres outside the Budapest region. The declining share of Budapest reflected the decentralization process: in 2000, over half of the malls were located in the Budapest regions, and its share fell to 40% by 2009. The process was stopped by the crisis and, lately, by the introduction of tighter national regulations on land use. The recent recovery was focussed on the Budapest metropolitan area, reflecting the trends of the highly contested CEE markets (Nagy 2014a).

The hypermarket network was expanding as dynamically as the malls (the number of stores rose from 44 to 170 between 2001 and 2010). The competing international chains (Tesco, Spar, Auchan, Cora) and domestic actors were seeking to get a full coverage of the market in the heyday of retail investments. The global crisis ended the expansion resulting in market exits (Delhaize/Cora). However, it has not brought major spatial restructurings in the segment. Recently, malls and hypermarkets are operating in almost one-third of the urban centres in Hungary[15], and the Budapest metropolitan area still holds its share (about 30% of the stores). The slowdown reflects the shift in corporate strategies towards security, efficiency, and organisational diversity.

The above trends reflect an on-going centralisation of retail and consumption spaces. The depth of socio-spatial polarisation and the marginalisation of various consumer groups can be tracked clearly at local (LAU2) scale. A constant and overall decrease could be observed in the number of stores in all the municipalities with non-urban status[16]: the number of retail shops per municipality fell from 15.4 (1994) to 8,0 (2013). Smaller places were affected more heavily; while the number of retail stores was increasing at macro-scale until 1999, municipalities with less than 2000 inhabitants[17] had not been targeted by any large scale developments and the number of shops per municipality declined from 8 (1994) to 5 (2014).

The shift towards an unbalanced spatial structure of consumption spaces was reflected by the increasing number of localities with no retail services as well as by emerging monopolistic and oligopolistic (2–5 shops[18]) local markets. Today, this group of settlements has a very limited supply of locally available goods, high retail prices hitting immobile social groups, and higher transportation and time costs affecting the mobile groups. The number of villages without any shops increased from 40 (2003) to 318 (2015); such places were concentrated in South- and West-Transdanubia and North-Hungary–overlapping with rural areas dominated by small villages (less than 1000 inhabitants). As a result of the constant decline, the share of such municipalities with no services rose from 42% to 51% in the discussed period. The process led to i) a spatially more compact areas of marginalised markets; ii) an increasing share of places with no or a single retail shop; iii) and a growing number of marginalised local markets in regions with higher share of urban population (Great Plain) or with higher-income level (Northwest-Hungary).

The food market was affected by similar trends and spatial patterns, apparently marginalising consumers through their daily practices in the discussed regions. Additionally, the villages around the capital city and in the Great Plain were also 'vacated'–and emerged as scenes to emerging oligopolistic markets–by the concentration of daily goods' retail floor space in major cities (Figure 5). The spatial restructuring of the non-food retail market had a similar yet more concentrated pattern that was polarised further during the crisis (2006–2014).

[15] At the beginning of 2016, there were 347 settlements with township legal status.

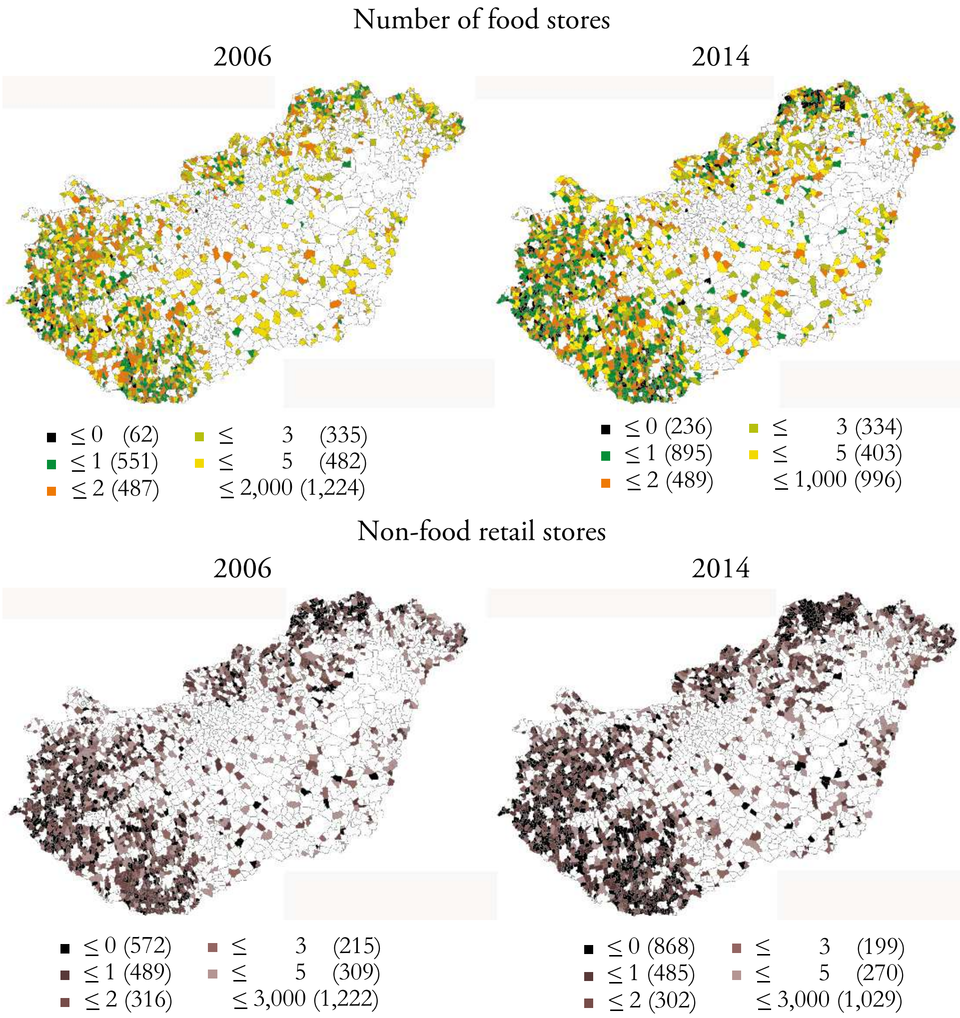
[16] The number of villages (by legal status) was 2920 in 1994 and 2809 in 2015. Many former villages got the legal status of 'town' in the investigated period; additionally, new independent municipalities were also established.

[17] The overall number of small settlements under 2000 inhabitants was shifting between 2347 (1999) and 2408 (2015), depending on the actual number of inhabitants in certain LAU2 units.

[18] Including several food shops, at least one providing clothing (often second-hand), a shop for farmers, and several general stores.

Figure 5 a); b); c) and d)

Spatial distribution of retail 'deserts', monopolistic, and oligopolistic markets in food and non-food sectors in 2006 and 2014



Source: Based on data of Hungarian Central Statistical Office County Yearbooks, 2006; 2014., edited by: G. Dudás.

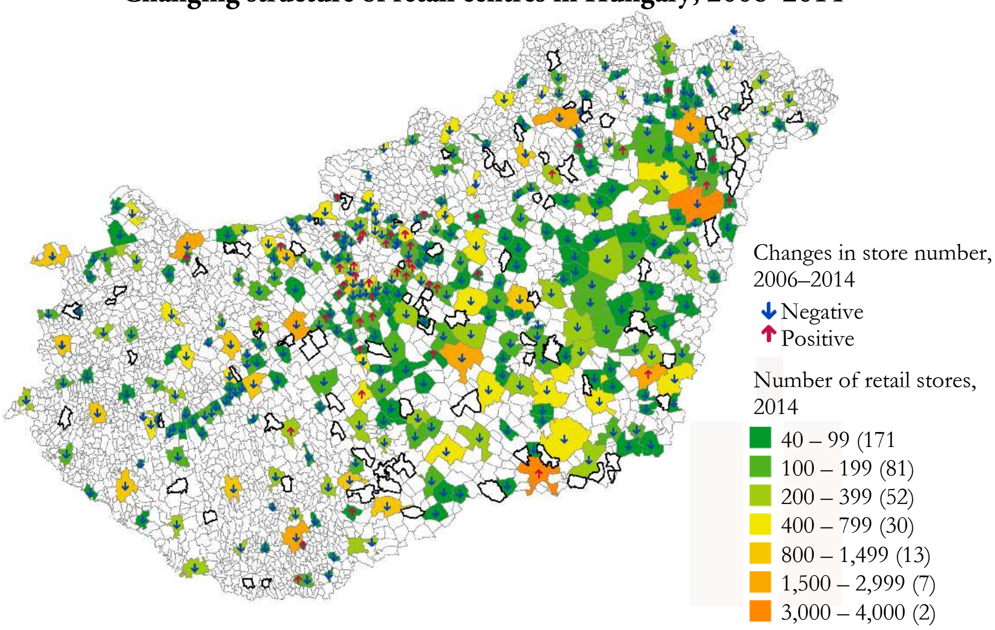
The structural changes in the retail sector redrew spatial hierarchies in consumption. In our earlier research (relying on data from 1996), we defined 40 shops per municipality as a threshold for being a central place in terms of scale and diversity (Nagy 1999).[19] We repeated the research at the peak of retail rush (2006) – keeping

[19] We refer here to Christaller's term. (Christaller 1966).

the same method and threshold for obtaining comparable results – and we identified 436 retail centres. (Nagy–Nagy 2008) However, by the end of the crisis, we found only 355 centres (dominantly, with town status) remaining. The changes within the group were substantial: there were 'newcomers' (mainly near major county towns and 10 municipalities near Budapest) benefiting from suburbanisation processes. The moderate decrease in the number of stores did not change the central role of major and medium-size county towns. Meanwhile, one-fifth of the former (2005) places were missing from the list, and the majority of smaller centres lost shops both in the food and non-food sector, which weakened their central functions. This process, along with the increasing number and spread of marginalised retail spaces, raised a major problem; the declining small retail centres were supposed to provide daily and non-daily goods for the surrounding, increasingly marginalised localities, and consumers. Thus, the scenario inevitably led to further polarisation.

Figure 6

Changing structure of retail centres in Hungary, 2006–2014



Areas bounded by bold black lines indicate the municipalities that did not reach the 40-shop threshold, and thus failed to get the consumption centre status by our definition.

Source: Based on data of Hungarian Central Statistical Office County Yearbooks, 2006; 2014., edited by: G. Dudás.

The differentiation of the retail network hit various regions differently, such as Transdanubian regions with high share of small villages–where the network had been scarce even before the crisis–and North Hungary as a whole sufferd from a thinning of consumption centres.

Explaining spatial patterns: Consumption, incomes, and poverty

The growth of consumption fuelled by rising incomes and later by bank loans stopped before the crisis and new trends emerged including a renewed focus on household consumption strategies, such as channelling declining incomes primarily into basic goods, utility fees, and paying debts (Nagy et al. 2015a). Although macro-processes largely shaped such decisions and local processes, substantial regional differences emerged in consumption patterns in the crisis period. In the pre- and the post-crisis period, the growth rate of consumption was much higher in the capital city and West Transdanubia, far exceeding the national average. In the prosperous years of the national economy, the Pest county and Central Transdanubia grew just as fast as the leading regions, but the quarterly data after 2014 suggests a widening gap between them and the two leading regions. The growth in the above regions was rooted in the increasing number of workplaces, growing wages in the private and the public sector, and in the high level of retail investments. Moreover, the average level of pensions in these regions was also higher than in the rest of Hungary that stabilised the demand.

The rate of growth in consumption was 2–3% lower in the East and South of Hungary than the national average, before and during the crisis. Moreover, while the rate of growth increased in the dynamic regions in 2015, it was slowing down in the lagging ones, generating a growing gap. Regional differences were also seen in the persisting structures of consumption in which food and utility expenses had much higher share in the East and South than in the well-off regions–while the families in the former areas spent significantly less on communication, travel, health, culture, and leisure. By adding the shrinkage of public services to these trends, we could see a much deeper gap between East-South and Northwest-Central regions of Hungary (Nagy et al. 2015b).

The post-crisis recovery seems to have deepened socio-spatial differences. In 2016, the growth of consumption in the capital city was four times higher than the national average (5%), followed by West-Transdanubia and Pest County (7–8%). Meanwhile, all other regions were still declining (by 2−5%). The growth of household incomes does not explain such a major difference in consumption as there was an increase in each NUTS2 region. The index of annual net income grew (at comparative prices) by 3–3% in 2013 and 2014 and by 4% in 2015 at macro-scale. However, spatial inequalities increased during the 2013–2015 period: in the eastern and southern regions of Hungary, the majority of counties (NUTS3 units) had no significant income growth, while in the dynamic regions' growth was above the national average. Such trends were fuelled not only by the uneven distribution of capital investments but also by hollowing out the public social care system and the resulting shrinkage in social transfers. The socio-spatial differentiation of incomes has a major explanatory power in understanding the polarisation of retail services: the areas affected mostly by the pre- and post-crisis restructuring are the poorest regions and highly dependent on state transfers, where marginality hits not only rural spaces but it also 'filters up' in the settlement hierarchy, hitting a number of small towns.

The above processes are reflected by income poverty data based on the Central Statistical Office 2005 Microcensus and the 2011 National Census. In the crisis-period, the share of the LAU1 regions hit by poverty[20] rose from 15% to 20% between 2005 and 2011, and further growth is being produced by recent austerity schemes of the government. Meanwhile, poverty grew as a spatially contagious phenomenon in East and South Hungary – and hit more LAU1 regions even in the West by 2011. The small-regions with the lowest share (<10%) of families hit by poverty concentrated in Central and Northwest Hungary in both periods; nevertheless, well off areas were shrinking and concentrated increasingly around Budapest and along the Budapest-Vienna and the Budapest-Balaton axes. These patterns mirror the uneven restructuring of the retail sector and the consumption spaces during crisis and in the following recovery period.

Conclusions

In the last two decades, retail restructuring produced an increasingly uneven landscape of consumption across CEE, contributing to socio-spatial polarisation in terms of access to goods and services to consumption practices and identities (Boros et al. 2007, Boros–Pál 2006, Eglitis 2011, Nagy 2014a, Patico–Caldwell 2002). Although the key agents – powerful retailers and the state – of such processes can be identified clearly at the national and local level, their relationships and embedding in various flows and networks shaping their strategies are highly complex. We sought to unfold this complexity in this paper:

(i) The spatial logic – the multiple territorial embedding–of retail capital was the driver of the centralisation processes in terms capital, organization, and consumption spaces in the pre-crisis neoliberal regulative context, enhancing urban-rural and metropolitan/non-metropolitan and intra-urban inequalities. The crisis made retailers focus on protecting their fixed assets and employing a highly selective location policy that deepened urban/rural gap, differentiated the urban network, and put metropolitan and intra-urban central spaces as prime targets for investments. The retailers' spatial strategies should be understood as the manifestations of the responses to the crisis of the financialised CEE markets (declining demand, indebtedness, flight of capital from peripheral markets) that differentiated consumers not only by income (Németh–Kiss 2007) but also by place of residence across Hungary. As a result, local and regional monopolies were produced/reinforced, and the population

[20] LAU1 regions where more than one-quarter of households lived in poverty.

of peripheral rural regions were marginalised, which made local households adopt new consumption practices, embrace informal activities, and accept new dependencies (e.g. the local state and personal networks controlled by powerful local agents) (Nagy 2014b).

(ii) Despite its limited scope, the state – primarily, the central government – adopted an interventionist policy in the crisis period. It rested on the stimulation of the domestic demand (reducing public utility costs and personal income taxes) that was socially selective due to the parallel steps towards shrinking and centralising social transfers and public services. As a result, deserving and undeserving consumer groups emerged, which led to 'double marginality' in peripheral spaces – produced by the spatial logic of retail capital and by state policies. Additionally, the introduction of tighter and increasingly complex regulation of retail – that did not deal with its socio-spatial consequences–stimulated market trends towards centralisation by (a) re-scaling regulative power from local to national level (on retail development, social transfers/local demand, and collective consumption), (b) rebalancing state-capital relations (enhanced administrative control over flows), (c) changing power relations between international and domestic capital, and (d) recasting state-capital-society nexus by shifting capital to finance state debt through taxation (VAT, corporate taxes) and administration. The limited transparency of re-regulation and the rising administration-related costs of operation hit mostly the small businesses in peripheral spaces–contributing to uneven development in Hungary.

The above processes were embedded into the systemic crisis of neoliberal capitalism (Harvey 2010) – the collapse of consumption growth rested largely on debts that hit CEE economies and their marginalised spaces heavily – and raised responses that enhanced the state control, caused a shift towards authoritarian schemes, and strengthened executive power, particularly in peripheral economies (Jessop 2010; Szivós–Tóth, 2013). However, the practices of policy making and the responses of local societies are shaped by national and local contexts that need further – in-depth – empirical research.

REFERENCES

AALBERS, M. (2009): Geographies of the financial crisis *Area* 41 (1): 34–42.

BAUD, C.–DURAND, C. (2012): Financialization, globalization, and the making of profits by leading retailers *Socio-Economic Review* 10 (2): 241–266.

BAUER, P.–ENDRÉSZ, M.–KISS, R.–KOVALSZKY, ZS.–MARTONOSI, Á.–RÁCZ, O.–SCHINDLER, I. (2013): Túlzott lakossági eladósodás: okok, trendek, következmények *MNB Szemle*, Különszám, 2013. október. https://www.mnb.hu/letoltes/bauer-2.pdf (downloaded: December 2016)

BOROS, L.–HEGEDŰS, G.–PÁL, V. (2007): A globalizáció néhány hatása Szeged kereskedelmére In: Sikos T. (ed.): *A bevásárlóközpontok jelene és jövője* pp. 79–103., Selye János Egyetem Kutatóintézete, Révkomárom.

BOROS, L.–PÁL, V. (2006): The effect of a hypermarket on the retail of Szeged *Geographica Timiensis* 15 (1-2): 79–86.

BOURDIEU, P. (1984): *Distinction: A Social Critique of the Judgement of Taste* Harvard University Press, Cambridge MA

CHRISTALLER, W. (1966): *Central Places in Southern Germany* Prentice Hall, New Jersey.

CLARKE, D. B. (1996): The limits to retail capital In: Wrigley, N.–Lowe, M. (eds.): *Retailing, consumption and capital* pp. 284-301., Longman, Harlow.

COE, N.–HESS, M. (2005): The internationalization of retailing: implications for supply network restructuring in East Asia and Eastern Europe *Journal of Economic Geography* 5 (4): 449–473.

COLLIERS INTERNATIONAL (2007): *Colliers Real Estate Review. Central and Eastern Europe* www.colliers.com (downloaded: December 2010)

COLLIERS INTERNATIONAL (2014): The changing face of retail. Eastern Europe. http://www.colliers.com/-/media/files/emea/easterneuropeaninformation/2014eeresearchpage/retail/ee-2014-the-changing-face-of-retail-final.pdf?la=en-gb (downloaded: December 2016)

DICKEN, P. (2007): *The Global Shift. Mapping the Changing Contours of the World Economy* Sage, London, UK.

EGLITIS, D. (2011): Class, culture, and consumption: representations of stratification in post-communist Latvia *Cultural Sociology* 5 (3): 423–446.

EHRLICH, K.–KRISZAN, A.–LANG, T. (2012): Urban development in Central and Eastern Europe – between peripheralization and centralization? *disP - The Planning Review* 48 (2): 77–92.

ESPON (2014): *ECR2 Economic Crisis: Resilience of Regions* ESPON and Cardiff University, Luxembourg.

GFK HUNGARY (2012) *Gfk Vásárlóerő Magyarország* Gfk Hungary, Budapest

GOSS, J. (2006): Geographies of consumption: the work of consumption *Progress in Human Geography* 30 (2): 237–249.

HADJIMICHALIS, C. (2011): Uneven geographical development and socio-spatial justice and solidarity: European regions after the 2009 financial crisis *European Urban and Regional Studies* 18 (3): 254–274.

HARVEY, D. (2010): *The enigma of capital and the end of capitalism* Oxford University Press, New York, US.

HARVEY, D. (2011): Roepke Lecture in Economic Geography–Crises, Geographic Disruptions and the Uneven Development of Political Responses *Economic Geography* 87 (1): 1–22.

HESS, M. (2004): Global production networks. Dealing with diversity In: HAAK, R.–TACHIKI, D. S. (eds.): *Regional Strategies in a Global Economy* pp. 31–52., IUDICIUM Verlag, Munich, Germany.

HETESI, E.–ANDICS, J.–VERES, Z. (2007): Életstíluskutatási eredmények fogyasztásszociológiai interpretációs dilemmái *Szociológiai Szemle* 17 (3-4): 115–134.

http://www.oecd-ilibrary.org/economics/how-s-life-2015_how_life-2015-en (downloaded: December 2016).

http://www.tarki.hu/hu/research/gazdkult/fogyasztoi_szegmentacio.pdf (downloaded: 12/2016).

HUDSON, R.–HADJIMICHALIS, C. (2014): Contemporary crisis across Europe and the crisis of regional development theories *Regional Studies* 48 (1): 208–218.

JAMESON, F. (1974): *Marxism and Form* Princeton University Press, Princeton, NJ.

JESSOP, B. (2004): Developments in Marxist Theory In: NASH, K.–SCOTT, A. (eds): *The Blackwell Companion to Political Sociology* pp. 7–17., Blackwell, Oxford.

JESSOP, B. (2010): The 'return' of the national state in the current crisis of the world market *Capital and Class* 34 (1): 38–43.

KELLER T. (2008): *Fogyasztói szegmentáció* TÁRKI, Budapest.

KSH (2008): Bevásárlóközpontok és hipermarketek *Statisztikai Tükör* 2008/144 http://www.ksh.hu/docs/hun/xftp/idoszaki/bevkp/bevkp08.pdf (downloaded: 12/2016).

LANG, T. (2015): *Socio-economic and political responses to regional polarisation and socio-spatial peripheralisation in Central and Eastern Europe: a research agenda Hungarian Geographical Bulletin* 64 (3): 171–185.

MILLER, D. (1999): *A theory of shopping* Polity Press, Cambridge.

NAGY E. (2014a): Bargaining, learning and control: Production of consumption spaces in post-socialist context *Europa Regional* 20 (1): 42–54.

NAGY, E. (1999): A kereskedelmi ingatlanfejlesztések néhány sajátossága és térszerkezet-formáló szerepe a kelet-közép-európai városokban *Alföldi Tanulmányok* 17 (1): 13–29.

NAGY, E. (2005):_Retail restructuring and emerging spatial patterns of consumption: new aspects of development disparities in Hungary In: BARTA, GY.–G. FEKETE, É.–SZÖRÉNYINÉ KUKORELLI, I.–TIMÁR, J. (eds.): *Hungarian Spaces and Places: Patterns of Transition* pp. 161–181., Centre for Regional Studies of the Hungarian Academy of Sciences, Pécs.

NAGY, E. (2009a): Internal trade In: KOCSIS, K.–SCHEWEITZER, F. (eds): *Hungary in Maps* pp. 182–187., Geographical Research Institute of the Hungarian Academy of Sciences, Budapest.

NAGY, E. (2009b):_Regulation and planning of retail spaces on the emerging markets: neoliberal schemes for urban restructuring? In: PALÓCZ, É.–SZALAY, L. (eds.): *Public and Private Services in the New Global Economy: XIX. International Conference of RESER* pp. 1–17., TÁRKI, Budapest.

NAGY, E. (2010): Változó erőviszonyok - átalakuló terek: A kereskedelmi tőke térszerkezet-alakító szerepe az élelmiszer-kiskereskedelemben In: SIKOS, T. T. (ed.): *Fenntartható fogyasztás és növekedés határai: új trendek a kereskedelemben* pp. 177–194., Selye János Egyetem Gazdasági Kar, Szent István Egyetem Gazdaság- és Társadalom-tudományi Kar, Révkomárom.

NAGY, E. (2014b): Függőség és egyenlőtlen fejlődés a vidéki terekben – az élelmiszer-ágazat változó erőviszonyai és lokális beágyazottsága In: Nagy, E.–Nagy, G. (eds.): *Polarizáció – függőség – krízis: Eltérő térbeli válaszok* pp. 78–88., MTA KRTK RKI Alföldi Tudományos Osztály, Békéscsaba.

NAGY, E.–TIMÁR, J.–NAGY, G.–VELKEY, G. (2015a): The Everyday Practices of the Reproduction of Peripherality and Marginality in Hungary. In: LANG, T.–HENN, S.–SGIBNEV, W.–EHRLICH, K. (eds.): *Understanding Geographies of Polarization and Peripheralization: Perspectives from Central and Eastern Europe and Beyond* pp. 135–155., Palgrave Macmillan, Basingstoke.

NAGY, E.–TIMÁR, J.–NAGY, G.–VELKEY, G. (2015b): A társadalmi-térbeli marginalizáció folyamatai a leszakadó vidéki térségekben *Tér és Társadalom* 29 (1): 35–52.

NAGY, G.–NAGY, E. (2008): A városok gazdasági potenciálja *Falu Város Régió* 15 (3): 32–42.

NÉMETH, N.–KISS, J. P. (2007): Megyéink és kistérségeink belső jövedelmi tagoltsága *Területi Statisztika* 47 (1): 20–45.

OECD (2015): *How's Life? Measuring Well-being. 2015.* OECD iLibrary.

PALÓCZ, É.–SZALAY, L. (eds.) (2009): *Public and Private Services in the New Global Economy: XIX. International Conference of RESER* TÁRKI, Budapest.

PATICO, J.–CALDWELL, M. L. (2002): Consumers exiting socialism: ethnographic perspectives on daily life in post-communist Europe *Ethnos* 67 (3): 285–294.

SMITH, N. (1991): *Uneven Development. Nature, Capital and the Production of Space* 2. ed. Blackwell, Oxford, UK.

STENNING, A.–SMITH, A.–ROCHOVSKA, A.–SWIATEK, D. (2011): *Domesticating Neoliberalism. Spaces of Economic Practice and Social Reproduction in Post-socialist Cities* Wiley-Blackwell, Oxford, UK.

SZALAI, E. (2006): *Az újkapitalizmus és ami utána jöhet* Új Mandátum Kiadó, Budapest

SZIVÓS, P.–TÓTH, I. GY. (szerk.) (2013): *Egyenlőtlenég és polarizálódás a magyar társadalomban* TÁRKI Budapest

WARD, T.–LELKES, O.–SUTHERLAND, H.–TÓTH, I. GY. (2009): *European Inequalities: Social Inclusion and Income Distribution in the European Union* TÁRKI Social Research Institute Inc., Budapest.

WRIGLEY, N. (1996): Sunk costs and corporate restructuring: British food retailing and the property crisis In: Wrigley, N.–Lowe, M. (eds.): *Retailing, consumption and capital* pp. 116-136., Longman, Harlow.

WRIGLEY, N.–COE, N. M.–CURRAH, A. (2005): Globalizing retail: conceptualizing the distribution-based transnational corporation (TNC) *Progress in Urban Geography* 29 (4): 437–457.

ZUKIN, S.–MAGUIRE, J. S. (2004): Consumers and consumption *Annual Review of Sociology* 30: 173–197.

VISUALIZATIONS

Brand wars in cyberspace: a GIS solution

Ákos Jakobi
Eötvös Loránd University,
Budapest, Hungary
E-mail:
jakobi@caesar.elte.hu

Hajnalka Lőcsei
Eötvös Loránd University,
Budapest, Hungary
E-mail:
locsei@caesar.elte.hu

In the first sight, it seems to be obvious that Nike, the well-known US sports equipment company, is predominant in Great Britain; however, the European (German) giant, Adidas also happens to be very popular in the country. The duel between the global first and second sportswear companies can be traced back to not just their sales numbers in a bid to gain larger market share, but also to soft factors, such as reputation. The only problem is that reputation can be at best measured by well-organized questionnaire data of population samples, which are sometimes difficult to administer. However, this issue can be also solved by capturing the presence of indirect knowledge about the brands, for example, from the geographically localizable content of cyberspace. The indirect application of digital footprints in scientific analyses has notable examples (Girardin et al. 2008; 2009; Järv 2012), in which the conclusions are based on the examination of digital data that have been generated as by-products of socio-economic phenomena.

Despite the fact that content on the Internet is placeless, there still lies the possibility of identifying web content geographically. The crawler and data mining methodology of collecting a large amount of content data that could be connected to addresses with known geographical position allows us to map the spatial occurrence of certain notions, phrases, words, or brand names, among others. Completely new aspects of regional inequalities of the information age appear because of analysing geocoded web content. The queries of keywords reflect that geocoded web content is not spatially random but follows certain geographical characteristics of the society. Contrarily, visual interpretations of results may reveal new inequality patterns and serve as an evidence of presumed, but not yet tested assumptions.

Regional Statistics, Vol 6, No 2. 2016: 173–176; DOI: 10.15196/RS06209

Our following example represents results of finding geolocated web contents in relation to the keywords 'Nike' and 'Adidas'. In our analysis, the postal addresses found on the websites have supported the determination of the spatial relevance of the web content. It was assumed that there is a good chance of the provided postal address to be connected with the published content. For example, if a company advertises its goods or services on a website, then we might also find the company's address on the page, and thereby interpret the content as attached to the company's geographical location. Certainly, the absence of address information on a website will not allow for geolocating the provided content, and the content would be attached to multiple geographical locations when multiple addresses are found.

The analysis was based on the 'G-Search' technology provided by ESRI Hungary Ltd., which applied a map-based crawler engine to collect and systematize web content automatically. Starting from an initial website database, the engine read all the found addresses and web links into a temporary database, from where – following the newly found links – further searches were made to collect as much data as possible. By the application of intelligent algorithms, the relevant texts were saved as an attachment to geographical addresses. This final database made it possible for us to run queries of keywords in order to find and determine the geographical location of the selected web content.

To explore the penetration of the brands Nike and Adidas in cyberspace, we queried the keywords in the collected UK-related database (since our postal address database was available for mainland UK, our analysis focussed only on that region), and then we mapped the location of the discovered addresses. The result maps demonstrated the spatiality of the published web content, namely the geographically identifiable places of web pages mentioning the selected keywords. Consequently, they are not (or not necessarily) figures denoting the location of shops selling those goods, but the figures of the location of cyberplaces (Batty 1997) mentioning them.

Concerning the first set of results obtained, we noticed that the number of geolocated websites referring to the keyword Nike was much larger and the density of address points was more consistent than the queries of the keyword Adidas (see top left and top right maps). The overall picture was obvious: Nike has won the brand war in the UK (if just those two companies are counted). However, it is also assumed that some localities could have the opposite outcomes, namely Adidas might overcome Nike in some areas of the country. Therefore, and in order to get a more detailed picture, we created a 10×10 km grid network over mainland UK and compared in each grid cells the number of Nike and Adidas hits to decide the local dominance proportions (see also Jakobi 2015). The new map reflected the geography of the Nike-Adidas duel in the UK cyberspace by depicting the locations where Nike has small or great advantage over Adidas, and the places where Adidas has higher penetration rates in cyberspace than Nike. Apparently, there are cells with equal counts of Adidas and Nike hits; in addition, the absence of coloured cells is observable at places where none of the keywords has been found. It is also interesting that the map of the cyberspace duel is not spatially random but follows geographical characteristics. The density of hits is apparently higher in the urban areas, while smaller in the less populated peripheries. However, it is also evident that the dominance of Nike related websites is stronger in main urban agglomerations, e.g. around London, the Midlands, Manchester, or Liverpool, among others. The dominance of Adidas related websites are only noticeable in some cells located at the peripheries surrounding the urban areas or in Cornwall. It seems that Adidas has a chance to overcome Nike in cyberspace only when it is about a small number of local hits.

REFERENCES

BATTY, M. (1997): Virtual Geography *Futures* 29 (4-5): 337–352.

GIRARDIN, F.–CALABRESE, F.–FIORE, F. D.–RATTI, C.–BLAT, J. (2008): Digital Footprinting: Uncovering Tourists with User-Generated Content *Pervasive Computing, IEEE* 7 (4): 36–43.

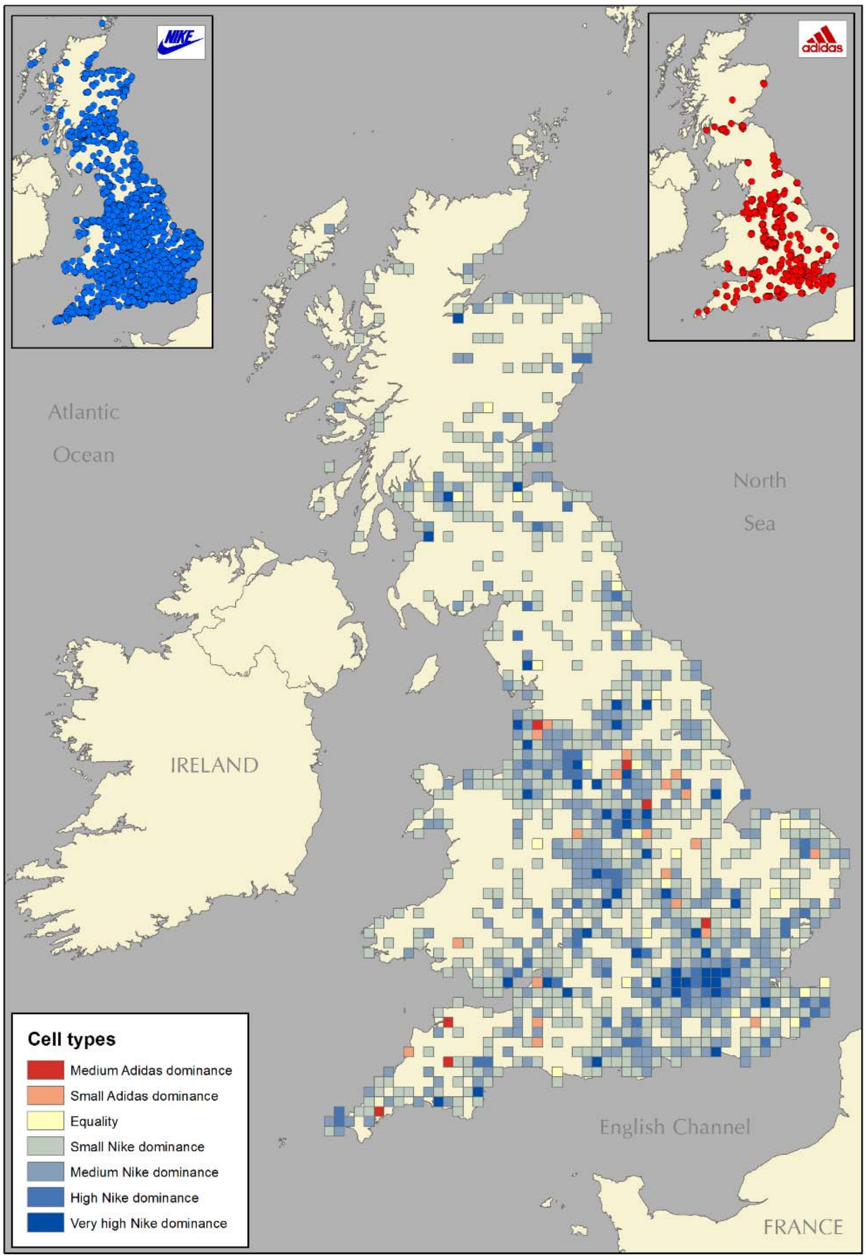
GIRARDIN, F.–VACCARI, A.–GERBER, A.–BIDERMAN, A.–RATTI, C. (2009): Quantifying urban attractiveness from the distribution and density of digital footprints *International Journal of Spatial Data Infrastructures Research* 4: 175–200.

JAKOBI, Á. (2015): A grid: aggregált és dezaggregált rácsmodellek a területi egyenlőtlenségek vizsgálatában *Területi Statisztika* 55 (4):322–338.

JÄRV, O.–AHAS, R.–SALUVEER, E.–DERUDDER, B.–WITLOX, F. (2012): Mobile Phones in a Traffic Flow: A Geographical Perspective to Evening Rush Hour Traffic Analysis Using Call Detail Records *PLoS ONE* 7 (11): e49171.

Software: ArcGIS

Local dominance of Nike and Adidas related websites in 10×10 km grid cells in mainland UK (centre), and the density of geolocated websites referring to Nike (top left) and Adidas keywords (top right)



Source: Own construction of the authors, base map from DIVA-GIS.

The world's economic centre of gravity

Géza Tóth
Institute of World and Regional Economics, University of Miskolc, Miskolc, Hungary
E-mail: geza.toth@ksh.hu

Zoltán Nagy
Institute of World and Regional Economics, University of Miskolc, Miskolc, Hungary
E-mail: nagy.zoltan@uni-miskolc.hu

The world's economic centre of gravity

Recently, an analysis has been conducted on the world's economic centre of gravity (Economist 2012, Grether-Mathys 2010, Quah 2011). Our aim is to examine all the countries of the world's economic centre of gravity together, while examining the continents separately. In our study, the geometric centre of the area for all countries is weighted by nominal GDP obtained from the World Bank database for the years 1970–2014 (GDP data at current US $ market prices). Although this method can be problematic, particularly for wide-area countries, it can address the most important trends of spatial displacements. Here, we used the official map of Eurostat as a base map with ETRS89 projection.

During the study, analyses on all the countries were carried out, while the continents were analysed separately (America, Africa, Asia [including Australia and Oceania], and Europe). In our model, Russia was considered as a part of Europe. In the model, changes in the country's borders related to influencing factors including the interpretation of the illustration on the gravitational movement trends were examined, and not just the fluctuations from one year to another. We illustrated each year for reasons of visibility, in the global study; however, in the case of continents, only a period of six years was taken into account.

Globally, the economic centre of gravity moved from the northwest to the southeast, essentially without a significant overshoot. The movement from the northwest part of Spain towards Tunisia indicates that the Northern hemisphere is still dominant, but it is steadily weakening against the Southern hemisphere. In 1970, the centre of gravity was also located in the Western Hemisphere. It shows the North American and European superiority, which had clearly moved to the Eastern hemisphere after 2008, thereby indicating a loss of position in the global economy of these regions.

In *Africa*, the centre of gravity was moving essentially close to the geometric centre of the continent during the analysed period, primarily in the Central African Republic, the Republic of Congo, and the Democratic Republic of Congo. In 1970, the centre of gravity was located in the southern part of this region. Although we can see a very small movement between the continents, in 2014, the gravity centre – after a considerable spatial dispersion – shifted northwards.

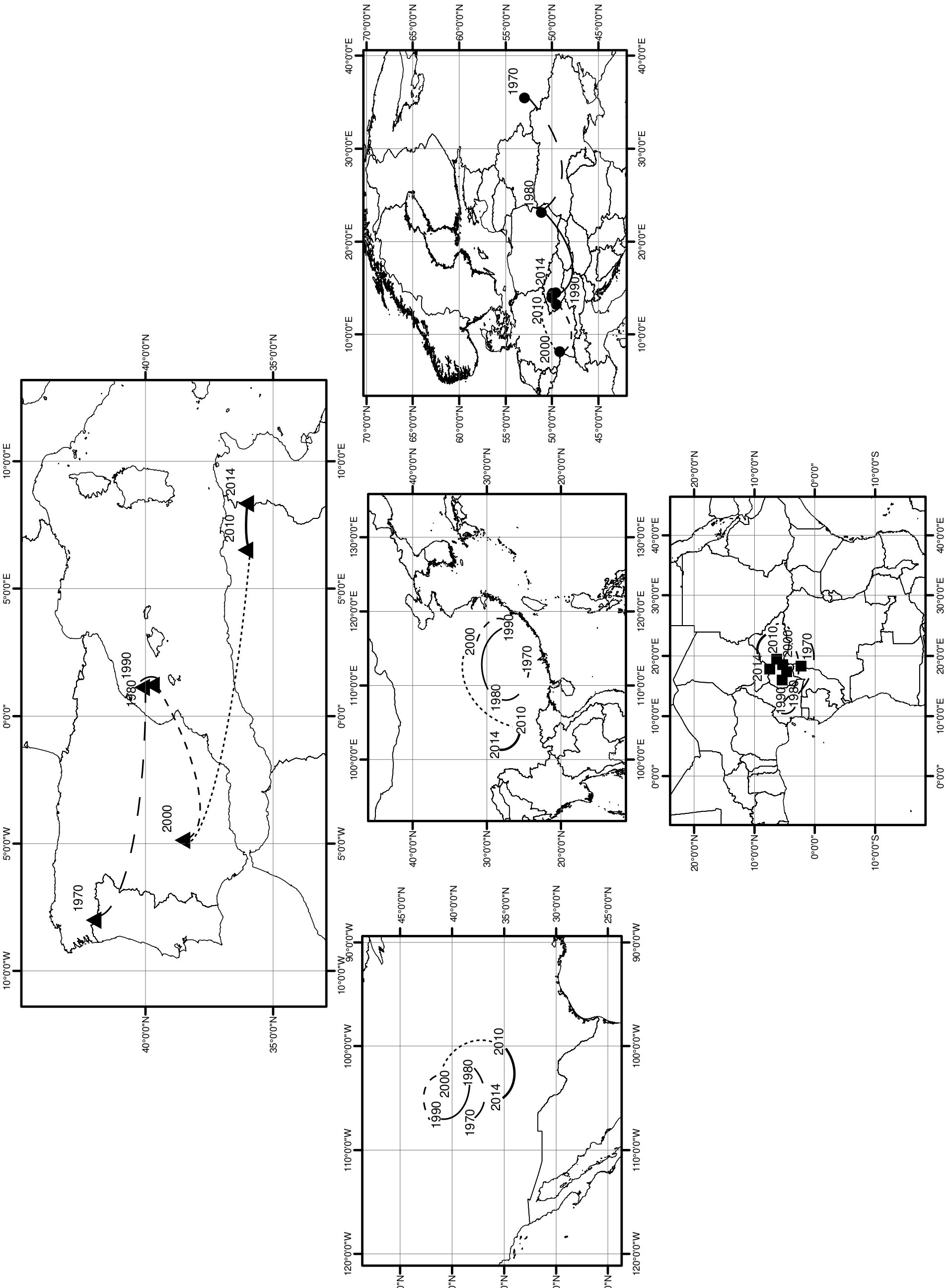
The economic centre of gravity of the *Americas* moved towards the United States, during the whole period. Beyond this, however, no change has been identified; however, we can find several underlying processes. First, from 1970 to 1999, the centre of gravity gradually shifted to the southeast, as a demonstration of the economic strengthening in Central and South America. This process came to a halt from about 2000 to 2004 and then intensified until the year 2011 when the southernmost part became the economic centre of gravity. Since then, the centre of gravity has again moved towards the north-northwest, owing to the economic recovery in North America.

The economic centre of gravity of *Asia, Australia, and Oceania* has been in the territory of the People's Republic of China, from 1970 to the present day, which provides enough evidence on the significance of the country in this vast region. In the 1970s and 1980s, the gravity centre mainly moved to the southeast of China, which is due to Australia's economic weight. This changed during the 1990s when Japan and South Korea were dominant economic powers in the region. In the 21st century, however, the centre increasingly shifted to the west, with one axis pointing to the oil-rich Middle Eastern countries and the other axis to India.

Finally, at the outset, analysis concerning *Europe* reveals that Russia and Soviet Union could not be broken down by continents, thereby affecting the study's results. The reason behind this decision, of course, was that Russia's economy in terms of performance in the European part of the continent is much more significant than that of Asia. At the outset, when examining Europe and Russia together, it is important to draw attention to the fact that the band's geometric centre is situated in the territory of Russia that is in Asia; therefore, we had to evaluate the position and movement of the economic centre of gravity in comparison to Asia.

In 1970, the centre of gravity moved from the Russian territories to the southwest towards France by the end of the 20th century. In the 21st century, however, because of the gradual economic strengthening of Central and Eastern Europe, it headed towards the northeast, and it was found in the Czech Republic in 2014.

The world's economic centre of gravity



REFERENCES

ECONOMIST (2012): *The world's shifting centre of gravity* Jun 28th 2012

GRETHER, J. M. –MATHYS, N. A. (2010): Is the world's economic centre of gravity already in Asia? *Area* 42 (1): 47–50.

QUAH, D. (2011): The Global Economy's Shifting Centre of Gravity *Global Policy* 2 (1): 3–9.

Data and maps

http://ec.europa.eu/eurostat/web/gisco/geodata
http://data.worldbank.org/indicator/NY.GDP.MKTP.CD