MORAVIAN GEOGRAPHICAL REPORTS
EDITORIAL BOARD

Editor-in-Chief:
Bryn GREER-WOOTTEN (York University, Toronto)

Executive Editor:
Bohumil FRANTÁL (Institute of Geonics, Brno)

Technical Editor:
Tomáš KREJČÍ (Institute of Geonics, Brno)

Marketing Editor:
Stanislav MARTINÁT (Institute of Geonics, Ostrava)

Linguistic Editor:
Martina Z. SVOBODOVÁ (BM Business Consultants, Brno)

INTERNATIONAL ADVISORY BOARD

Pavel CHROMÝ (Charles University, Prague)
Marina FROLOVA (University of Granada)
Marián HALÁS (Palacký University, Olomouc)
Jan HRADECKÝ (University of Ostrava)
Karel KIRCHNER (Institute of Geonics, Brno)
Sebastian LENTZ (Leibnitz Institute for Reg. Geography, Leipzig)
Damian MAYE (University of Gloucestershire, Cheltenham)
Ondřej MULIČEK (Masaryk University, Brno)
Josef NAVRÁTIL (University of South Bohemia, České Budějovice)
Philip OGDEN (Queen Mary University, London)
Ján OŤAHEL (Institute of Geography, Bratislava)
Martin J. PASQUALETTI (Arizona State University, Tempe)
Michael ROTH (Nürtingen-Geislingen University)
Michael SOFER (University Bar Ilan, Ramat-Gan)
Milan TRIZNA (Comenius University, Bratislava)
Steven TUFTS (York University, Toronto)
Antonín VAISHAR (Institute of Geonics, Brno)
Dan VAN DER HORST (University of Edinburgh)
Miroslav VYSOUPIL (Palacký University, Olomouc)
Maarten WOLSINK (University of Amsterdam)
Jana ZAPLETALOVÁ (Institute of Geonics, Brno)

PUBLISHER
Institute of Geonics, The Czech Academy of Sciences, v. v. i.
Identification number: 68145535

MAILING ADDRESS
MGR, Institute of Geonics of the CAS, v. v. i.
Department of Environmental Geography
Drobného 28, 602 00 Brno, Czech Republic
(e-mail) mg@geonika.cz
(home page) http://www.geonika.cz

Brno, June 30, 2017

PRINT
NOVPRESS s.r.o., nám. Republiky 15, 614 00 Brno
© INSTITUTE OF GEONICS OF THE CAS 2017
ISSN 1210-8812 (Print)
ISSN 2199-6202 (Online)

Articles:
René BRAUER, Mirok DYMITROW
Human Geography and the hinterland: The case of Torsten Hägerstrand’s ‘belated’ recognition .......... 74

Pavlína NETRDOVÁ, Vojtěch NOSEK
Exploring the variability and geographical patterns of population characteristics: Regional and spatial perspectives ............................................. 85

Kristína BILKOVÁ, František KRIŽAN, Marcel HORŇÁK, Peter BARLÍK, Pavol KITA
Comparing two distance measures in the spatial mapping of food deserts: The case of Petržalka, Slovakia ......................................................... 95

David FIEDOR, Zdeněk SZCZYRBA, Miloslav ŠERÝ, Irena SMOLOVÁ, Václav TOUŠEK
The spatial distribution of gambling and its economic benefits to municipalities in the Czech Republic ... 104

Maria GUNKO, Tatjana NEFEDOVA
Coping with employment issues through commuting: Evidence from Central Russia .................. 118
Human Geography and the hinterland:  
The case of Torsten Hägerstrand’s ‘belated’ recognition

René BRAUER a,*, Mirek DYMITROW b

Abstract
Seeing Human Geography as a nexus of temporally oscillating concepts, this paper investigates the dissemination of scientific ideas with a focus on extra-scientific factors. While scientific progress is usually evaluated in terms of intellectual achievement of the individual researcher, geographers tend to forget about the external factors that tacitly yet critically contribute to knowledge production. While these externalities are well-documented in the natural sciences, social sciences have not yet seen comparable scrutiny. Using Torsten Hägerstrand’s rise to prominence as a concrete example, we explore this perspective in a social-science case – Human Geography. Applying an STS (Science and Technology Studies) approach, we depart from a model of science as socially-materially contingent, with special focus on three extra-scientific factors: community norms, materiality and the political climate. These factors are all important in order for knowledge to be disseminated into the hinterland of Human Geography. We conclude it is these types of conditions that in practice escape the relativism of representation.

Keywords: knowledge production, hinterland, social science, Human Geography, Torsten Hägerstrand, STS

1. Introduction

Human Geography is replete with concepts. Seen as constituents of thoughts – either as mental representations (cf. Locke, 1690; 1975; Hume, 1739; 1978), abilities (cf. Wittgenstein, 1953) or so-called Fregean senses (cf. Peacocke, 1992) – concepts are crucial for most psychological processes, including categorisation, inference and decision making. While psychological processes both affect and reflect the multiple ways we engage with the world (Pinker, 2007; Dymitrow and Brauer, 2016; Dymitrow et al., 2017), it is often forsaken that science – seen as a construct of the human mind – is an inseparable part of this reciprocity (cf. Fleck, 1936; 1986). Since concepts, as Margolis and Lawrence (1999, p. 1) put it “often reflect deeply opposing approaches to the study of the mind, to language, and even to philosophy itself”, scientific conceptual advancement, then, is unlikely to be obtained merely by means of intellectual consensus (Dymitrow and Brauer, forthcoming). Let us consider the circumstances.

New concepts and ideas are introduced into science constantly. Some ideas become popular while others fall out of favour; this is part of the natural progression of any growing academic field. It is not only the academic merit of competing concepts, however, or their internal validity that decides if these will be incorporated or abandoned (Johnston and Claval, 1984; Latour, 1993; Shapin, 2010). Classical philosophy of science and the idea of a disinterested academia may suggest this, but from 50 years of critical sociological studies of ‘science in action’ we know that this is not the case (Sismondo, 2011; 2012). Sociologically identified factors influencing the progression of science are: community norms (e.g. Merton, 1973; Mitroff, 1974; Shapin, 1982), material factors (e.g. Collins, 1981; Latour, 2005; Bennett and Joyce, 2013), and the wider political climate (e.g. Haraway, 1991; Edwards, 1997; Livingstone, 2010), to mention but a few. The overwhelming majority of classical sociological studies, however, have focused on the traditional scientific disciplines, such as physics, mathematics or biology (e.g. Shapin et al., 1985; Bloor, 1991; Collins and Pinch, 1993; Chalmers, 2013). Conversely, the social sciences have not yet seen a similar level of scrutiny, although this is slowly changing (e.g. MacKenzie, 2009; Lamont, 2009; Camic et al., 2012). Understanding scientific process as a human pursuit for knowledge, it is fair to assume that the social sciences – including Human Geography – operate in a similar fashion (e.g. Barnes, 2001; Law, 2004; Cloke and

a School of Hospitality and Tourism Management, Faculty of Arts and Social Sciences, University of Surrey, Surrey, UK (*corresponding author: e-mail: r.brauer@surrey.ac.uk)
b School of Business, Economics and Law; Department of Economy and Society – Unit for Human Geography, University of Gothenburg, Sweden
Johnston, 2005; Johnston and Sidaway, 2015). This paper sets out to investigate these particularities of knowledge production within a social science context, to further investigate the validity of this assumption.

To achieve that, we depart from Law’s (2004) model of how research/science operates on a sociological level. We apply this model to a specific social science case: the case of Torsten Hägerstrand and his concept of time geography. Time geography and Hägerstrand’s life lend themselves well to this type of analysis, because Hägerstrand’s rise to fame was not a smooth, straight-forward progression as the majority of accounts that discuss the life path of Hägerstrand might suggest (cf. Duncan, 1974; Hägerstrand, 1983; Sollhe, 1991; Morrill, 2005; Pred, 2005; Öberg, 2005; Buttimer and Mels, 2006; Buttimer, 2007; Lenntorp, 2008; Persson and Ellegård, 2012; Ellegård and Svevin, 2012). It hence represents a good starting point for a Science and Technology Studies (STS) type of analysis, which commonly departs from controversy or failed projects (e.g. Law, 1992; Latour and Porter, 1996; Venturini, 2010).

STS is an offspring of science studies, which investigates how social, political and cultural values affect scientific research and technological innovation, and how these, in turn, affect the interconnected network of society, politics and culture (Latour, 1987). STS departs from a socio-constructivist view of knowledge production, which epistemologically shifts the definition of how science works from philosophy to sociology (Latour et al., 2011). An STS approach views science as a process that “teaches fallibility, not absolute truth […]”, recognising the provisional character of all [scientific] knowledge” (Edwards, 2010, p. 438). This ‘provisional character’ applies to ‘knowledge about knowledge’ as well, limiting in this sense its predictive capabilities (Collins and Evans, 2008, p. 140). Since most STS knowledge was produced within a natural science setting, the question remains how applicable is this understanding within a social science context – or, here, in Human Geography?

2. How to study social science sociologically

In order to address this question, we depart from our interpretation of Law’s (2004) sociological model of how (social) science research is created and validated within a specific discipline. We are interested in investigating community norms, material factors and the wider political climate surrounding the inception of Hägerstrand’s ideas, as these all have been identified as important factors behind knowledge production within the natural sciences. Hägerstrand is known for his ‘seemal’ and ‘highly innovative’ work on migration, cultural diffusion and time geography, while also being a major figure within Human Geography (cf. Lenntorp, 2008). Hägerstrand’s academic profile offers a research-friendly context for examining social science in the making – from obscurity to worldwide recognition. At this point, a word on scope is in place. Although we do engage with Hägerstrand’s life and time on a minute level, neither Torsten Hägerstrand nor the development of time geography are the focal point of our analysis. Rather, by using different socio-material dimensions of his recognition and rise to fame as the analytical foci for our study, we dare to extrapolate his case to any conceptual advancement in Human Geography (and in social sciences at large) where recognition and popularity can be historically attributed to extra-scientific factors (which – we argue – almost always is the case). Furthermore, accounts of the life path of Hägerstrand often take a romanticising view of Hägerstrand on the circumstances of his recognition. Such accounts usually focus on his contributions and not on the circumstances of his recognition (e.g. Helmarfrd, 2005). In other words, the aim of this paper is to determine the specifics of the knowledge dissemination process within a social science context, using an STS perspective (cf. Latour, 1999, pp. 25–79), and add to the literature of how Hägerstrand was recognised for his ideas.

The principal method in our analysis is a historical deconstruction of Hägerstrand’s path towards recognition in the style of seminal STS, used to study the natural sciences in action (cf. Latour, 1987; Shapin et al., 1985; Bloor, 1991). This approach stems from the postmodernist critique of historic representation and is often used as an alternative method to explore the conventional history of science (cf. Schuster, 1995; Jardine, 2003; Phillips, 2012). Munslow (2006) distinguishes between the reconstruction, construction and deconstruction of historic accounts. A historical reconstruction and construction epistemologically treat the past as static, implying that their form of representation is an accurate account of the transpired events. Deconstruction, on the other hand, differs epistemologically in that the past is conceptualised as continuous event or as “the process of making history” (Munslow, 2006, p. 75). The historical deconstruction implied here aims to outline the external conditions for an historic event to take place in the first place, because “…it is only by doing this that we challenge the belief that there is a discoverable and accurate representable truthfulness in the reality of the past” (Munslow, 2006, p. 4).

Having employed a detailed content analysis of relevant accounts relating to the life and times of Hägerstrand, in line with the principles of historical deconstruction, we have followed the central methodological provocation of STS research, namely that “it could be otherwise” (Woolgar and Lezaun, 2013, p. 43). Such an approach re-introduces the ‘messiness’ of creation back into the analysis, as opposed to a ‘smooth’ historical reconstruction (Collins, 1981; Latour, 1999; Law, 2004; Brauer et al., 2016). Primarily, this establishes an undetermined timeline, without inherent directionality, undetermined in the sense that many different potential timelines could have arisen at any particular point in time. Methodologically, such an approach is very liberal in terms of guidelines; it is the data material that ‘decides’

---

1 The difference between reconstruction and construction lies primarily in the methodological choices regarding how historical facts are used (cf. Phillips, 2012, p. 27).

2 The data material for this study represents autobiographical and biographical accounts, eulogies, Festschriften, Hägerstrand’s original works and historical documents – to mention only a few. See the references for a full list.

3 Hägerstrand himself sees time in a similar fashion. He makes a distinction between the actual manifestation of time (the life path) as opposed to all potential possibilities (the space-time prism) (Hägerstrand et al., 2009, pp. 218–224). Here, we are primarily interested in the potential aspect of ‘how it could have gone otherwise’, as this is what we refer to when invoking an underdetermined timeline.
where the investigation will proceed, not the researcher. Venturini (2010, p. 260) summarises this approach in three methodological ‘commandments’:
1. “you shall not restrain your observation to any single theory or methodology;
2. you shall observe from as many viewpoints as possible;
3. you shall listen to actors’ voices more than to your own presumptions”.

We interpret these methodological commandments as follows. First, although we introduce a theoretical model of science (see the following section), this model is merely used as an analytical starting point, as we are “aware of the need to avoid any methodological strait-jacketing and remain open and creative in our thinking” (Tribe and Liburd, 2016, p. 45). Second, we ‘follow’ (cf. Latour and Porter, 1996, p. 204) Hagerstrand’s trail wherever it may lead us, which means that the employed historical deconstruction accounts for as many viewpoints on Hagerstrand’s life as possible. Lastly, we place the encountered actors’ voices above our ‘own presumptions’ in line with a ‘hierarchy of credibility’ among the different individuals and accounts involved. In other words, the closer a source to Hagerstrand and the time in question, the more credibility it attains (cf. Shapin et al., 1985). This also means that our own preconceived notions place last, while Hagerstrand’s own biographical account supersedes secondary accounts of his life. We begin our investigation by introducing the aforementioned model of science used as the basis for our analysis.

3. A sociological model of science

If science is not an abstract, disinterested pursuit of knowledge, then how does it work? The implication from the sociological scrutiny of science is that it essentially represents a human endeavour. In this sense it may be instructive to think of research using an economic metaphor of the classical science canon ‘nanos gigantium humeris insidentes’, as so eloquently expressed by Isaac Newton (translation: we are dwarfs standing on the shoulders of giants). The giants, however, are not passive: they represent powerful allies (cf. Shapin et al., 1985; Latour, 1987; Latour, 1988). The cost involved in the unmaking of a proposition is accepted as true or false, criticised or praised (Fleck, 1936) 1986; Latour and Woolgar, 1979; Law, 2004). The more alliances can be mobilised, the harder it becomes to defeat the newly-created proposition. Once enough allies are mobilised, a proposition becomes accepted as true (Latour et al., 2011).

Figure 1 is a visualisation of the process of transformation that every research project goes through, i.e. the making of science. Initially, a research focus needs to be established. This is achieved by drawing inspiration from previously established knowledge, the so-called hinterland (box 1 in Fig. 1). This allows for the identification of ontological categories, i.e. a system of ‘created truth’, agreed upon by way of methodological practices in line with the scientific standards of a particular hinterland (box 2 in Fig. 1). Ontological transformation hence reduces an overwhelmingly complex reality into an understandable and manageable size by removing and filtering out redundant (‘unwanted’) data. Epistemological transformation is achieved by agreeing upon a process of how knowledge is accumulated and structured, i.e. how claims of truth can be distinguished from false ones according to the established standards. Methodological transformation, which generates data and offers hints to potential correlations, isolates relevant information in line with procedures established within the methodological hinterland. The outcome of these three types of transformation (box 3 in Fig. 1) renders a theoretical idealisation of the observed reality, which subsequently becomes a simplified representation of that reality in correspondence to the established hinterland. As these simplified representations ultimately influence new studies, the whole process is infinitely repeated.

This model, however, runs into problems of representation, because previous knowledge (upon which its construction was based) inherently influences how reality is to be interpreted for every new study. As a consequence, STS researchers have called this dilemma the multiple reality assumption (cf. Mol, 2002). This interpretation is at odds with the conventional assumption – for example within multi-method research – that the more different approaches are implemented to solve a problem, the better our understanding of it. Instead, the multiple reality assumption implies that depending on what research is chosen to serve as an alliance, a new interpretation of the same reality is created (see Fig. 2). On a theoretical level, this gives rise to the inescapable relativism of ideas that has laid the foundation for much criticism towards classical definitions of science, which cannot circumvent this impasse philosophically (Kuhn, 1970; Feyerabend, 1988; Sismondo, 2012). The practical consequence of this contingency is that scholars can be referring to the same object, which nonetheless is conceptualised differently depending upon what parts of the hinterland the chosen approach departs from (cf. Mol, 2002; Law, 2004). The particularities of sociological knowledge production as outlined by STS scholars emphasise the process of science, i.e. the praxis of doing science. The implication is that although, philosophically, consolidation of contradicting knowledge claims cannot be achieved, in actual praxis it is possible (cf. Collins and Evans, 2002). STS-scholars usually identify extra-scientific factors as being important to escape the relativism implied by the multiple reality assumption by materially reinforcing and fixing one particular interpretation (Law and Urry, 2004; Latour, 2005; Mol, 2010).

Here, we want to focus upon those aspects of science that shape this outcome within a social science context. For the sake of clarity of argument, we limit our analysis to three factors: community norms, material factors and the political climate, and how they correspond to the case of Torsten Hagerstrand. Before we begin our analysis, however, we need to briefly introduce our protagonist and the ‘controversy’ surrounding his recognition – this section also reflects the common representations of his life, from which our historical deconstruction will depart.

---

4 We have used the term ‘controversy’ given that the cited literature regards the time lag of his recognition and incorporation into the geographic hinterland as an eventuality rather than a process. This is also the reason for choosing to put quotation marks around the word ‘belated’ in the title, given that the main line of our argument is to establish that there was nothing inevitable about this incorporation into the hinterland of Human Geography. Rather, it was a combination of Hagerstrand’s own effort and external forces, the latter being the primary focus of this article.
4. Torsten Hägerstrand and his ‘belated’ recognition

Torsten Hägerstrand (1916–2004) is considered one of the most prominent figures in Human Geography of the last century (Johnston, 2007), having significant impact on the discipline itself and public planning, both in Sweden and internationally (Öberg, 2005). He received many decorations in recognition for his contributions and Pred (2005) called him a ‘giant of the discipline’. In 2016, as geographical communities throughout the world commemorated the 100th anniversary of his birth, Hägerstrand continues to be referred to as “one of the most celebrated Nordic social scientists of the 20th century [who] has greatly influenced the development of a broad range of research areas” (University of Stockholm, 2016). Hägerstrand is most famous for his work on the diffusion of ideas and as the founding father of time geography (Persson and Ellegård, 2012). Concepts of his time geography were formally introduced to the wider English-speaking audience in his 1970 article: ‘What about people in regional science?’ (Hägerstrand, 1970). Moreover, he kept developing and refining his theory of time geography his entire life (cf. Hägerstrand et al., 2009). Hägerstrand was also involved in public planning, partly because his early work was largely ignored within Human Geography. A great number of works have been written in appraisal of his intellectual legacy (Hägerstrand, 1983; Sellhe, 1991; Öberg, 2005; Morrill, 2005; Fred, 2005; Buttner and...
Mels, 2006; Buttimer, 2007; Lenntorp, 2008; Persson and Ellegård, 2012). What these works usually argue, however, is that the conceptual roots of time geography were to be found in his early writings, like his doctoral thesis, and that his recognition was primarily due to his own actions. Nevertheless, it took almost seven more years for his ideas to be recognised (Duncan, 1974). Why did it take so long? As will become clear in the elaboration below, we have identified several extra-scientific factors that were vital for his recognition and all of them were outside of Hägerstrand’s zone of control. Before we turn to this, we need to define the concepts of ‘recognition’ and ‘success’ in a scientific context.

5. Deconstructing a success story

‘Success’ is not the unilateral achievement of a single person’s actions; rather, it represents the product of a multi-factorial chain of events that must have come together in order to warrant broader recognition (Latour, 1987). In a science context, STS researchers define the words ‘success’ and ‘recognition’ as synonymous (cf. Porter, 1978; Latour, 1988; Gieryn, 1999). But if we define recognition as ‘scientific success’, we also have to define science in this regard. Law (2004) defines science as “an activity that involves the simultaneous orchestration of a wide range of appropriate literature and material arrangements. It is about the orchestration of suitable and sustainable hinterlands” (Law, 2004, p. 29, italics in the original). In this sense, the scientific success and recognition of Hägerstrand represents a successful installation within the scientific hinterland of Human Geography. A deconstruction of Torsten Hägerstrand’s success can now be recounted, in particular focusing on community norms (and paradigms) as well as material and political factors that allowed its incorporation into Human Geography’s hinterland. In order to understand the particularities of his recognition, however, we also need to establish a timeline for his recognition.

Hägerstrand published his doctoral thesis in 1953 (Hägerstrand, 1953), laying the conceptual roots that many of the biographical accounts of Hägerstrand’s life mention, nevertheless it took almost seven more years for him to be recognised for them⁵, becoming part of the then-established quantitative revolution (Duncan, 1974). Some biographical accounts simply ignore this period. Hägerstrand himself treats this period only briefly in his auto-biographical reflections (Hägerstrand, 1983) and the majority of other accounts attribute it to some form of language barrier, by claiming that “many of his publications were in Swedish, however, Hägerstrand spoke fluent English and the language barrier is not reflected in Hägerstrand’s publication record. Instead, we suggest that placing this recognition in a wider context of the paradigmatic changes that happened within Human Geography at the time, offers a better explanation. This will be elaborated next.

5.1 Community norms and paradigms

In the case of Hägerstrand, prior to his recognition, he made attempts to promote his ideas abroad. He lectured in Oslo and Copenhagen in 1954, he visited Edinburgh as a guest lecturer in 1957, and he conducted an academic course at the University of Munich in the early 1960s. Despite all this, his efforts were unsuccessful, even his own senior students were “unsympathetic to his ideas” (Duncan, 1974, p. 128). Hägerstrand changed the angle in his work on diffusion by emphasising his empirical findings rather than his (time geographical) methodology. His methodological approach broke with convention in that it looked for regularities in migration patterns, and made use of computation and codification irrespective of locality. Such an approach did not fit well into the reigning regional paradigm of contemporary Human Geography, which focused primarily on descriptive accounts (Hägerstrand, 1983).

The first recognition of his work took place in the winter of 1959/60, when Hägerstrand attended a conference in Seattle, which by then was a stronghold for the new (quantitative) geography (Buttimer, 2007). In those years the paradigm changed, it was not Hägerstrand’s ideas that changed substantially, but rather Human Geography itself. Between 1950 and 1960 Human Geography underwent a paradigm shift (Johnston, 1979), which also changed how the work of Hägerstrand was subsequently viewed (Duncan, 1974; Brauer and Dymitrow, 2017). In most countries Geography Departments underwent two

⁵ The date of Hägerstrand’s “recognition” cannot be clearly defined: most of the cited accounts assume the publication of his doctoral thesis as the locus of his ideas. As such, this publication is treated as a point of “formal” establishment within the Human Geography hinterland. The exact point in time of Hägerstrand’s “recognition” cannot be given a specific date either, as it was a cumulative process (see Duncan, 1974, pp. 114–120). This is one of the main points of the argument presented here, as the incorporation was a complex process requiring several factors to be in place before the “recognition” could occur. We use the seven year time span for narrative purposes, as the same biographical accounts regard the 1960 invitation to the Seattle conference as a keynote speaker as the ‘official recognition’.

⁶ From the about-page of ‘Honors of the Association of American Geographers’ (AAG): “1968 – Outstanding Achievement: Torsten Hägerstrand and Joseph E. Spencer”.
substantial changes during this period. Firstly, the discipline of Geography split into Human and Physical Geography and secondly, there was a shift from the regional paradigm to a more quantitative paradigm (Johnston and Claval, 1984). Every country underwent its own individual change and the dates differ significantly. In United States, in the early 1950s the newly formed discipline of Human Geography experienced specialisation towards different sub-disciplines, including political, economic, historical and urban geography, to mention but a few. This dispersal fragmented the previously unified regional paradigm, undermining its legitimacy. In the mid-1950s, a climate of strong emphasis on positivism and more ‘scientific’ methods slowly transformed Human Geography, turning it towards quantitative methods that have now found their place amidst previous specialisations (Johnston and Claval, 1984).

As such, we suggest that Hägerstrand’s recognition and ‘instalment’ within quantitative geography was only possible once this new hinterland had been established. Only after new quantitative ideas had disseminated and created their own hinterland within the wider body of Human Geography knowledge first, then was Hägerstrand’s wider recognition possible. As such, Hägerstrand was literally ahead of his time.  

5.2 Material factors

Just as the changes within Human Geography were outside the control of Hägerstrand, so were other changes taking place at Lund, his home university. This became a cornerstone for his recognition, i.e. his work with computation. In a 1955 article, Hägerstrand’s hints at these benefits reaped from being in physical proximity to the academic environment of the University of Lund can be identified (Hägerstrand, 1955)7. This becomes even more apparent with the establishment of Sweden’s second computer in 1956, Siffermaskinen i Lund (SMIL). Bo Lenntorp describes this historical contingency by claiming “[i]t is the fortunate combinations, constellations that facilitate development and change [in science]. SMIL, Hägerstrand and Human Geography was one of these” (Lenntorp, 2006)8. At that time, computers were gargantuan machines requiring significant expertise of the staff to run and maintain them (Edwards, 1997). Hägerstrand only mentions this pivotal role of computers in passing in one of his own biographical accounts:

“… my school-mate since secondary school, Carl-Erik Fröberg, who had just come back from a stay in the United States […] introduced me to the concept of random numbers […] and handed over to me a thin pamphlet on the Monte Carlo Method” (Hägerstrand, 1983, p. 248).

Not only did his ‘school-mate’ introduce him to mathematical concepts that would define his doctoral thesis, Fröberg would also go on to hold the Chair in Numerical Analysis at Lund University. The chair was especially established for him and SMIL through governmental funds, due to the government’s vested interest in developing their own computer. Fröberg would teach courses in numerical analyses, which Hägerstrand attended and their influence can be seen in Hägerstrand’s work afterwards (cf. Hägerstrand, 1951; Hägerstrand, 1952; Hägerstrand, 1955).

The trip to the United States mentioned above was also financed by the Swedish government, in order to evaluate the possibilities to build its own computer (Fröberg and Sigurd, 1962). The contacts that were established during this trip with pioneers of computation like John von Neumann were essential in establishing the computer at Lund (Lund University, 2011). Hägerstrand is also mentioned as one of the applicants of SMIL, arguing for the continued use of the computer in a book published by Fröberg (Fröberg and Sigurd, 1962). Another positive impact cited in that book was the development of the programming language ALGOL 60 in order to improve the memory capacities of SMIL. This language became eventually a global standard, pushing computer development forward (Sperber et al., 2010). The government funding for the computer eventually ran its course, which led to its shutdown just before 1970. Nevertheless, the prestige of having had access to such a high-tech device over several years granted legitimacy to his words. Therefore, claims like “geographer[s have to] accept the computer as an everyday aid” (Hägerstrand, 1967b, p. 3) came from a position of authority, undoubtedly aiding his recognition.

In that sense, Hägerstrand undoubtedly benefitted substantially from having a computer available to him at Lund, the most advanced computer at one point in time (cf. Lund University, 2011). His physical proximity to SMIL can also be regarded as a materially afforded competitive advantage for Hägerstrand’s intellectual development and subsequent recognition. Therefore, the material proximity afforded by the environment of the computer raised the status of Hägerstrand and conceptually influenced his work. This contingency made his work accessible and trustworthy for dissemination within the quantitative hinterland of Human Geography, which eventually led to his recognition.

5.3 Political factors

In terms of political factors, which profoundly came to shape the career path of Hägerstrand, the connection to World War II must be mentioned. One such factor is the influence of Edgar Kant, an Estonian professor of Economic Geography. Kant introduced Hägerstrand to many different points of view, and became his mentor forming Hägerstrand’s later views (Öberg, 2005). Kant only migrated to Sweden, however, because of the Red Army invasion of his native Estonia (Jauhiainen, 2005). Another political factor that is connected to World War II is why he was invited to the aforementioned conference in Seattle in the first place. This conference put him in contact with American

---

7 Although Hägerstrand was recognised for his quantitative ideas, his version of time geography always mixed quantitative understandings with qualitative reasons. As such, quantitative/qualitative time geography is more a result of institutional demarcation rather than Hägerstrand’s own views. For a full description of Hägerstrand’s view on his version of time geography, see Hägerstrand, et al., 2009.

8 In a footnote, Hägerstrand mentions that the arguments he put forward within this article were based on his experience with a punched card tabulator, which he was allowed to borrow courtesy ofCarl-Erik Quensel, a professor of statistics at Lund University (Hägerstrand, 1955, p. 240). At that point in time he did not yet have access to a real computer.

9 Translation into English by the authors; original quote in Swedish: “som vanligt är de lekyssamma kombinationer, konstellationerna som skapar utveckling och förändring. SMIL, Hägerstrand och kulturgeograf var en sådan”, as recorded in a lecture given by Bo Lenntorp in celebration of the 50th anniversary of SMIL at the University of Lund, 2006-10-16.
quantitative geographers, which came to be the beginning of Hägerstrand’s wider recognition. They only invited him, however, because “USA visas were not granted on political grounds for their first two choices – Hans Bobek and Walter Christaller – [politically safer was] Torsten from ‘neutral’ Sweden” (Buttimer, 2007, p. 140).

Other political factors include Hägerstrand’s involvement in planning. Although this was mostly Hägerstrand’s secondary field of interest, it helped him in his career. He regarded this experience “valuable [to understanding] how the transformation of localities and region is bound up with events in society at large [...] but in general it represented a distraction” from his theoretical interests (Hägerstrand, 1983, p. 253). It also left a mark in his use of rhetoric as he used this involvement as an argumentative tool. For example, he claimed that the benefits of his quantitative approach represented a possibility to give insights into the time-efficiency of the policies, i.e. a quantitative measure to evaluate policies (Hägerstrand, 1967b, p. 18). Furthermore, Hägerstrand used examples from projects that he had worked on with government funding (e.g. the Öresund Strait project). He even used this connection to legitimise his critique of theoretical Human Geography in his most famous article “What about people in regional science?”, where he asserts that geography should serve as an instrument to guide policy and planning (Hägerstrand, 1970, p. 1). Once again, we can observe that there were factors outside of Hägerstrand’s zone of control. These political factors are important in changing the outcome of science as they reinforce or undermine a certain type of knowledge. In the case of Hägerstrand these political factors clearly gave him access to scientific networks and created a position of authority, aiding the incorporation of his ideas into the newly established quantitative hinterland of Human Geography knowledge.

6. “Genius is one percent inspiration, ninety-nine percent perspiration”

This famous quote is often attributed to the American inventor Thomas Edison. We can observe this tenacity in the case of Hägerstrand as well. The obstacles that were in place surely must have felt like insurmountable challenges for Hägerstrand. For one, Hägerstrand’s quantitative ideas were initially unpopular with geographers, and his recognition was only possible once the scientific paradigm in geography had changed to a more quantitative one, leaving him ‘out dry’ for roughly ten years. But he was not without ‘luck’. Hägerstrand had a material competitive advantage due to access to computers that the Lund environment provided. Also, his recognition was aided on account of Sweden’s neutral status during WWII: Hägerstrand’s Swedish nationality enabled him to acquire the crucial invitation to participate in the seminal conference in Seattle. In conclusion, all these extra-scientific factors came to influence the dissemination of his ideas and concepts into Human Geography’s hinterland. This, however, can create the impression that his recognition was primarily due to ‘chance’ events, a perspective which sometimes can be perceived as blemishing for the contribution of an individual. Perhaps this is the reason why some biographical accounts omit these aspects. Nothing could be further from the truth, however, and Hägerstrand’s involvement in planning best exemplifies his tenacity, and surely is worthy of admiration. We acknowledge that his tenacity also represents a vital contributing factor to his wider recognition, a fact explored in the rest of this section, which focuses on Hägerstrand’s involvement with spatial planning. Table 1 is an illustration of the key life-events that occurred in Hägerstrand’s life, which cumulatively led to his recognition, only after these were in place, was his incorporation in the Human Geography hinterland possible.

Torsten Hägerstrand had a lifelong involvement with spatial planning, which fitted well with his quantitative ideas and his computational knowledge. In general, however, he regarded his involvement in planning merely as a “distraction” from his real passion, which was research (Hägerstrand, 1983, p. 253). Therefore, to understand this involvement it is important to acknowledge that the previously mentioned split between Human and Physical Geography occurred in Sweden as well. In Sweden, this had happened in 1948, changing the discourse of the newly established Swedish Human Geography as it adopted a more social science focus. Effectively this changed the legitimacy of Human Geography as a subject. Swedish human geographers were searching for societal relevance for their work. In higher education it went so far that Human Geography was in jeopardy of getting its funding rescinded. Hägerstrand explains his initial interests in planning in the following terms:

“Since geographers had an uncertain future on higher levels in the school system we were several university geographers who felt it is our obligation to try to open a new labour market for our advanced students” (Hägerstrand, 1983, p. 252).

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1916</td>
<td>Born in Moheda (Sweden)</td>
</tr>
<tr>
<td>c. 1932</td>
<td>Start of friendship with Carl-Erik Fröberg</td>
</tr>
<tr>
<td>1937</td>
<td>Enrolled at Lund University</td>
</tr>
<tr>
<td>1953</td>
<td>Publication of his PhD thesis, Lund</td>
</tr>
<tr>
<td>1954</td>
<td>Lectured in Oslo and Copenhagen</td>
</tr>
<tr>
<td>1956</td>
<td>Establishment of SMIL at Lund University</td>
</tr>
<tr>
<td>1957</td>
<td>Guest lecturer at the University of Edinburgh</td>
</tr>
<tr>
<td>1959/60</td>
<td>Seattle conference</td>
</tr>
<tr>
<td>1966</td>
<td>Grant of 1.1 million SEK for regional studies</td>
</tr>
<tr>
<td>1968</td>
<td>Received an Outstanding Achievement Award from the Association of American Geographers</td>
</tr>
<tr>
<td>1970</td>
<td>Publication of ‘What about people in Regional Science?’</td>
</tr>
<tr>
<td>1979</td>
<td>Received Victoria Medal from the Royal Geographical Society</td>
</tr>
<tr>
<td>1992</td>
<td>Awarded Lauréat Prix International de Géographie Vautrin Lud (Vautrin Lud Prize)</td>
</tr>
<tr>
<td>2004</td>
<td>Died in Lund</td>
</tr>
</tbody>
</table>

Tab. 1. Key events in the life of Torsten Hägerstrand

Note: Events that are spread out over time are not included within this illustration of key events (e.g. the establishment of the Lund School of Geography starting in the late 1950s and spanning to the end of the 1960s).

Sources: Duncan, 1974; Hägerstrand, 1983; Sollbe, 1991; Buttimer, 2007; Persson and Ellegård, 2012

10 Other countries had their own versions of this existential conflict (see Johnston and Claval, 1984)
The labour market he refers to denotes the involvement of geographers in public planning. In this way, the new quantitative paradigm was very welcome in that it clearly separated the ‘scientific’ subject of Human Geography from ‘common knowledge’. Before the new labour market could be opened, however, Hägerstrand had to make geographical education relevant for the public planning context. This effort eventually led to geographers securing the labour market, whereupon geography was made relevant and enabled Hägerstrand’s involvement in the early planning stages of the Öresund Strait project. This tenacity helped to secure the position of Human Geography as an academic subject at Swedish universities.

Hägerstrand’s tenacity paid off doubly, as all these efforts allowed him to secure research grants from various government departments. One grant of 1.1 million SEK represented the biggest single research grant in Sweden up to that date, for any social science research (Sollbe, 1991). This money was also used to support Hägerstrand’s research, as it paid for seminars and publications, which later allowed him to establish a platform for his ideas. This and other research grants allowed the Department of Geography at Lund University to pay the researcher’s adequate wages and stop their “brain drain” (Öberg, 2005, p. 342). Furthermore, it aided with the funding of journals, financing seminars and travel expenditures, creating scholarships to send researchers abroad and monetary incentives for scholars to come to Lund – i.e. providing monetary resources to create and finance the so-called ‘Lund School of geographical thought’ (Buttimer and Mels, 2006). This, among other things, allowed Hägerstrand to enter the epistemic community of quantitative geography (Hägerstrand, 1983). In fact, this represents the particular network structure that became vital in Hägerstrand’s international recognition and subsequent success (Duncan, 1974).

As stated previously, Hägerstrand faced rejection in his early years. We argued that the concurrent paradigm shift within Human Geography offers a better explanation for his ‘belated’ recognition. Viewed in this light, it is somewhat ironic that his efforts devoted to the ‘distraction’ (his involvement in planning) eventually aided his main passion (research). As such, our analysis highlights another central ‘extra-scientific’ factor, i.e. the character of Torsten Hägerstrand. Much has been written about his character in the biographical accounts; nevertheless the tenacity he manifested is the last ‘factor’ we would like to stress as being vital for his recognition. This is exemplified through his involvement in planning. The research also points to other, more personal, aspects that were important for his success, such as his wife “being a lifelong unpaid research assistant” (Hägerstrand, as cited in Sollbe, 1991), the psychological effect of his son committing suicide, or his Swedish ethnicity. Although all of these factors represent potential fruitful avenues for future inquiries, for the sake of brevity and clarity of argument, they have been omitted in this paper.

7. Community norms, materiality and politics … some concluding remarks

Each scientific discipline has its own set of tools – ontological, epistemological and methodological – which form a hinterland of concepts and ideas that define that very discipline’s ‘identity’. Human Geography is not an exemption in this respect. Although new concepts and ideas are constantly introduced and fall out of favour, respectively, their success or demise depends not only on internal but also on external factors as well as the ‘qualities’ of the people involved. Hägerstrand claimed as much himself, when he stated that “[i]deas do not succeed one another in de-natured space; rather they grow out from roots in progressively changing life experiences” (Hägerstrand, 1983, p. 9). We have identified three extra-scientific factors that correspond to these ‘natured spaces’ within Human Geography – community norms, materiality and politics – which all starkly influenced the dissemination of Torsten Hägerstrand’s ideas.

At this point, a note on Human Geography within the context of our analysis is necessary. Being more diverse than any other of the traditional disciplines, Human Geography represents a specific case of knowledge production. Its diversity – in the sense that ‘old practices are not completely replaced by newer ones’ (cf. Sheppard, 1995) – creates ‘niches’ for diametrically opposing epistemologies under the banner of one discipline (cf. Barnes, 2001). Since these niches (i.e. hinterlands) have to be established in the first place, however, there is not just one Human Geographical hinterland, but several, and they are constantly changing and evolving (cf. Johnston, 2008). While this certainly makes Human Geography particularly susceptible to alterations on account of competing epistemic practices, the underlying hows and whys often become blurred or even completely omitted from analysis. Effectively, such accounts create a deceptive image of a linear knowledge production within Human Geography, whereas much of it is actually purely circumstantial, incidental or provisional. When assessing the quality of geographical knowledge and especially its state-of-the-art, these factors should unconditionally be taken into consideration.

Despite Human Geography’s diversity, the mechanisms behind its recognition function are similar to those of other scientific disciplines. While these types of external factors are well-documented within the natural sciences, the social sciences have not been scrutinised to a similar extent. As we can observe, research, scientific success and recognition have material and sociological contingencies within the particular contexts of their creation. Therefore, the assumption that all scientific disciplines function sociologically in a similar way despite some apparent differences (cf. Camic et al., 2012; Dymitrow and Brauer, 2017), seems justified. While academic success is usually considered the intellectual achievement of the individual researcher, we tend to forget about the externalities that contribute greatly to knowledge production. We agree with Annemarie Mol (2002) that it is these types of norms, materialities and socio-economic conditions that in practice escape the relativism of representation.

Acknowledgment

The authors would like to thank Ron Johnston, Simon Duncan and Kajsa Ellegård who have all contributed with literature suggestions, which formed the starting point for our analysis. Furthermore, we would like to extend our gratitude to Mats Frödlund, who commented upon early drafts of this paper. Last but not least, we would like to thank the Unit for Human Geography at the University of Gothenburg for inspiring us to dig deeper into the works of Torsten Hägerstrand, and thus setting the ground for this paper.
References:


SISMONDO, S. (2012): Fifty years of The Structure of Scientific Revolutions, twenty-five of Science in Action. Social Studies of Science [online], [cit. 04.05.2014]. Available at: http://sss.sagepub.com/content/early/2012/03/29/0306312712439255.extract


Please cite this article as:
Exploring the variability and geographical patterns of population characteristics: Regional and spatial perspectives

Pavlína Netrdová a, Vojtěch Nosek a

Abstract

The variability and geographical patterns of population characteristics are key topics in Human Geography. There are many approaches to exploring and quantitatively measuring this issue. Besides standard aspatial statistical methods, there is no universal framework for incorporating regional and spatial aspects into the analysis of areal data. This is mainly because complications, such as the Modifiable Areal Unit Problem or the checkerboard problem, hinder analysis. In this paper, we use two approaches which uniquely combine regional and spatial perspectives of the analysis of variability. This combination brings new insights into the exploration of the variability and geographical patterns of population characteristics. The relationship between regional and spatial approaches is studied with models in a regular grid, using variability decomposition (Theil index) as an example of the regional approach, and spatial autocorrelation (Moran’s I) as an example of the spatial approach. When applied to empirical data based on the Czech censuses between 1980 and 2011, the combination of these two approaches enables us to categorise the studied phenomena according to the regional and spatial nature of their variability. This is a useful advance, especially for assessing evolution over time or comparisons between different phenomena.

Keywords: geographical patterns, regional variability, spatial autocorrelation, Theil index decomposition, Moran’s I, Czech Republic

Article history: Received 18 April 2016; Accepted 3 January 2017; Published 30 June 2017

1. Introduction

Many scientific disciplines study the socio-economic characteristics of administrative regions or other spatial units and the differences between them. Although these disciplines often study the same phenomenon, the terminology is inconsistent and the applied methods differ. The reasons are manifold: the researchers’ different professional backgrounds and routines in the field, varying research goals, issues concerning the availability of suitable data, etc. The result is that while some authors refer to variability, others use terms such as inequality, geographical concentration, spatial concentration, agglomeration, and polarisation. More importantly, while researchers with economic or regional science backgrounds usually prefer ‘pure’ quantification of regional differences by calculating some variability measures on a macro-scale level (such as NUTS 2 or NUTS 3 levels in the case of the EU1), geographers often try to look beyond the predefined

regions as a unit of analysis and focus on the micro-scale level (for example, municipalities). In this article, we aim to overcome the methodological divide between different approaches to measuring variability and geographical patterns by examining their relationship and joint use in empirical research.

The simplest way to measure the geographical variability of areal data is to apply standard statistical measures of variability to geographical data. These methods can be considered aspatial, however, as they do not work with the spatial information inherent in the data (Fotheringham et al., 2000). Therefore, they tend to not reveal much about the geographical organisation of phenomena in space and relationships to higher regional structures. Moreover, when studying geographical patterns, there are two main concerns associated with aspatial methods when applied to areal data – the Modifiable Areal Units Problem (MAUP) and the checkerboard problem. The MAUP deals with the fact

1 The term NUTS refers to “Nomenclature of Territorial Units for Statistics”, a geocode standard for subdivisions of countries in the EU.
that areal data are sensitive to the definition of boundaries used for the construction of the units, which may not be meaningful for the analysed problem (Openshaw, 1984; Wong, 2009; Klapka et al., 2016). The checkerboard problem stems from the fact that the geographic positions of regions (and potential neighbourhood effects) is ignored even though they are based on spatial data (Guimaraes et al., 2011).

It is clear that aspatial methods are far from sufficient when assessing geographical patterns and variability, and this has been recognised by many authors (Arbia, 1989, 2001; Rey and Montouri, 1999). There are two fundamental approaches to addressing the MAUP and checkerboard problem and bringing spatial or regional perspectives into the analysis of variability. We call the approach to bringing spatial clustering into the analysis the “spatial approach”, and the approach to analysing variability at several regional levels and quantifying the importance of respective regional levels the “regional approach”. In this paper, the spatial approach is represented by spatial autocorrelation measures, which can quantify the level of spatial clustering in the whole study area and uncover local clusters, thus bypassing the checkerboard problem. We employ the example of Moran’s I as the most commonly used index of global spatial autocorrelation. The regional approach is represented by variability decomposition, which quantifies the importance of respective regional levels, thus addressing the MAUP problem. In this latter case, we use the example of the decomposition of the Theil index, a typical decomposable index used in inequality studies.

Sometimes aspatial methods are combined with spatial (Arbia, 2001; Lafourcade and Mion, 2007) or regional (Brühlhart and Traeger, 2005; Rey, 2004) approaches. It could be argued that differences of methodology account for the reason regional and spatial approaches are typically used separately and, as a consequence, have not previously been combined with aspatial methods by authors. This paper also responds partly to the challenges posed by Rey and Janikas (2005), with the goal of demonstrating the importance of studying the interrelationship of (regional) variability and spatial autocorrelation.

The main goal of this paper is to demonstrate how the conjoint use of regional and spatial approaches helps to uncover and understand the variability and geographical patterns of population characteristics that are unclear when only aspatial methods are used. We argue that using both approaches conjointly can offer more comprehensive and innovative results, as documented in some empirical studies (Blážek and Netrdová, 2012; Nosek and Netrdová, 2010).

In order to fulfil this goal and to interpret results correctly, we need to understand the relationship between these two approaches, explore it on simulated data, and test it on empirical data. Moreover, we can categorise empirical data on the basis of their spatial and regional perspective of variability, which is useful especially when assessing evolution over time or comparisons across different characteristics.

The paper is organised as follows. In section 2, we describe and discuss the theoretical-methodological aspects of the approaches and highlight their potential complementarity. In section 3, we specify the methods and data used, where we also stress the importance of testing statistical significance and distinguishing between stochastic and spatially contingent components of measured values. The relationships between the methods are fully analysed in section 4. In section 5, we present empirical examples from the Czech Republic. Section 6 summarises and concludes the paper.

2. Theoretical and methodological background

The variability of geographical phenomena is often studied with only aspatial methods, which are invariant to permutations across units and do not incorporate information about the absolute or relative position of the respective unit in the calculation (Fotheringham et al., 2000). Basic variability measures such as the variance and standard deviation are useful for quantifying absolute levels of variability, but they do not meet the independence of scale requirement and it is therefore difficult to use them for regional analyses. This can be solved by using the coefficient of variation. The drawback of the coefficient of variation is that it is calculated from the distribution’s mean, which is not resistant to the extreme values typical for asymmetrically distributed geographical phenomena (Imre et al., 2012; Korčák, 1938). A good way to assess the uneven distribution of phenomena in space is to construct a Lorenz curve, a graphical representation of the distribution of a studied variable (such as wealth) in a society or space. One statistic with a straightforward interpretation that can be easily derived from the Lorenz curve is the Gini coefficient, which, due to its relative independence of the mean is very popular, and probably the most widely used variability measure in the social sciences.

Each of these measures, including the Gini coefficient, satisfies the condition of anonymity, a property of being insensitive to any spatial permutation (Sen, 1972). This condition, however, is not always a desirable property of a variability measure and is more a pitfall (Arbia, 2001), especially from a geographical point of view. The total insensitivity to the geographical position of units under analysis leads to the same results when units with high concentration values are adjacent as when units are located in the opposite part of the study area. The problem of ignoring neighbourhood effects is known as the checkerboard problem (Guimaraes et al., 2011). Another well-known problem associated with the analysis of areal data is the Modifiable Areal Units Problem (Openshaw, 1984; Wong, 2009). This problem refers to the sensitivity of variability or other statistical measures to the exact delimitation of areal units. There are two components of this sensitivity: the zoning effect (the dependence of results on the changing zonal boundaries); and the scale effect (the dependence of results on the level of aggregation, e.g. from municipalities to a regional level of analysis) (Arbia, 1989). Marcon and Puech (2003) address this issue. They state that variability is measured at a single level (typically at a chosen administrative level). Since observations may differ at different geographical levels, however, it may be useful to measure concentration at different geographical levels simultaneously. When a less fragmented (i.e. more aggregated) regional structure is used, some local specifics may remain hidden in the regional means, and some interpretations may be biased.

Problems associated with aspatial methods can be partly overcome by analysing spatial variability through the concept of spatial autocorrelation. This approach enables us to incorporate the neighbourhood effects into measuring variability and to identify whether there is a significant spatial pattern. In this way, spatial autocorrelation addresses the issue connected with the checkerboard problem. There are many ways to measure spatial autocorrelation depending on the nature and properties of the data (Anselin, 1988; Cliff and Ord, 1973). In general, two forms of measuring spatial autocorrelation can be
identified – global and local. The indicators of global spatial autocorrelation measure the extent of spatial clustering of “similar” values. In local form, they identify exact spatial clusters and reveal their character. Examples of global spatial autocorrelation statistics include Geary’s c, Getis and Ord’s G, and Ripley’s K; local indicators include Getis and Ord’s Gi, and Ord and Getis’ O (Getis, 2007). In recent studies, the most frequently used indicator of global spatial autocorrelation is Moran’s I, which is based on covariance and has many similarities with Pearson’s product-moment correlation coefficient. Anselin (1995) introduced local Moran’s I, the local indicator of spatial association (LISA) statistics. LISA cluster maps show statistically significant units in four types of spatial association.

The relationship between the aspatial concept of variability and the spatial approach to variability (quantified by spatial autocorrelation measures) has been studied by many authors (Arbia, 2001; Rey, 2004). Although one may expect an empirical relationship, as supported by empirical data (Rey, 2004), no theoretical or mathematical relationship exists. Consider the example of when all values in a studied area are modified in the same way. For instance, when values are increased or multiplied by a positive constant, the measure of spatial autocorrelation remains unchanged but the measure of aspatial variability changes quite markedly. This also applies conversely when we fix the values of variability and make random shuffles or specific arrangements of data in spatial units. As shown, there is no theoretical relationship between the overall variability and spatial variability. Therefore, observed spatial patterns of variability are only of an empirical nature. Rey’s (2004) finding of a strong positive relationship between measures of variability in state incomes and the degree of spatial autocorrelation in the US over the 1928–2000 period has no methodological justification and is therefore purely empirical. As shown, there is no theoretical relationship between the overall variability and spatial variability. Therefore, observed spatial patterns of variability are only of an empirical nature.

The utilisation of spatial approaches in assessing geographical variability helps to control the checkerboard problem, but invariably the MAUP effects remain. The reason is that both aspatial methods and spatial autocorrelation methods, representing a spatial approach, only work on one level of analysis. There are, however, typically more geographical levels that can be considered (for instance, a municipal level as a micro-scale and administrative regions, say NUTS3, as a macro-scale). Moreover, it is desirable not only to quantify variability at various, yet still single, geographical levels, but also to be able to quantify the relative importance of geographical levels compared with others. The regional approach in this way enables us to assess geographical variability at different scales and with distinct delimitation of regions, thus controlling the MAUP. First, in this approach, it is important to distinguish between overall variability, as measured between units at the most detailed sub-regional level, and regional variability, as measured between regional means. Second, there are two types of regional variability. Simple regional variability quantifies the differences between regional means, while relative regional variability quantifies the ratio between simple regional variability and overall variability. The latter enables us to quantify the importance of regional levels in overall variability and thus to assess the relative importance of a specific geographical level on the differentiation of particular phenomena.

For the regional approach, the aspatial methods are not sufficient. Unfortunately, the most commonly used coefficient to measure variability, the Gini coefficient, cannot be decomposed without a residuum into between-group and within-group components (for Gini coefficient decomposition, see Lambert and Aronson (1993); Mussard et al. (2003) necessary for quantifying relative regional variability, unless a spatial weight matrix is brought into the equation (see Rey and Smith, 2013). Gini decomposition without residual was proposed by Okamoto (2009), although the between-region variability of the Gini decomposition is null only if the distribution within each sub-region is identical to all the others. Decomposition enables us to quantify the share of a selected regional level and “scale down” the variability. This reveals the most important regional/local levels, and has many practical implications. For these purposes, it is possible to use indices from the generalised entropy class, which are decomposable without residuum. Of these the Theil index is the most widely used measure of regional variability (Cowell and Flachaire, 2007). The convenience of the various variability measures is illustrated by Shorrocks and Wan (2005) or Subramanian (2004), and Litchfield (1999) describes the axiomatic approach to properties of variability measures.

The relationship between an aspatial concept of variability and a regional approach to variability (represented by simple regional variability) is similar to the spatial concept and is only of an empirical nature. As an example, consider when regional means remain the same, but the values within these regions change. Thus the same values for simple regional variability show different values for overall variability depending on the exact modifications of the data. This also applies conversely when we fix the values of overall variability and make random shuffle or specific arrangements of data in regions influencing the values of simple regional variability.

Methods of quantifying simple regional variability (aspatial) and relative regional variability (regional approach) use a predefined regional structure, and methods of measuring spatial autocorrelation (spatial approach) also require some predefinition. To determine spatial lag and measure spatial autocorrelation, one must define a spatial weight matrix that operationalises the concept of “near” spatial units that can influence each other (Cliff and Ord, 1973). By changing the regional structure, for example on different hierarchical levels, we can calculate the regional variability to assess the relevance of the specific regional delimitation and hierarchical level to a given spatial process. By changing the spatial weights matrix, for example by extending the distance of influence, we can assess the distance over which the spatial process operates. This idea, hidden in both approaches, highlights their similarity. The only difference is that one approach takes a discrete view of spatial processes embedded in regions when measuring regional variability, while the other takes a continuous view on spatial processes with distance-decay influence when measuring spatial autocorrelation. It can be argued that the method of repeating measures of regional variability for many regional systems “floating” in a given area and the method of using a discrete spatial weights matrix based on regions are interchangeable. But this contradicts the foundation of these approaches and we expect that combining the two (discrete and continuous) can help to better understand the regional and spatial consequences of a given process.
The difference as well as complementarity of regional and spatial approaches is summarised in Figure 1. There are four types of results when regional and spatial approaches are combined. If both relative regional variability and spatial autocorrelation are high, the characteristic under analysis may be considered spatially dependent and bounded in regions. The characteristic is concentrated, and within predefined regions. Sometimes the characteristic concentrates across regional borders, meaning it is spatially dependent but with no relation to regions. If the characteristic does not form spatial concentrations nor concentrate in regions, it is both spatially and regionally independent. It is theoretically possible that the characteristic does not form any spatial concentrations yet it concentrates on a regional level. This characteristic would be spatially independent yet bounded in regions.

![Fig. 1: Typology of areal data based on regional and spatial perspectives](Source: authors’ conceptulasion)

### 3. Methods

The ways of measuring spatial autocorrelation and relative regional variability as tools for quantifying the spatial and regional approaches to geographical variability are manifold. In this paper, spatial (spatial autocorrelation) and regional (relative regional variability) approaches are represented by two specific measures: Moran’s I and the Theil index (T) and its decomposition. The relationship between spatial and regional approaches is demonstrated in the examples of these two most widely-used methods and without loss of generality, some conclusions based on these two methods can be drawn.

The formulas for Moran’s I (Equation 1) and Theil index T (Equation 2) can be written as:

\[
I = \frac{n}{n-1} \sum_{i,j} w_{ij} (y_i - \bar{y})(y_j - \bar{y}) \left/ \sum_{i,j} w_{ij} \right.
\]

\[
T = T_B + T_W = \sum_{j=1}^{k} \frac{y_j}{y_i} \ln \frac{y_j}{y_i} + \left( \sum_{j=1}^{k} \frac{n_j}{n} \sum_{i=1}^{n_j} \frac{y_i}{y_j} \ln \frac{y_i}{y_j} \right) = T_B + T_W
\]

where:

- for I (Moran’s I), \(n\) = number of units, \(i\) = index for individual units, \(j\) = index for regions, \(k\) = number of regions, \(\bar{y}\) = mean of the variable under analysis, \(w_{ij}\) = spatial weight matrix;

- and for T (overall Theil index), \(T_B\) = between-region component of Theil index, \(T_W\) = within-region component of Theil index (see Anselin, 1988; Anselin, 1995; Elbers et al., 2008; Shorrocks and Wan, 2005).

While \(T_B\) is a measure of simple regional variability, the share of simple regional variability in overall variability: \(T_B / T\) measures relative regional variability. All computations regarding the Theil index and its decomposition were performed in MS Excel and EasyStat 1.0 (Novotný et al., 2014), and all computations regarding Moran’s I were performed in GeoDa 1.4.0 (Anselin, 2003; Anselin et al., 2004).

As the formulas show, to measure spatial autocorrelation requires a spatial weight matrix (\(w_{ij}\)) that operationalises the position and proximity of geographical units (Anselin, 1988; Cliff and Ord, 1973; Getis and Aldstadt, 2004). The selection of a particular spatial weight matrix is often considered crucial, and is said to have a significant effect on the resulting spatial autocorrelation values (Anselin and Rey, 1991). The selection of a spatial weight matrix is especially important when only one variable is studied and the results may vary significantly, or when systems with different regional structures are compared (Nosek and Netrdová, 2014). For studying general patterns and for interpretation (zero, high, or low Moran’s I), however, the choice of a spatial weight matrix is not that important and the simplest spatial weights matrix, first-order contiguity, suffices (Stakhovych and Bijmolt, 2009). We have confirmed this by testing 18 different spatial weight matrices (this exercise is not included in this text). All analyses in this paper therefore use the 1st order queen spatial weights.

When using spatial autocorrelation, inference is commonly used and integrated in the software developed for this type of analysis. The inference is generally based on the comparison of random and empirical distributions of data in the studied area. To assess the significance of Moran’s I against a null
In the 100 by 100 grid, 100 ‘regional’ units (10 by 10) are distributed randomly and in several specific ways in a pseudo-random data. The log-normal distribution is often into account. did not take decomposition and relative regional variability however, considered only a variability on a single level and interpret the results properly, one must understand how of a regional approach and Moran’s I as a representation model values of Theil index decomposition as a representation and spatial approaches and their relationship

In order to utilise regional and spatial approaches conjointly and interpret the results properly, one must understand how their theoretical and methodological aspects are related. We model values of Theil index decomposition as a representation of a regional approach and Moran’s I as a representation of a spatial approach in a regular grid by running series of simulations. Besides modelling the values we can study the relationship between these two approaches. The relationship between variability and spatial autocorrelation has not been studied in detail, with a few exceptions (Arbia, 2000, 2001; Rey, 2004; Rey and Janikas, 2005). Each of these authors, however, considered only a variability on a single level and did not take decomposition and relative regional variability into account.

The model consists of 10,000 log-normally distributed pseudo-random data. The log-normal distribution is often considered to represent socio-geographical data with the most accuracy (Novotný and Nosek, 2009). These data were distributed randomly and in several specific ways in a regular square tessellation with 100 rows and 100 columns. In the 100 by 100 grid, 100 ‘regional’ units (10 by 10) are specified. In addition to overall variability measured by \( T \) (the differences between 10,000 units), simple regional variability measured by \( T_R \) (the differences between 100 regional means) and relative regional variability measured by \( T_R / T \) (the share of simple regional variability in overall variability) can be calculated. The set of pseudo-random numbers helps to minimise the effect of differences in variability and the regular tessellations minimise the effects of regional delimitation. Different regular tessellations (triangles, squares, and hexagons) were compared in Boots and Tiefelsdorf (2000). Using different types of regular tessellations is far beyond the scope of this paper and not important for achieving its goals.

The relationship between relative regional variability \( (T_R / T) \) and spatial autocorrelation (Moran’s I) is expected to be rather complex. A strong positive relationship between the interregional inequality share and spatial clustering was found by Rey (2004), who pointed out the ease of change of this result through re-shuffling. Theoretically, when high spatial autocorrelation is observed, both high and low relative regional variability can be present. One might also assume that with very low spatial autocorrelation it is theoretically and mathematically impossible to have high relative regional variability. Figure 1 shows the theoretical possibilities which may occur. The main purpose of the simulation with the model data is to examine this relationship and to document the fact that certain values of spatial autocorrelation result in very different values of relative regional variability, and vice versa.

To simulate \( T_B / T \), the value in each unit in the 100 by 100 regular square tessellation remains the same but the regional borders of 100 regions shift (assuming the grouping in regions is exhaustive, mutually exclusive and there are no enclaves nor exclaves). In this exercise, \( T_B / T \) can vary from 0 to 1 depending on the delimitation of regional borders and their correspondence with spatial clusters. The values in the regular square tessellation, which are generated to have fixed values of \( T_R / T \) (10% through 90%, with a maximum close to 100%), are rearranged within this tessellation with the aim of customising the levels of spatial autocorrelation. Several random shuffles and specific arrangements (ranging from arrangements with assumed maximal upper levels of positive spatial autocorrelation, such as by a concentration of high values in the same half of the tessellation, to chessboard arrangements with a supposed maximal negative spatial autocorrelation), were applied to test whether different relative regional variability can show different spatial concentrations when spatially distributed in specific ways. Another goal was to test the hypothesis of a statistically insignificant spatial autocorrelation when data with different levels of variability are randomly distributed. This relationship is depicted in Figure 2.

For fixed \( T_B / T \), significant values of Moran’s I are observed even when the data are randomly shuffled. This confirms the assumption that with some non-zero relative regional variability (even when \( T_B / T \) is only 10%), insignificant spatial autocorrelation cannot be observed. In fact, quite a clear pattern can be determined. With increasing \( T_B / T \), Moran’s I increases with roughly similar values. This relationship may therefore be considered purely stochastic in nature. The situation is slightly different when specific arrangements for calculating spatial autocorrelation are used (simulating a limited number of empirical relationships). The relationship for low values of \( T_B / T \) is rather loose (for \( T_B / T = 10\% \), it varies from + 0.5 to −0.2). With increasing relative regional variability, the range of possible Moran’s I
values narrows. With extreme values of relative regional variability, Moran’s I is close to its possible maximum. From a stochastic perspective, there is a clear positive correlation. From the simulations, it is evident that only three basic combinations are possible: high $T_B / T$ – high Moran’s I, low $T_B / T$ – high Moran’s I, and low $T_B / T$ – low Moran’s I.

These findings may help with the interpretation of empirical results in several ways. For example, the low values of relative regional variability do not necessarily mean that the studied phenomenon has a weak spatial pattern or none at all. This can be caused by the regional delimitation which may not be appropriate for the particular phenomenon, in which case, spatial autocorrelation statistics would provide significant added value.

5. Empirical evidence

In this section the theoretical assumptions and modelled data are tested empirically using the case of the Czech Republic. The data set contains empirical data from four population censuses (1980, 1991, 2001, and 2011). All data were recomputed to the most current structure at a municipal level in 2011 and are thus directly comparable. We chose the Czech Republic as an empirical example because comparable data were available at a very detailed level in a very fragmented regional structure (6,251 municipalities). Besides the municipal level, the data were studied at the Czech regional structure – 13 NUTS 3 units (administrative macro-regions). At this regional level, the Prague region has been merged with its surrounding region (Central-Bohemian Region) in order to represent geographical processes more accurately (see Hampl, 1999). Due to their consistent statistical nature, characteristics with a minimum structural variable of 0 and a maximum structural variable of 1 were chosen for the study. Further, the characteristics were chosen based on their supposed behaviour in the geographical context of regional variability and spatial autocorrelation (following the results documented in Netrdová and Nosek, 2009). Based on these two requirements, the following characteristics were chosen:

- Unemployment (economic) – the unemployed population, normalised by the economically active population;
- Agriculture (economic) – the population employed in agriculture, normalised by the economically active population;
- Education (social) – the university-educated population, normalised by the population over the age of 15 years; and
- Age (demographic) – the population 65 years and older, normalised by the overall population.

Figure 3 captures the empirical results: the spatial approach to geographical variability (Moran’s I) on the vertical axis; and the regional (Theil index decomposition) on the horizontal. The setup is based on the theoretical assumptions presented in Figure 1. There are three combinations (types) represented in this empirical case. The unemployment rate (red) proved to be both spatially dependent and bounded in regions. This result was expected due to the fact that regional delimitation matches quite well the labour market delimitation (based on work flows).
The only exception is the year 1991, when the regional and spatial patterns were not yet developed. The employment in agriculture (green) is a typical representative of a spatially dependent characteristic with a weak relation to regions. In this case, the administrative regions do not fit with relatively highly agricultural areas determined to a large extent by physical geography. The share of university educated (blue) has always been regionally less dependent with clusters around larger cities. The clustering (concentration) has been steadily increasing since 1980. Demographic characteristics, such as the share of the 65 years and older population, are typical representatives of both spatially and regionally independent characteristics. The slightly higher values of both Moran’s I and \( T_B \) in 1980 is caused mainly by a different age structure in Sudetenland.

For a proper interpretation and understanding of the empirical results, it is critical to also consider the local level. Although both aforementioned methods (regional and spatial) are suitable for studying the geographical variability of the studied phenomena, several important questions remain unanswered:

- What role does “the spatial” play in the distribution?
- What is the nature of spatial clustering – can we identify development axes or nodes, areas of peripheries, and so on?
- In what localities does statistically significant clustering occur?

Answering these questions is common in most geographical research (see for example, Sun and Jones, 2013) and crucial for a geographical contextual understanding of the studied processes, which is naturally more important than mere quantifications of the differences.

Local statistics of spatial autocorrelation present the most suitable way to support simple graphical visualisation by identifying and testing spatial clusters. Local statistics have many advantages over simple visualisation, as well as when compared with global statistics showing the average for the entire studied area. They eliminate the problems of analysing spatial aggregated data, help to discover deviations from global statistics and thus help to better map spatial processes (Fotheringham, 1997; Unwin and Unwin, 1998). The advantages of LISA cluster maps are documented using empirical examples with Czech data.

In the category of variables “spatially dependent and bounded in regions” (see Fig. 1), one cannot assess the character of clustering and its relation to regional organisation. There may be differences, however, in the type of clustering. Areal clusters, axes, centres, or other specific patterns may form. From the studied characteristics, unemployment proved to have the highest relative regional variability, as well as global spatial autocorrelation. As shown in Figure 4, the unemployment rate forms spatial clusters with low-low types of clusters organised in axes connecting Prague with other regional centres in Bohemia.

Fig. 4: LISA cluster maps for empirical Czech data, weighting scheme queen contiguity 1st order
Note: The High-Low type of spatial association indicates that a municipality with a value above the mean is surrounded by municipalities with values below the mean, and so on. The significance cut-off value of 0.05 is used (after carrying out 9,999 permutations). The permutation procedure was performed using GeoDa 1.4.0.
Source: authors’ calculations based on Czech Statistical Office (population census 2011)
(Plzeň, Liberec, České Budějovice). The high-high types of unemployment rate clusters are located in structurally affected regions that were formerly oriented on heavy industry (northern Bohemia, the Ostrava region) and peripheral regions (such as southern parts of Moravia). The regional distribution of clusters corresponds with the evidence of high regional variability in the unemployment rate – in the majority of NUTS 3 regions, low-low or high-high clusters dominate.

The question of the different character of spatial clusters is also relevant when variables belong to the second category “spatially dependent with weak relation to regions” (Fig. 1). We can only assume that spatial clusters are formed across regional borders, i.e. they do not comply with the regional structure. Different spatial patterns are documented by employment in agriculture and the share of the population that is university-educated. Employment in agriculture is regionally more bounded than expected; however, the $T_D/T$ is lower at the higher geographical level (NUTS 3), indicating the possibility that the concentrations run across these regions. This characteristic can, to a large extent, be determined by physical-geographic attributes (visible in Fig. 4). Spatial clusters of the low-low type are located especially in border regions (regions with higher altitude) and in metropolitan areas (around Prague, Brno, Ostrava, and so on) that are organised in “areal clusters”. The share of the university-educated population shows a slightly different pattern. $T_D/T$ is relatively low, but Moran’s I is rather high because the university-educated population is concentrated in bigger cities around which high-high clusters can also be found (especially university cities such as Prague, Brno, Ostrava, Olomouc, Hradec Králové, and so on). There are not many low-low types of clusters for this variable.

Local analysis is also important for the category of “both spatially and regionally independent variables” (see Fig. 1). For these variables, no statistically significant spatial autocorrelation was observed. As these results were obtained for the whole area under study (in our case 6,251 municipalities in the Czech Republic) as average values, we cannot be sure whether the variable does not cluster in the whole area or whether there are some local clusters. Although the share of the population 65 and over did not appear to have levels of spatial autocorrelation as high as other variables, the LISA cluster map in Figure 4 shows significant spatial clusters similar to that of the agriculturally-employed population and the share of the population 65 and over (visible in Fig. 4). Spatial clusters of the low-low type are formed across regional borders, i.e. they do not comply with the chosen regional structure and thus can directly face regional variability for different regional structures can uncover the effects of that structure on interpretations of the final results.

In summary, the regional variability and spatial autocorrelation approaches are strongest when used conjointly rather than separately (Rey, 2004). They produce important complementary findings about spatial aspects of variability. In the relative regional variability approach, differences are attributed to geographical levels, while global spatial autocorrelation and its local form help to uncover local specifics that are unrelated to regional structure.

It is also important to mention the restrictions on how the methods proposed in this paper should be used. First, very detailed data are required for measuring spatial autocorrelation. This is often a problem in the social sciences. Analyses are therefore often limited to a few variables and frequently to data from population censuses. Second, the dependence of spatial autocorrelation statistics on the subjective choice of a spatial weighting scheme may be considered important. Our tests, however, suggest that the choice of spatial weight matrix does not influence the final interpretation. Finally, although the combination of both methods helps mitigate the checkerboard problem and MAUP, they should still be taken into account. For example, the results of the regional variability analyses depend on the chosen regional structure and thus can directly face MAUP. On the other hand, comparing simple and relative regional variability for different regional structures can uncover the effects of that structure on interpretations of the final results.

The combination of spatial and regional viewpoints can have interesting implications for public policy as well. It is clear that attributing processes to different regional levels (and thanks to local analyses, also to specific localities) can have strong practical implications. In the geographical context of the European Union, it is highly relevant to study the role of international borders and/or other regional borders (such as NUTS 2, the basic regional units for the EU’s convergence policy). The study of geographical patterns and variability should be extended in research that deals with local or micro-regional data and when comparing the regional structure of socio-demographic and socio-economic indicators (such as in Kladivo et al., 2012). Not only can this help to study cross-border cooperation (or segregation), but it can also help to understand the EU’s integration process better.

Despite the interesting findings of this study, there are still many avenues for future research. The generalisations presented here should be tested repeatedly with other...
variability and spatial autocorrelation statistics. Importantly, these methods should be employed in empirical research. By capitalising on the advantages and complementarity of both approaches, interesting and innovative outcomes can be uncovered to reach a better understanding of both regional differences and dependencies, as well as the spatial effects of variability.

Acknowledgement

This work was supported by the Czech Science Foundation (GACR) under Grant No. 15-10493S – “Evolutionary dynamics of spatial differentiatation of socioeconomic phenomena and the role of regions in Czechia spatial and multilevel approach”.

References:


ANSELIN, L., SYABRI, I., KHO, Y. (2004): GeoDa: An introduction to spatial data analysis. Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign, USA. [online]. [cit. 05.04.2013]. Available at: http://www.csiss.org/events/workshops/geodaGA.pdf


HAMPL, M. [eds.] (1999): Geography of Societal Transformation in the Czech Republic. Prague, Department of Social Geography and Regional Development, Faculty of Science, Charles University in Prague.


Please cite this article as:

Comparing two distance measures in the spatial mapping of food deserts: The case of Petržalka, Slovakia

Kristína BILKOVÁ a, František KRIŽAN a,*, Marcel HORŇÁK b, Peter BARLÍK a, Pavol KITA c

Abstract
Over the last twenty years or so, researchers’ attention to the issue of food deserts has increased in the geographical literature. Accessibility to large-scale retail units is one of the essential and frequently-used indicators leading to the identification and mapping of food deserts. Numerous accessibility measures of various types are available for this purpose. Euclidean distance and street network distance rank among the most frequently-used approaches, although they may lead to slightly different results. The aim of this paper is to compare various approaches to the accessibility to food stores and to assess the differences in the results gained by these methods. Accessibility was measured for residential block centroids, with applications of various accessibility measures in a GIS environment. The results suggest a strong correspondence between Euclidean distance and a little more accurate street network distance approach, applied in the case of the urban environment of Bratislava-Petržalka, Slovakia.

Keywords: food access, food deserts, distance measures, GIS, Bratislava-Petržalka, Slovakia

1. Introduction
Food is a core element of the basic economies and quality of life of individuals (Sadler et al., 2016). The “food environment” has recently become a focus in numerous scientific disciplines (Caspi et al., 2012; Glanz, 2009; Glanz et al., 2016; Charreire et al., 2010; Lytle et al., 2017; McKinnon et al., 2009; Pinard et al., 2016). The concept of a “food environment” can be interpreted in many different ways (McKinnon et al., 2009): in this paper, it is perceived as a retail environment where food retail is operated. Food retail covers both small-scale and large-scale retail units. Not surprisingly, the food environment influences consumer food selection and health outcomes (Gustafson et al., 2013). As noted by Glanz et al. (2009), research on the food environment has revealed that good access to supermarkets may be associated with greater fruit and vegetable consumption, more affordable prices and reduced BMI (Body Mass Index). Methodologies employed to assess the food environment include sales analysis, menu analysis, nutrient analysis and geographic analysis (McKinnon et al., 2009). Between 2007 and 2015, the most frequent methodology used to study the food environment was geographic analysis, utilised in 65% of all the articles (Lytle et al., 2017). Hence, geographical approaches, mostly based on the measurement of accessibility to food stores, still remain one of the most frequent (McKinnon et al., 2009), in spite of the criticism of some researchers (Caspi et al., 2012; Lytle, 2009, Minaker et al., 2013).

In the two last decades, attention has been increasingly paid to food access solutions (Shannon, 2014; Walker et al., 2010). According to Andreyeva et al. (2008, p. 1387): “… access to healthful food is a critical domain of securing high-quality nutrition”. Not surprisingly, access to food is increasingly considered as one of the main attributes of life-quality research, as low food accessibility is frequently closely related to phenomena of social inequalities, marginality or transport disadvantages (see, e.g. Hendrickson et al., 2006, or Raja et al., 2008). For geographers, research on the spatial distribution of food sources (and food stores, specifically) is routinely applied, but this approach invites challenges concerning the methodology and data sources leading to proper results and correct interpretations. Similarly, the identification of food deserts is a relatively new phenomenon in geography. Especially in the transitive societies of post-communist Europe, this was not an issue
until a short time ago. Only a few studies have examined the issue of food access and food deserts in Slovak conditions (see for instance, Križan et al., 2015).

Measuring access to food sources generates multiple questions: Jaskiewicz et al. (2016) warn that most of the methods used to measure food accessibility have certain limitations. Various approaches bring different results and limitations, and are therefore often not comparable. The development of GIS (Geographic Information System) tools more generally, has brought revolutionary progress in visualisation as well as analytical research methods utilised in food accessibility mapping.

An assessment of the two most frequently used GIS approaches to distance measurement of food accessibility in a post-socialist urban environment is the main aim of this paper. First, Euclidean distance applications are utilised, and then the shortest network distance is evaluated. Specific attention was paid to a comparison of the application of both measures to assess food accessibility in Bratislava-Petržalka. In this paper, the authors compare the results generated by food environment accessibility measurement in Bratislava-Petržalka, which is the largest residential neighbourhood from the communist period in Central Europe, with a specific post-communist urban structure different from the urban environment typical for Canadian or U.S. cities.

2. Food deserts as a subject of geographical research

As a consequence of expert discussions, the term “food deserts” was introduced and used for the very first time in the mid-1990s (Cummins and Macintyre, 1999; Reisig and Hobbiss, 2000; Wrigley, 2002).

The phenomenon known as food deserts is broadly and generally defined, but particular definitions might differ in their context according to the scientific focus of their authors. A food desert can be defined as an “...area, where foods are expensive and relatively unavailable” (Cummins and Macintyre, 2002, p. 2115). As noted by McEntee and Agyeman (2010, p. 165), the literature has agreed on a general definition of food deserts, defining them as “areas of relative exclusion where people experience physical and economic barriers to accessing healthy food”. For the purposes of this paper, a food desert is perceived as a territory where (respecting relevant criteria) consumers have no access to large-scale retail units offering cheap and healthy food compared to local small-scale retail units.

Guy and David (2004, p. 223) describe the main attributes of food deserts and their residents as follows:

- the residents may be physically disadvantaged in terms of mobility and accessibility;
- the residents may also be economically disadvantaged due to low incomes;
- such residents will probably have poor nutrition since they eat cheaper foodstuffs of lower quality;
- they will be geographically disadvantaged because of poor choice of food stores in their living environment; and
- local small-scale food stores supply only limited selection of foods at higher prices compared to large-scale stores.

Scientific literature on food deserts may be classified according to various factors. Settlement-spatial aspect is surely one of them. Based on this, we may distinguish (i) urban and (ii) rural food deserts. The issue of food deserts was initially researched in British cities which explains why most of the relevant literature on urban food deserts covers British urban environments. The issue was generally received in academic periodicals, however, later covering rural environments as well.

Numerous interdisciplinary studies have been published on this issue in the last twenty years (Walker et al., 2010). Essentially, the following three approaches within food deserts research may be distinguished: (i) the medical approach (Budzynska et al., 2013; Glanz et al., 2012); (ii) the spatial approach (Widener et al., 2015; Chen and Clark, 2015); and (iii) the economic-social approach (Hendrickson et al., 2006; Raja et al., 2008). Each particular approach is based on different input data and methods applied to analyse them (black population, obesity, low income, pregnant women, etc.). All approaches are united in measuring the distance, which stems from the elementary definition of food deserts. Authors do differ, however, in their approaches aimed at the identification and mapping of accessibility (Shaw, 2006). As McEntee and Agyeman (2010) indicate, an easily applied universal method for the identification of food deserts has yet to be developed.

With numerous concepts developed on food deserts issues and their criticism, a new concept named ‘food oases’ has appeared, contributing a somewhat more complex and critical view upon food deserts (Walker et al., 2011). The term ‘food oasis’ was implicitly described for the first time in the study by Short et al. (2007) depicting the impact of small-scale stores on food safety. Although the text of the study does not define a food oasis as a phenomenon, it is perceived as an antonym to food deserts. Thus, food oases represent a concentration of food stores highly accessible for low-income communities (cf. Walker et al., 2011, 2012).

3. Methods and data

The analysis was carried out in a GIS environment. Geographic information systems were originally developed as a tool to assess and visualise information of a geographical nature. The evolution of regional sciences and related scientific disciplines using GIS as an interdisciplinary analytical instrument has engendered a significant integration of spatial analysis and information systems (see Goodchild, 1987). GIS have been increasingly applied in economics, too (Cliquet, 2006). As Cromley and McLafferty (2002, p. 234) underline: “GIS necessarily emphasise accessibility, the geographical dimension of access.” Recent research on food deserts primarily makes use of an approach based on Geographic Information Systems (GIS)-based analysis (Shannon, 2015). GIS-based spatial analysis has become an essential tool for food system research, and the proximity of residences to large supermarkets or supercentres is a commonly-used proxy for access (Mulrooney et al., 2017).

Since healthy and inexpensive foods are usually offered by large-scale retail units (as reported by Andreyeva et al., 2008, or Križan et al., 2015), the analysis presented here is focused on mapping accessibility to supermarkets. Accessibility may be effectively examined by various accessibility measures (Handy and Niemeier, 1997; Vale et al., 2015). Many recent studies assess accessibility of services via the perceptions of respondents, which allows researchers to become more sensitive to various accessibility aspects perceived by residents (see for instance, Vojnovic et al., 2014). In this
According to Apparicio et al. (2008), the following four approaches to measuring distance are typically used: Euclidean distance (straight-line distance), Manhattan distance (distance along two sides of a right-angled triangle opposed to the hypotenuse), shortest network distance and shortest network time (Fig. 1). In this paper we compare two of them: a) Euclidean distance and b) shortest network distance.

A selection of appropriate territorial units representing the residential areas was necessary for the application of all the above-mentioned accessibility measures. Our ambition was to respect the specific urban structure of Petrážalka, consisting of urban blocks of various sizes and ground shapes. Unlike most modern planned towns, Petrážalka's ground plan is far from a grid pattern with square or rectangular shapes. In the territory of Petrážalka, 143 residential areas (or residential localities) were identified, each clearly delimited by the street network as a particular block of residential buildings. The accessibility were measured from the centroids of individual blocks in a GIS environment. Individual residential blocks are of various shapes, areas and population sizes. This was not an issue in this paper; however, as our aim was to examine the application of various accessibility measures rather than any detailed geographical interpretation of food deserts in the study area.

Data on the location of retail units and their attributes come from an extensive field survey supported by a VEGA project (contract No. 1/1143/12). There were 4,089 retail units located in the city in 2011. Their location corresponds well with the population distribution and daily routines of consumers (Križan et al., 2014).

4. Study area

Bratislava is the capital city of the Slovak Republic and is located in the southwest of the country (Fig. 2). The city of Bratislava is divided into 17 boroughs, of which Petrážalka is the most populated. As of 2011, Petrážalka registered 105,842 residents (Statistical Office of the Slovak Republic), which was 25.7% of the total population of the city. With a population density up to 3,700 thousand residents per km², the site ranks among the areas with the highest population densities in the country (Buček and Korc, 2013). Petrážalka was designed as one of the most ambitious projects of the former communist regime and represents one of the largest prefabricated housing estates in Central Europe.

In the last 25 years Bratislava has witnessed a considerable transformation, as demonstrated by vast suburbanisation processes unprecedented in Slovakia (Novotný, 2016; Tóth, 2012; Šveda and Podolák, 2014), although this transformation has not been accompanied by proper transformation of the urban transport infrastructure and capacities (Seidenglanz et al., 2016). As in other post-communist cities (see Maryáš et al., 2014), these changes have certainly had huge effects on the urban retail environment (Križan et al., 2014).

The retail sector is one of the most transformed sectors of economic activity (Križan et al., 2016). The formation of Petrážalka has had its own specific pattern. According to the 1970s and 1980s planning practices, there was a discrepancy between living spaces and the structures where food stores were supposed to operate. First, these facilities were mainly localised in the middle of residential blocks as small supermarkets, and a long time after that some extra retail facilities without conceptions of planning were opened. Generally, by the beginning of the 1990s, retail capacity was not sufficient to meet the growing demands of the area’s inhabitants (Spišišek, 1994).

As noted by Mládek (1994), like other similar urban structures, Petrážalka suffered from a general lack of services for its residents. One of the reasons was the fact that the
Fig. 2: Study area. Source: authors’ elaboration

Fig. 3: Accessibility of large grocery stores in Petržalka. Source: authors’ elaboration based on field research

Fig. 3: Accessibility of large grocery stores in Petržalka.

Source: authors’ elaboration based on field research

location of most of these facilities had been planned for the main urban axis of the neighbourhood, which, however, was never completed. In 1993, there were 207 food stores in Petržalka, representing 50% of all retail capacities with an area reaching over 14 thousand square metres (Spišiak, 1994).

Later, this under-developed food environment witnessed a severe retail atomisation process. By now, the food environment has gradually been concentrated into large-scale retail units. This process is largely followed by consumer behaviour adaptation (Bilková et al., 2016).

Today, over 31% of the overall food store capacity of Bratislava is located in Petržalka (75 stores in total), which seems to be sufficient. Petržalka is home to more than 50% of all retail capacities of the city (28,295 m²), with several large-scale shopping centres and numerous smaller food stores located there. The study area comprises 18 large-scale stores (hypermarkets, supermarkets) and 58 small-scale stores (Fig. 3). The floor areas of the retail units vary between 10 m² and 5,800 m² (Križan et al., 2015).

5. Results

Most consumers use passenger cars to reach food stores in Bratislava (Bílková et al., 2016). In Petržalka, nearly 70% of consumers prefer passenger cars or walking. This is one of the arguments for our selection of relevant accessibility measures and types of distances.

When assessing food environment accessibility, one might discuss consumer perceptions and food environment on the one hand (cf. Moore et al., 2008), or consumer
satisfaction with food stores accessibility on the other (Bílková et al., 2016). The question is whether consumer perceptions are really linked with accessibility values to various food stores as indicated by various accessibility measures. Though we respect the fact that distance to food outlets is a significant predictor of healthy food perceptions (Barnes et al., 2016), we are not able to simply compare the food stores accessibility measurement results with consumer perceptions, since our data on consumer perceptions do not cover all residential blocks in the study area.

The following part of the paper presents the application of the three above-mentioned accessibility measures to the examination of the accessibility to large-scale retail units in Petržalka. The first case covers the application of Euclidean distance measures, the second one applies the street network distance method. The Euclidean distance approach utilises buffer tools, while the latter method rests on service areas in a network analyst tool. The general picture of the large-scale stores accessibility in Petržalka is shown in Figure 3. Not surprisingly, the Euclidean distance approach brings somewhat less accurate results, finding most of the inhabited area as food oases, generally defined as opposed to food deserts (represented by residential blocks located over 1 km from the nearest large-scale unit: see Krížan et al., 2015).

The nearest supermarket accessibility measure (Acc1) is the first of the proposed accessibility measures employed in our study (see Fig. 4). Application of the Euclidean distance approach reveals residential blocks which might be identified as potential food oases, with accessibility below 1 km. More realistic results have been shown by applying the street-
network distance approach (similar to the approach used by Sadler et al., 2013). Based on this measure, nearly one-half of the 143 surveyed residential blocks may be considered as potential food deserts. While the first approach reveals potential food deserts only in the marginal areas of Petrážalka, the latter detects some areas also in its central part as a consequence of the complicated access road network. The average Euclidean distance to the nearest large-scale retail unit is 667 m, while the average street-network distance is higher at 1,054 m. There is a significant positive relationship between the Euclidean distance and street network distance measures: $r = 0.83 \ (p < .0001)$.

The second accessibility measure (Acc2) applied in this study basically detects the number of large-scale retail units accessible within 1 km distance (Fig. 5).

The Euclidean distance application has shown that no large-scale store is accessible from 14.7% of the residential blocks identified in northern parts of Petrážalka and in its peripheral western and eastern neighbourhoods. On the other hand, one fifth of the blocks demonstrate good accessibility (within 1 km) to three or more large stores.

The street network distance method, however, shows such accessibility of three large stores only for 1 locality. On average, the Euclidean distance measure results in 1.8 large stores within 1 km distance, while only 0.7 stores have been identified by using the street-network distance approach (see Tab. 1). There is not a significant relationship between the Euclidean distance and street network distance measures in this case: $r = 0.47$ (Tab. 2).

The third approach applied in this study is based on accessibility of three supermarkets operated by three different companies (Acc3) within various radii from the residential blocks. In this case, the variability in distance measurement was most evident (see Fig. 6). For the radius up to 3,000 m, the Euclidean distance method application shows the cumulative share of such blocks reaching 54%, while it is only 20% if the street network distance approach is applied. With the radius up to 5,000 m, the cumulative shares of the residential blocks rise up to 94.4%, and 78%, respectively. The average Euclidean distance to three large stores in Petrážalka is 3,006 m, while the average street-

### Tab: 1: Descriptive statistics of food access measures (Acc)

<table>
<thead>
<tr>
<th>Acc</th>
<th>N</th>
<th>Mean (m)</th>
<th>S.D.</th>
<th>Median (m)</th>
<th>Minimum (m)</th>
<th>Maximum (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acc1a*</td>
<td>143</td>
<td>667.4</td>
<td>362.28</td>
<td>604.0</td>
<td>70.5</td>
<td>2,301.9</td>
</tr>
<tr>
<td>Acc1b**</td>
<td>143</td>
<td>1,053.6</td>
<td>535.58</td>
<td>990.2</td>
<td>89.4</td>
<td>2,633.8</td>
</tr>
<tr>
<td>Acc2a</td>
<td>143</td>
<td>1.8</td>
<td>1.34</td>
<td>2.0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Acc2b</td>
<td>143</td>
<td>0.72</td>
<td>0.79</td>
<td>1.0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Acc3a</td>
<td>143</td>
<td>3,005.8</td>
<td>1,067.72</td>
<td>2,876.6</td>
<td>1,113.8</td>
<td>7,685.5</td>
</tr>
<tr>
<td>Acc3b</td>
<td>143</td>
<td>4,722.1</td>
<td>1,458.99</td>
<td>4,885.2</td>
<td>1,637.2</td>
<td>9,856.0</td>
</tr>
</tbody>
</table>

### Tab: 2: Pearson correlations (A) and Spearman rank correlations (B) between alternative types of distance

<table>
<thead>
<tr>
<th>Accessibility measure</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acc1</td>
<td>.83</td>
<td>.86</td>
</tr>
<tr>
<td>Acc2</td>
<td>.47</td>
<td>–</td>
</tr>
<tr>
<td>Acc3</td>
<td>.84</td>
<td>.85</td>
</tr>
</tbody>
</table>

Note: All coefficient values are significant at the $p < 0.0001$ level. Source: authors’ elaboration

---

**Fig. 6: Accessibility (in metres) of three large grocery stores operated by different companies (Acc3)**
Source: authors’ elaboration based on field research
network distance is 4,722 m. There is a significant positive relationship between the Euclidean distance and street-network distance measures: $r = 0.84$ (Tab. 2).

6. Discussion and conclusions

Two approaches to measuring accessibility to supermarkets in an urban environment have been evaluated in this paper, applying methods developed with the use of Geographical Information Systems. When evaluating spatial accessibility, the choice of distance type is likely to generate different results, potentially leading to significant measurement errors.

As noted by Glanz et al. (2016, p. 286), the science and art of measuring retail food store environments has expanded and matured significantly in the past decade, although standardised measures are not used routinely and there is much work to be done. We realise that food accessibility might affect food consumption, but there are many more factors acting in this. Moreover, selection of food store is not always linked with its distance (Ledoux and Vojnovic, 2013). In spite of that, research on accessibility measures concerning the food deserts issue can be considered as highly relevant (cf. Barnes et al., 2016).

This research contributes to a better understanding of accessibility, especially due to the fact that it represents one of the first attempts to investigate quantitatively, food deserts on a town-wide scale in Slovakia. So far, food deserts research has mainly been case-based and often descriptive rather than analytical. More specifically, this research is the first to deliver a comprehensive analysis of the distances that may possibly influence the results perceived by retail chains’ marketing experts.

The quality of food access in Petržalka, as a typical post-communist housing estate, is surprisingly comparable with conditions observable in U. S. or Canadian cities (Apparicio et al., 2007; Leete et al., 2012; Sparks et al., 2011; Jaskiewicz et al., 2016). In the case of Euclidean distance measures, the mean distance to a large grocery store is below 1 km, but ranges from 0.07 to 2.3 km (Tab. 1). There are only 1.8 large grocery stores within a 1-km buffer radius around a building centroid, and the mean distance to the nearest three different large grocery stores is 3 km. In the case of street-network distance, the results are less favourable: the mean distance to a large grocery store hardly exceeds 1 km, the number of large grocery stores within a 1-km service area around a building centroid reaches only to 0.7, and the distance to the nearest three different large grocery stores is 4.7 km. Related to this, we should emphasise that the geography of access to retailers in terms of their strategies.

Despite a significant positive correlation between the Euclidean and street-network distances (cf. Apparicio et al., 2008; Sparks et al., 2011), spatial and empirical perspectives suggest that more realistic results are derived from the shortest network distance method compared to the Euclidean distance approach. Therefore, the shortest network distance application seems to be more appropriate for investigating accessibility to food stores in urban environments.

Retail markets are highly saturated, which emphasises the need for managers to understand the existing competitive structure for putting in place strategies which will allow retail chains to survive (Križan et al., 2014; Sinha, 2000). Of course, this study has some limitations in terms of measuring any results of place marketing and the geographic extension of the research, and this may give direction to possible future research. GIS represents a helpful visualisation and analytical instrument useful for identification and assessment of food accessibility (Charriere et al., 2010; McEntee and Agyeman, 2010; Shannon, 2015). In addition, this type of analysis could provide a useful tool to retailers in terms of their strategies.

The conclusions of this study cover only one of the numerous aspects of food environment research focused on accessibility measures and distance types. Surely, further detailed research on how accessibility affects consumer behaviours is necessary. Apart from distance, the temporal aspects of the food environment and its dynamics are important (Farber et al., 2014; Widener and Shannon, 2014). Recently, more attention has been paid to research on temporal changes in the food environment (Widener et al., 2017).

Further research on the accessibility to food stores and the utility of types of distance measures could also cover the less urbanised post-communist environments (e.g. small towns) of Central Europe, as well as rural environments. Due to the low density of existing over-ground communication networks in rural areas, however, this will probably call for more specific approaches.

Acknowledgement

This work was supported by the APVV project named: Consumer Society and Consumer Regions. Stratification of Post-Communist Society (No. APVV-16-0232), the VEGA project named Specifics of time-space human behaviour under the impact of socio-economic changes (No. 1/0082/15) and by the VEGA project (No 1/0246/17).

References:


Please cite this article as:

The spatial distribution of gambling and its economic benefits to municipalities in the Czech Republic

David FIEDOR a, Zdeněk SZCZYRBA b, Miloslav ŠERÝ b, Irena SMOLOVÁ b, Václav TOUŠEK b

Abstract
Gambling is a specific type of economic activity that significantly affects many aspects of society. It is associated mainly with negative impacts on the lives of individuals and their families, but it also has a positive economic impact on the public budgets of states, regions and municipalities. In this article, we focus on a geographic assessment of the development of gambling in the Czech Republic, which is based on a spatial analysis of data on licensed games and data on the revenues of municipalities arising from gambling. It turns out that the occurrence of gambling is strongly influenced by binary centre/periphery dichotomy, with the exception of the Czech-Austrian and Czech-German border areas which are characterised by a high concentration of casinos resulting from more rigid regulation of gambling on the other side of the border. In this research, the authors develop an innovative scientific discipline within Czech human geography: The geography of gambling.

Keywords: gambling, regulation, municipalities, spatial patterns, economic benefits, Czech Republic

Article history: Received 1 August 2016; Accepted 12 May 2017; Published 30 June 2017

1. Introduction
Gambling is analysed in this contribution as it is connected with a specific type of socio-economic activity that has a number of significant impacts on a society and its economy (Williams, Rehm and Stevens, 2011). From this viewpoint, gambling appears to be an important social phenomenon. Its importance is also underlined by the fact that it is an historical phenomenon that has been accentuated during the 20th century (McMillen, 1996).

The Czech Republic has been experiencing gambling since at least the 5 or 6th centuries BCE (Szczyrba et al., 2015, p. 1). According to Mravčík et al. (2014), the gambling landscape of the Czech Republic is highly developed, such that it is one of the European countries with a high availability of gambling. Moreover, international data collected by the Australasian Gaming Council underlines a very strong position for Czech gambling on a global scale (Ziolkowski, 2016).

In spite of these facts, there has been a rather insufficient effort in research on gambling in the Czech Republic. This paper extends the present state of understanding of gambling in the Czech Republic in terms of its geography. Furthermore, the paper attempts to contribute to current theory regarding the geography of gambling, because a principal aim is to assess the spatial patterns of both the diffusion and economic aspects of gambling based on data from the Czech Republic. Specific research questions can be formulated as follows.

First, what is the spatial distribution of Electronic Gaming Machines (hereinafter EGMs) in the Czech Republic? Are there any spatial patterns or is their deployment irregular? Does the binary dichotomy “centre vs. periphery” apply for EGMs? Does the phenomenon of a borderland have any importance?

Second, are there any spatial patterns regarding the revenues from gambling that accrue to the budgets of individual municipalities? Again, the existence of the binary dichotomy “centre vs. periphery” is evaluated. Is the size of the population of settlements important in this context? Considering the fact that the attributes of gambling that are being tackled are not static and evolve over time, we will be interested in their recent dynamics, which, in the case of gambling, are strongly associated with changes in the legislative framework. The most recent significant legislative change affecting gambling in the Czech Republic took place in 2011.
Third, attention will be paid to the intensity of the dynamics of the attributes of gambling under evaluation – in the period before and after the aforementioned legislative change. Has the level of spatial concentration of gambling increased in recent years? With regard to the year of implementation of the legislative change (see below) and the currency of the data that is needed, we will also investigate changes in the character of the attributes of gambling under evaluation in the period 2010–20141.

In connection with this research, it is necessary to mention the Amendment to the Lottery Act No. 300/2011 Coll., that changed the system by which gambling companies are taxed, and contributed to an increase in the tax revenues paid to municipalities. At the same time, municipalities were granted competencies to regulate gambling operations in their jurisdictions. Mainly for this reason, the research questions were formulated not only to capture the extent of the above-mentioned attributes of gambling, but also to capture the dynamics of their evolution by comparing the years before and after the Amendment to the Lottery Act, i.e. by comparison of the available data from the years 2010 and 2014.

2. Theoretical background

At this point in time, betting and its associated gambling or lotteries have become a rapidly-growing industry. Never in history has gambling had such a great economic potential as that which can be observed in recent decades (Shaffer and Hall, 2001). This is largely due to technological progress in the development of gaming machines, and especially the development of the Internet, which has increased the number of opportunities for gamblers and the overall revenue from gambling (Griffiths, 1999). With the development of gambling, the interest of geographers in research into the spatial context of the spread of this phenomenon has also increased. The gambling landscape in an area is the subject of research not only for purely scientific purposes, but also in terms of institutional assessments of the real social impact that gambling brings. Gambling deepens the impacts of adverse socio-economic factors on the population in regions affected by gambling. Thus, it is often found that people living within 700 metres of casinos, gambling centres or betting shops, have twice as high a probability of becoming problematic players than residents living at a distance of more than three kilometres away from gambling premises.

The relationship between the interests of the gambling industry and the public sector, both at the national level and at regional and local levels, can also determine the proliferation of gambling in an area. At various levels of public administration, so-called gambling policies are formulated, by which countries, regions and sometimes even municipalities, make decisions about the form of gambling, the level of its taxation, location, opening hours of gaming equipment, etc. (Fijnaut and Littler, 2007). At the national level, countries apply different approaches to the control and regulation of gambling or define the parameters of the distribution of gambling centres and machines in the area. Different approaches, however, may occur over time even within a particular country. Such events therefore affect the dynamics of the spatial distribution of gambling over time.

Spatially, the gaming landscape is concentrated mainly in cities and large urban agglomerations (Klebanow and Gallaway, 2015), which, among other things, helps to create an infrastructure for urban tourism. Cities are a key factor in the location of gaming landscapes. On the other hand, the casinos and their complexes change the face and the functional structure of cities, often in terms of the transformation of cities in order to increase their attractiveness (McCarthy, 2002). In urban areas, geographical research has uncovered the spatial distribution of gambling establishments using the concept of availability (see Robitaille and Herjean, 2008; Krizan, Bilkova, Kita and Hornik, 2015; Russnak et al., 2016; Fiedor, 2016), and the concept of accessibility (see Doran and Young, 2010; Young, Markham and Doran, 2009; Young, Markham and Doran, 2012). Marshall (2005) distinguishes several attributes of availability: geographic, temporal and social. Other authors discuss ‘geo-temporal availability’, to which they also add financial availability (Moore et al., 2011; Thomas et al., 2011). Most empirical studies have confirmed a link between the geographic accessibility of gambling establishments and the prevalence of problem gambling (Pearce et al., 2006; Welte et al., 2009; Welte et al., 2004). For example, Pearce et al. (2006) found in their study that people living within 700 metres of casinos, gambling centres or betting shops, have twice as high a probability of becoming problematic players than residents living at a distance of more than three kilometres away from gambling premises.

The spatial spread of gambling is also significantly connected with socio-economically deprived areas, i.e. regions with a population with lower social status and economic problems (Orford et al., 2010). The reasons for the expansion of gambling in this type of territory have been explained, inter alia, by the lack of leisure opportunities for socially disadvantaged population groups (Beckert and Lutter, 2009). Gambling takes place primarily in the environment of gambling establishments and the community of gamblers includes people with low incomes, often unemployed and recipients of social benefits. Gambling deepens the impacts of adverse socio-economic factors on the population in regions affected by socio-economic disadvantages (Abbott et al., 2013). On the other hand, in economically stronger regions, gamblers prefer the Internet or live casino games. Both of these factors have a direct impact on the spatial distribution of gambling in the territory. Wardle et al. (2014) showed in their study that in the UK, the areas with a high density

---

1 The data for the year 2015 were not analysed due to their unavailability during the processing of this paper.
of gambling equipment correlate with areas of socio-economic deprivation (formerly economically thriving ports and older industrial cities).

As shown in some studies, for example by Hampton (2010), peripheral border regions are also economically interesting areas for gambling. The gambling landscape has a tendency to develop in this area, often because on one side of the border, gambling is strictly regulated, while on the other side the legislation is more indulgent. This legislative inequality creates economic opportunity for agents working in the gambling industry. Casinos in the border areas primarily work because of the demand from abroad and have only a probable effect on the local population (Felsenstein and Freeman, 2002). Local casinos increase revenues from tourism and are a source of so-called gambling tourism (Zagoršek and Jaklic, 2009). On the other hand, the peripheral regions that are not determined primarily by location, but rather by the social and economic characteristics of the population, are associated with gambling to a much lesser extent. The same can be said for regions with strong entrenched traditionalism and religiosity (Diaz, 2000).

3. Geographical context of the study

In the Czech Republic, with its experience of many years of gambling, research on gambling is carried out, but it is not as well established as in Western European countries. In terms of the subject, local research, which is implemented mostly within the medical sciences, focuses primarily on the socio-pathological and medical manifestations of gambling in society, the availability and functioning of addiction services, and possibilities of prevention and treatment of gambling addiction (Mravčík, 2015). Some attention has been devoted also to a comparison of conditions for the operation of gambling in the environment of the European Union, and for the issues of the operation and taxation of gambling in the Czech Republic (Vlčková, 2011).

In the scope of Czech geographical research, the topic of gambling became a part of its subject orientation only a short time ago, and only in a marginal way. An overview of gambling and associated problems in the Czech Republic, including a summary of the historical context, legislation, prevalence and research base and agenda, was recently presented by Sz cyrba et al. (2015). A popular learned article by Perlin and Bednárková (2015) discussed the negative consequences of gambling for Czech society. The issues of accessibility and the regulation of gambling in the city of Olomouc and its suburban hinterland have been analysed by Fiedor et al. (2016).

Globally, gambling is most often associated with large cities, and only exceptionally with rural areas. The Czech Republic (Fig. 1) is a country with 10.5 million inhabitants, with the level of urbanisation of about 70%. The largest cities are Prague (1,200,000), Brno (380,000), Ostrava (300,000) and Plzeň (160,000). The settlement structure in the Czech Republic is characterised by a higher proportion of small towns (under 10,000 inhabitants) and six thousand small rural municipalities (Majerová, 2007). It is important to note that casino games or EGMs were operated in 1,926 municipalities in the Czech Republic in 2010, while in 2014 it was only in 1,560 municipalities. Considering this fact and the formulation of the research questions in the introduction, we evaluated the basic geographic aspects of gambling at the level of higher administrative units, i.e. the administrative districts of municipalities with extended powers (total 205) and the City of Prague. Exceptionally, we assessed the geographic aspects of gambling at the level of individual municipalities (6,250 municipalities in 2010; 6,253 municipalities in 2014).

4. Data and methods

The Czech Ministry of Finance has a privileged position in the area of data collection on gambling operations in the Czech Republic, as well as their monitoring and control. With regard to the research questions raised in the introduction, we can divide the data that is used into two main spheres: data on the number of EGMs operating in particular municipalities of the Czech Republic; and data on the tax revenues of municipalities from gambling. The data on the number of inhabitants in the years 2010–2014 were

![Fig. 1: The Czech Republic: an overview map. Source: authors’ processing](image)

Note: The current administrative structure of municipalities with extended powers is available at Regional Information Service: Ministry of Regional Development (2017)
drawn from the public database of the Czech Statistical Office and, like the number of EGMs in operation, the decisive date was the end of the year².

In 2013, the Ministry of Finance began to publish a regular informative summary of licensed EGMs and casino games (before 2013, the informative lists were provided only on request). Each line in the list of approved EGMs (or casino games) represents just one specific licensed machine/game. From the data, we can read the entire address where the licensed game is located, the type of game (video lottery terminals [hereinafter VL Ts], electro-mechanical roulette, slot machines, casino games, etc.), their operators, and even the series number of each licensed game. On the other hand, the published informative lists of licensed EGMs and casino games have their limits. Probably the most prominent pitfall in using these data is the fact that the lists only include games that were licensed by the Ministry of Finance (slot machines accepting domestic currency outside casinos are licensed by the municipalities themselves). To obtain this important component of the data, we contacted the Ministry of Finance with a request based on an application of Act No. 106/1999 Coll., on free access to information³.

This is the only possible way to get a complete overview of all the licensed EGMs and casino games. Nevertheless, since 2012, the Ministry of Finance no longer has the precise addresses at which the slot machines licensed by municipalities are operated, which partially obstructs research conducted at the local level. Another equally significant limitation on the use of the data is the fact that the information on the addresses of licensed games is sometimes incorrect (non-existent combination of a street name and house number, misprints of names, etc.). Besides the lists of licensed EGMs and casino games, the Ministry of Finance also issues lists of currently applicable municipal ordinances that regulate (or ban) gambling. Because of the general legislation, every municipality has an obligation to notify the Ministry of Finance, which then initiates administrative proceedings to remove licences to operate machines/games that are in conflict with a newly-adopted municipal ordinance. The list is updated and available on the website of the Ministry of Finance of the Czech Republic.

Data regarding budgets and accounting information for all levels of the state administration and local government is collected by the MONITOR information portal of the Ministry of Finance. In the analytical section of this portal, it is possible to export the required items of tax revenue of each municipality as an MS Excel spreadsheet. The items searched for our purposes are under numbers 1347 and 1351 (in 2010) or 1351 and 1355 (in 2014) in the lists of municipalities’ revenues classified according to the type of revenue. The sum of these items results in the total income of municipalities from gambling⁴, which is analysed here.

The total tax revenues of municipalities were exported from the analytical part of the MONITOR portal by exporting class 1 Tax revenues according to the category of the budget structure.

Analysis of the spatial distribution of EGMs and casino games or municipal revenues from gambling is constructed on the basis of the proportion of the number of licensed EGMs and casino games per thousand inhabitants of the administrative unit (administrative district of a municipality with extended powers) or per the total amount of tax revenue. To capture any changes in the number of EGMs or casino games and changes in revenues from gambling in the period 2010–2014, we used the so-called change index (CI), i.e. the proportion of the number of EGMs and casino games (i.e. the income from gambling) in 2014 and 2010 reported as a percentage, i.e. multiplied by 100 (\(\text{CI} = 100\) \(\approx\) no change; \(\text{CI} > 100\) \(\approx\) an increase; \(\text{CI} < 100\) \(\approx\) a decrease in the number of EGMs and casino games in a given administrative unit). For the analysis of the proportion of gambling-based revenues in total municipal tax revenues we used a multiple box plot (Potter et al., 2006). The data showed in the chart (medians, box plot boundaries – the upper and lower quartiles, the non-outlier range values) were set to correspond with the nature of the data. For the sake of clarity, we have omitted outlier values, which are values exceeding the size of the box by 1.5 times. We are aware of losing full variability of the original data, on the other hand the outliers are extreme values which should have distorted the analysis. Specific municipalities with extreme values are mentioned in the manuscript and they are discussed below in the text. We used the STATISTICA software for drawing the boxplot.

Following the analysis of the spatial distribution of EGMs and casino games, we evaluated the change in the concentration of EGMs between 2010 and 2014 using a Lorenz curve and the Gini coefficient. In our evaluation, we used the number of EGMs including casino games and the population of each municipality. The Gini coefficient can be calculated as the ratio of the area bounded by a diagonal (the axis of the first quadrant of the Cartesian system of coordinates) and the Lorenz curve to the total area below the diagonal. In practice, however, the so-called Brown formula (Brown, 1994) is usually used:

\[
G = \frac{1}{2} \sum_{i=0}^{n} \left( x_i - x_{i+1} \right) \times \left( y_i + y_{i+1} \right)
\]

where \(x_i\) and \(y_i\) are the relative cumulative frequencies of both characteristics, \(x_0 = 0\) and \(x_n = 1\). It uses values from 0 to 1, where the value of the Gini coefficient approaching zero indicates the most even distribution of a phenomenon in the researched space and, vice versa, a value close to 1 reveals an extremely unequal distribution.

² Our analyses include the total population of the spatial units in question. At first, we planned to relativise the data using the population aged 18 and over, but then we refrained from that because of the lower availability of data on the spatial distribution of this age category and also because of the fact that the revenues from gambling that accrue to the budgets of municipalities can theoretically benefit all age groups.

³ The Czech Ministry of Finance archives all requests for information and data which were handled and thus it is possible to find on their website (www.mfcr.cz) all the data about the location of slot machines licensed by municipalities with which the authors worked (Ministry of Finance of the Czech Republic, 2015).

⁴ The total revenues of municipalities from gambling do not include revenues from administrative fees for licences for slot machines because they cannot be set apart from all the administrative fees. In comparison with the above-mentioned items, however, the amount these fees represent is negligible.
5. Results

First, an analysis of the spatial distribution of EGMs and casino games is presented. In the next part, we deal with the problems of revenues from gambling, both at the municipal level and in an aggregate form at the level of administrative districts of municipalities with extended powers. In both parts of the results, we also deal with the dynamics of the changes of these attributes of gambling in the Czech Republic (through a comparison of the years 2010 and 2014).

5.1 Spatial distribution of EGMs and casino games in the Czech Republic

The total numbers of EGMs and casino games differ significantly when we compare 2010 and 2014. While in 2010 the total number of gambling units in operation in the Czech Republic was 111,213, in 2014 it was only 69,713 (Ministry of Finance of the Czech Republic, 2015, 2016a). This drop was significantly influenced by the Amendment to the Lottery Act No. 300/2011 Coll., which came into force on 1st January 2012 and granted municipalities the right to regulate gambling on their territory and also significantly increased taxation on gambling operators. In the two years in question, the major share of the gambling market (not only in the Czech Republic, but also in all administrative districts) was occupied by VLTs, which were licensed by the Ministry of Finance. Until 2012 the municipalities could not regulate their numbers (Szczyrba et al., 2015).

In 2010, slot machines licensed by municipalities constituted over 30% of all the EGMs in the country, but in 2014, this proportion was less than 18%. This is connected with the fact that in a majority of the administrative districts of municipalities with extended powers, there are more machines licensed by the Ministry of Finance. In 2010, however, there were 44 administrative districts in which there were more machines licensed by municipalities than by the Ministry of Finance (Fig. 2).

These districts were administrative districts with a smaller total number of EGMs and casino games in operation that were typically demographically smaller (e.g. Konice, Nepomuk or Pacov), or administrative districts located in the Central Bohemia region (Benešov, Brandýs nad Labem-Stará

---

Fig. 2: Spatial differentiation of EGMs and casino games according to the administrative districts of municipalities with extended powers in 2010 and 2014 (as of 31st December).
Sources: Czech Statistical Office (2015); Ministry of Finance of the Czech Republic (2015, 2016a); authors’ processing
Boleslav, Černošice, Kolín). On the other hand, in 2014 the number of administrative districts where the number of slot machines licensed by municipalities decreased amounted to a total of 13. In some administrative districts, the number of slot machines licensed by the municipalities was lower in 2010, while in 2014 the number was higher: this was the case in the administrative districts of Děčín, Valašské Klobouky and Železný Brod. This change was accompanied by a significant decrease in the total number of licensed EGMs and casino games in these administrative districts. The low number of administrative districts in which there were more slot machines licensed by municipalities is primarily due to a simpler regulatory process. Slot machines are licensed or prohibited by municipalities themselves, while the EGMs and casino games licensed by the Ministry of Finance can only be banned by municipalities through an administrative procedure.

The absolute numbers of EGMs and casino games in operation in 2010 were clearly linked with the administrative districts with the city with the highest population in the Czech Republic. In 2010, for example, the City of Prague had 16,156 EGMs and casino games in operation, while in 2014 it was 9,044, which is again the highest number among all of the spatial units that were observed (in relative terms, in 2010 it was 14.52% and in 2014 12.93% of all the EGMs and casino games in operation in the Czech Republic). It is also necessary to pay attention to the administrative district of Olomouc, as on a long-term basis it has the largest number of licensed EGMs and casino games among cities with at least 50,000 inhabitants (Mravčík et al., 2014; Fiedor, 2016). In terms of the absolute numbers of licensed EGMs and casino games, the administrative district of Znojmo was sixth in 2010, and even fourth in 2014 (although it is in 21st place according to its population). More than half of the total number of licensed EGMs and casino games within the administrative district of Znojmo were located in Chvalovice, a village between the towns of Znojmo and Retz (on the E59 highway, a former duty free zone) at the Czech-Austrian border. When the relative numbers of licensed EGMs and casino games (i.e. per 1,000 inhabitants) are compared, gambling is the most popular in the administrative districts on the Czech-Austrian and Czech-German borders (Fig. 2). This is due to the municipalities (e.g. Dolní Dvořiště, Česká Kubice or Rozvadov) with large casinos with hundreds of EGMs for visitors from neighbouring countries. Unfortunately, we do not have any data about visitors to casinos (here or in other locations in the Czech Republic), unlike other countries (Makarovč, Macur and Rončevič, 2011).

Even the distribution of the lowest values of the relative indicator of the number of EGMs and casino games per 1,000 inhabitants shows certain spatial patterns. Figure 2 shows clearly that the regions with the lowest share of slot machines relative to the population are located in areas on the borders of three or four administrative regions. In both the years under observation, the lowest category (less than three licensed EGMs and casino games per 1,000 inhabitants) involved three administrative districts of municipalities with extended powers, which are located in these areas, namely, Bystrčice nad Pernštýnem (in the border area between the Vysočina region, South Moravia region and Pardubice region), Konice (in the border area between the Olomouc, South Moravia and Pardubice regions), and Královice (in the border area between the Plzeň, Karlovy Vary, Ústí and Central Bohemia regions). Outside the above-mentioned regions, the proportion of EGMs and casino games relative to the population was generally lower in administrative districts which were located near the borders of administrative regions.

Most of these regions can be identified with the so-called areas of the internal peripheries of the Czech Republic. By internal peripheries we mean those found in the research by Musil and Müller (2008). These authors used not only locational characteristics and economic indicators but also social indicators to outline the internal peripheries. The internal peripheries defined by these authors are spatially bound to the frontiers of the regional units and have socio-economic characteristics that can be used to explain the lower share of EGMs and casino games. First, there is a very low value for the density of the population, pointing to a small population in these districts. In addition, the local population has a higher proportion of older people aged 65 years or over and, on the contrary, a lower proportional representation of younger population groups. These social conditions can influence the demand for gambling in the sense that the number of potential participants in gambling is limited. This may also be supported by the fact that a significant number of the economically active population commute to work on a weekly basis. The economically active population of the internal peripheries is also increasingly employed in the primary sector of the economy, which is characterised by lower wages, which implies lower purchasing power among the local population, which acts as a brake on the development of gambling. We cannot, of course, forget the relative location of the regions with the lowest proportion of EGMs and casino games. The regions are located away from the backbone communications, which is another cause of low values for this indicator.

In addition to internal peripheries on the boundaries of the regional units, the lowest values of the observed indicators are found in certain traditional areas of Moravia and Silesia. In Moravia, it is the region of southern Moravian Wallachia (the administrative districts of Valašské Klobouky and Vizovice), while in Silesia, it is the cultural-historical area of Hlučín region (the administrative districts of Kravaře and Hlučín) and Jablunkov region (the administrative district of Jablunkov). In these regions, religious observance is also an important factor. The above-mentioned regions are the areas with the highest degree of religious observance in the Czech Republic. The higher proportion of the faithful in their populations may be a reason for the small number of EGMs and casino games in these regions (Díaz, 2000).

Figure 3 illustrates several important findings on the spatial patterns of changes in the indicators under evaluation between 2010 and 2014. It is possible to identify regions in which the absolute values of EGMs and casino games increased within the monitoring period. These regions are in a minority and can be divided into two categories. The first is that of the regions situated on the Czech-German or Czech-Austrian borders. In these regions we have demonstrated high levels of relative indicators presented in Figure 2. These regions (the administrative districts of Domžalice, Horšovský Týn, Mariánské Lázně, Tachov, Trhové Sviny and Vímrav) already had high numbers of EGMs and casino games in 2010 and they increased further in the period in question. Second, there were the regions (the administrative districts of Jablunkov, Kraslice, Mnichovo Hradiště, Moravské Budějovice, Ostrov, Poděbrady, Pražské and Vímrav) which had low numbers of EGMs and casino games (in absolute and relative terms).
in 2010 and where, despite an increase in the numbers of the entities being monitored, their absolute numbers in 2014 were still low.

In a substantial majority of the spatial units that were observed, the numbers of the EGMs and casino games is being monitored dropped. This tendency can be attributed to the above-discussed legislative regulation of gambling which took place in 2011. The most dramatic drop in the number of EGMs and casino games occurred in the above-mentioned regions of the internal periphery. This tendency may have been affected by the above-mentioned widespread conditions dampening the proliferation of gambling, in combination with the implementation of municipal ordinances aimed at the regulation of gambling. There are many examples of such responses of municipal councils in internal peripheries: a good example is Bystřice pod Perštejnem, where the former mayor Josef Novotný contributed significantly to a ban on gambling in all public places in the municipality and later, as a senator of the Upper House of the Parliament of the Czech Republic, was the first Czech politician to pursue the strict regulation of gambling. It is also necessary to draw attention to the dramatic drop in the numbers of EGMs and casino games in core regions around cities at a high level of the administrative hierarchy, e.g. the administrative districts of Brno, České Budějovice and Jihlava. The significant drop is due to generally binding regulations aimed at the complete elimination of EGMs and casino games by the above cities in the period in question.

The Lorenz curves (Fig. 4) show the spatial concentration of EGMs and casino games in municipalities and administrative districts with extended powers in the Czech Republic. In connection with legislative changes and the overall decrease in the numbers of EGMs and casino games, we can assume that gambling began to concentrate spatially. Both graphs in Figure 4, however, show a different reality and one can see how crucial is the choice of the level of an administrative unit. At the beginning of this section, we suggested that gambling in the years in question was connected with less than a third of the municipalities. Nevertheless, the graph shows that 80% of the population live in municipalities in which gambling was or still is carried on. At the municipal level, the spatial concentration of gambling in terms of the comparison between 2010 and 2014 remained virtually unchanged (the Gini coefficient for the

---

**Fig. 3:** Change index of the number of EGMs and casino games between 2010 and 2014
Sources: Ministry of Finance of the Czech Republic (2015, 2016a, 2016b); authors’ processing

**Fig. 4:** Lorenz curve showing the degree of spatial concentration of EGMs and casino games in the municipalities of the Czech Republic (left) and in the administrative districts with extended powers (right) in 2010 and 2014.
Sources: Czech Statistical Office (2015); Ministry of Finance of the Czech Republic (2015, 2016a); authors’ processing
year 2010 was 0.37 and it was 0.39 in 2014). At the same time the chart on the right (Fig. 4) shows that at the level of administrative districts with extended powers all the residents had access to gambling opportunities, as there were some EGMs or casino games in each administrative district in the two years in question. Additionally, we can also observe a higher spatial concentration of gambling in 2014 (the Gini coefficient in 2010 was 0.23 and it was 0.28 in 2014). Nevertheless, the difference in the spatial concentration of gambling is not large and the small values of the Gini coefficient for the level of administrative districts can mean higher spatial refinement and the already-mentioned fact that at the level of administrative districts, everybody has access to gambling opportunities.

5.2 Economic benefits of gambling to municipalities

In 2010, gamblers in the Czech Republic bet CZK 125.6 billion, of which CZK 93.8 billion was paid back to gamblers as winnings (Ministry of Finance of the Czech Republic, 2016a). In comparison with 2010, in 2014 both the volume of bets (CZK 138.1 billion) and the winnings paid out (CZK 106.7 billion) increased. In all likelihood this is mainly because of the aforementioned changes to the legislation, and the municipal revenues also changed substantially. The spatial analysis of municipal revenues is the subject of the next section.

While in 2010, the municipal budgets received CZK 1.5 billion from gambling, in 2014 it was already more than CZK 5.5 billion (MONITOR, 2015). It might be useful to outline the system of levies on gambling companies, which has been fundamentally changed by the Amendment to the Lottery Act in 2011. In 2010, the income of municipalities from gambling consisted of a local fee charged for the operation of a slot machine (provided the municipality had established this fee) and a contribution for purposes beneficial to the public (social, health, sports, environmental, cultural or other welfare services), part of which the operators of EGMs had to pay directly to the municipalities. The contribution for public purposes was paid as a progressive tax amounting to 6–20% of the profit made by gambling operators. In practice, large companies were divided and subsidiaries established on purpose to allow the parent company to divide its income and reduce the percentage of profit paid for purposes beneficial to the public.

At the beginning of 2012, the Amendment to the Lottery Act (No. 300/2011 Coll.) fundamentally changed the system of charges5. The administrative fees for a licence to operate EGMs or casino games were reduced and local fees, as well as the contributions to purposes beneficial to the public, were cancelled. In addition, the exception for the lottery and gambling companies consisting of an exemption from corporate income tax (19%) was cancelled and a flat tax rate for all gambling companies (20%) was introduced. Moreover, in the case of EGMs, the taxation basis consists of a proportional part (20% of the difference between the money put in and that paid back) and a fixed part (55 CZK for each EGM per day). The accounting definition of the taxation also differs according to whether it is for EGMs or lotteries and other similar games. In the case of EGMs, 80% of the taxation accrues to the municipal budget (depending on the number of EGMs), while the remaining 20% goes to the state budget. On the other hand, the income from lotteries and other similar games is divided as follows: 70% to the state budget and 30% to the budgets of the municipalities according to the classification of taxation (i.e. including those municipalities where gambling is not operated).

In 2010, the highest revenues from gambling accrued to the municipalities of the administrative districts of Vyškov, Vimperk and Kaplice (Fig. 5). The administrative district of Kaplice had the highest number of EGMs and casino games (per 1,000 of the population) of all the administrative districts; the situation in the administrative districts of Vyškov and Vimperk reflects the specific nature of the taxation of gambling in the Czech Republic before 2012, when gambling companies contributed only a part of their profits to purposes that were beneficial to the public (see above). The municipality of Komárno (700 inhabitants), in the area of the administrative district of Vyškov, is the domicile of the European Data Project6, a company producing EGMs for the European market, whose subsidiary companies, operating a network of casinos called Admiral Global Betting and American Game, contributed to the municipal budget to the tune of CZK 1 million7 in 2010 (mostly for the development of transport infrastructure). Neither EGMs nor casino games, however, were in operation in the municipality in that year. In the administrative district of Vimperk, there is a village called Stražný (400 inhabitants), which is home to several casinos. In 2010, this municipality received more than CZK 12 million from gambling taxation, which was designated mostly for investment projects (development of the sewerage system, public lighting, etc.) in the municipality (Ministry of Finance of the Czech Republic, 2016a). Through these examples we can document the importance of the presence of a gambling company in a municipality, not just until 2012, but also afterwards (see below).

The lowest incomes from gambling were observed in administrative districts located in the inner periphery of the Czech Republic and in the administrative districts of the Vysočina region, which mostly consists of small rural municipalities. In 2014, the highest income clearly accrued to the administrative districts of municipalities with extended powers located on the Czech-Austrian and Czech-German borders (Fig. 5). The map also shows the administrative districts with those of the largest cities which did not completely ban gambling by municipal ordinances at their centre. This ban mainly affected the revenues in the municipalities of Brno and České Budějovice. The lowest revenue from gambling in 2014 was accounted for in administrative districts on the borders of three regions (e.g. Bystřice nad Pernštejnem and Konice), which is caused by the small number of EGMs installed (see section 5.1).

5 The rate of the revenues was valid until the end of 2015. Since the beginning of 2016, the new taxation on gambling has established the rates of 23% (lotteries) and 28% (EGMs). At the same time, the rate of the fixed charge for every licensed EGM has increased to 80 CZK per day.
6 EDP was founded in 1996 and in terms of the number of employees (more than one thousand) it is one of the largest industrial companies in the South Moravian region. Its subsidiaries rank among the market leaders on the gambling market in the Czech Republic.
7 It was an exceptional year, because in 2009 it was CZK 28 million and in 2011 CZK 10 million (Ministry of Finance of the Czech Republic, 2016a).
The change in the absolute incomes of municipalities that constitute the administrative districts of municipalities with extended powers between 2010 and 2014 is shown in Figure 6. It is clear that only three administrative districts behave atypically (as a result of changes in the legislation and the generally expected increase in the revenues of municipalities): Bučovice, Vyškov and Vimperk. In the case of Bučovice and Vyškov, the drop in revenues from gambling is connected with a smaller number of EGMs and casino games in operation. On the contrary, the administrative district of Vimperk saw a significant increase in the numbers of EGMs and casino games in operation. An intensive revenue growth is present even in the administrative districts with an intensive decline in the numbers of EGMs and casino games. Only in the exceptional cases of some administrative districts can the increased revenue from gambling be seen as the current increase in the number of EGMs and casino games in operation in these regions.

The connection of gambling with the urban environment has been mentioned several times. Therefore, it is now necessary to examine the dependence of revenues from gambling on the size of a municipality. As every municipality already had at least some income from gambling in 2014, the analysis included all the municipalities of the Czech Republic. The box plot (Fig. 7) shows both the relationship...
between the incomes of municipalities from gambling and their sizes, as well as the dynamics that occurred after the aforementioned legislative changes. As already mentioned in section 3, the graph omits outliers. In 2010, the category of municipalities with up to 1,000 inhabitants showed no profits from gambling (there are nearly 5,000 municipalities with less than 1,000 inhabitants in the Czech Republic and several municipalities with low populations on the Czech-Austrian and Czech-German borders constituted outliers within this size category). On the contrary, in 2014, all the municipalities have some revenue from gambling (although in the category of municipalities with less than 1,000 inhabitants it is still a negligible amount). In 2010, the highest proportion of revenues from gambling in total tax revenues was observed in municipalities in the categories from 5,000 to 50,000 inhabitants (the differences between the categories within this range were minimal). It is possible to observe a relationship between the proportion of revenue from gambling and the size of a municipality, which is especially apparent in 2014. We can conclude that the proportion of revenue from gambling increases together with the increasing size of the municipality.

Fig. 7: Development of the proportion of revenues from gambling in the tax revenues of municipalities in 2010 and 2014 (N\textsubscript{2010} = 6,250; N\textsubscript{2014} = 6,253)
Sources: Czech Statistical Office (2015); MONITOR (2015); authors’ processing

Fig. 6: Change index of gambling revenues for municipalities with extended powers in the period between 2010 and 2014. Sources: Czech Statistical Office (2015); MONITOR (2015); authors’ processing
Considering the economic effects on the municipal budgets, confirmed by the research study of Osman and Šerý (2013), which resulted in a budget surplus in some municipalities enabled the return of taxes directly to municipalities. In 2010, the old system of taxation of gambling companies the economic benefits of gambling for municipal budgets. Germany (Mravčík et al., 2014). the more stringent regulation of gambling in Austria and Czech-German border area, which is characterised by a certain exception in this respect is the Czech-Austrian extended powers with some populous towns at their centre. of EGMs and casino games consisted of municipalities with the areas of the so-called internal peripheries by the lowest values of the concentration of EGMs and casino games. A majority of the regions characterised of concentration of EGMs and casino games in regions and a visible impact on the extent and regulation of gambling (Marshall, 2002; Sauer, 2000), with the liberal approach of legislation to gambling bringing its development and simultaneously higher tax revenues for the state. Even in the Czech Republic, it was confirmed that gambling is closely related to the urban environment (Mravčík et al., 2014). Gambling is found in peripheral regions to a noticeably lesser extent, although casinos and gaming clubs are also often found in border areas and suburbia (Abbott and Cramer, 1993; Felsenstein and Freeman, 2002), which is also the case in the Czech borderlands. Similarly to other countries, we can see some relation between a lower degree of concentration of EGMs and casino games in regions and a higher proportion of religious believers (Diaz, 2000).

6. Discussion and conclusions

The main objective of this paper was to assess the spatial patterns of the expansion of gambling in the Czech Republic, primarily by examining two of its attributes: EGMs and casino games, and the taxes on gambling companies that flow into municipal budgets.

In response to the research questions that were formulated, we can state that the binary dichotomy 'centre vs. periphery' applies in the context of the spatial distribution of EGMs and casino games. A majority of the regions characterised by the lowest values of the concentration of EGMs could be identified with the areas of the so-called internal peripheries of the Czech Republic delimited by Musil and Müller (2008), while a majority of the regions with higher concentrations of EGMs and casino games consisted of municipalities with extended powers with some populous towns at their centre. A certain exception in this respect is the Czech-Austrian and Czech-German border area, which is characterised by a high concentration of casinos, which can be associated with the more stringent regulation of gambling in Austria and Germany (Mravčík et al., 2014).

Similar spatial patterns can also be found in the case of the economic benefits of gambling for municipal budgets. In 2010, the old system of taxation of gambling companies enabled the return of taxes directly to municipalities, which resulted in a budget surplus in some municipalities in the administrative districts of Vimperek and Vyškov, as confirmed by the research study of Osman and Sery (2013). Considering the economic effects on the municipal budgets, we found that the relationship of the size of the population of a municipality and the proportion of the revenue from gambling in the total tax revenues of the municipality is very important (with the exception of the larger cities in the Czech Republic). Even in the case of the largest cities, the revenues from gambling represent a substantial part of the total tax revenues of the municipality, but because of the significantly differentiated sources of income of large cities, this income is not so crucial.

For the last research question, which was devoted to an evaluation of changes in the attributes of gambling in the comparison of the years 2010 and 2014, we can also give a clear answer. An expected trend that the regulation of gambling and resulting lower number of EGMs and casino games would lead to their growing spatial concentration (Jensen, 2017), has not been confirmed (see Fig. 4).

The results of this paper largely correspond to contemporary knowledge of scientific disciplines dealing with gambling. In recent years, changing legislation has had a visible impact on the extent and regulation of gambling in the Czech Republic. The literature often mentions the regular rotation of periods of liberalisation and regulation of gambling (Marshall, 2002; Sauer, 2000), with the liberal approach of legislation to gambling bringing its development and simultaneously higher tax revenues for the state. Even in the Czech Republic, it was confirmed that gambling is closely related to the urban environment (Mravčík et al., 2014). Gambling is found in peripheral regions to a noticeably lesser extent, although casinos and gaming clubs are also often found in border areas and suburbia (Abbott and Cramer, 1993; Felsenstein and Freeman, 2002), which is also the case in the Czech borderlands. Similarly to other countries, we can see some relation between a lower degree of concentration of EGMs and casino games in regions and a higher proportion of religious believers (Diaz, 2000).

These results, based on the analysis of the spatial distribution of gambling, must be subject to critical evaluation and possible limitations. First, the authors worked with data on legally authorised gambling establishments and betting games in casinos. This restriction cannot be disregarded, however, because information on illegally operated gambling cannot be obtained for the Czech Republic. Especially in cities that promote zero tolerance of hard gambling, illegally operated gambling venues may exist, but their number is not likely to be too large, as evidenced by Rossow and Bang Hansen (2015) in Norway. Neither are the official data on legally operating gambling entirely correct (incorrect or incomplete addresses of places where gambling devices and betting games in casinos are licensed), which was pointed out by Fiedor (2016) using the example of the city of Olomouc. It is, however, necessary to emphasise that when data are aggregated to higher-level territorial units, these errors do not manifest themselves. For our comparison of the periods before the major amendment to the Act on Lotteries No. 300/2011 Coll. and the period after the amendment, we used the years 2010 and 2014. Using data for individual years can be problematic, however, because in those years some more significant fluctuations in the revenues of municipalities can occur (until 2012, the operators of gambling establishments could return a part of their proceeds to anybody, including municipalities, for good public purposes). Nevertheless, the extreme cases of municipalities and administrative districts of municipalities with extended powers discussed in this paper also had similar positions on the gambling market in the Czech Republic in the following years.
By analysing the spatial distribution of EGMs and casino games we identified regions that are characterised by a very high concentration of EGMs and casino games, which can be considered as objects for further research. Specifically, these are the municipalities situated near the Czech-Austrian and Czech-German borders, where the concentration of gambling is also associated with higher tax revenues obtained from gambling. Another direction is local-level research of gambling in cities with large populations, which are the most affected by gambling. These results could be compared with the social status, crime rates and other selected social pathological phenomena of the population of these areas. The last but definitely not the least direction for further research would be to focus on the different gambling politics of the cities and to study the potential effect of regulations of gambling on the suburban areas and hinterlands of large cities.

Acknowledgement

The paper was supported by the project of the specific research MUNI/A/1419/2016: “Integrated research of environmental changes in the landscape sphere II” and by the Palacký University Internal Grant Agency (IGA_PrF_2016_026): “Landscape and people: Geographic analysis of environmental changes”.

References:


Please cite this article as:
Coping with employment issues through commuting: Evidence from Central Russia

Maria GUNKO a *, Tatyana NEFEDOVA a

Abstract

Commuting to work has become a widespread practice over the last several decades. Some relatively strong employment issues in rural Russian municipalities, as well as a significant wage gap between them and the major cities, fosters daily commuting as well as commuting with long periods of time spent at employment locations, known in Russia as otkhodnichestvo. Based on official statistical data and our own surveys conducted in Tula and Kostroma regions, we analysed the spatial differentiations between the share of commuters in the employed population of rural municipalities, as well as the variations in individual socio-economic characteristics between rural commuters and non-commuters, in general, and between the commuters themselves. Our analysis of the latter characteristics suggests that there is in fact more difference between the rural commuters and non-commuters in Russia than between daily commuters and those engaged in otkhodnichestvo, indicating similar motives and, perhaps, even personalities of commuters. Giving the identified differences, commuting on a larger scale does not seem to be a universal tool to cope with employment issues, as some population groups are more likely to engage in it than others.

Key words: Commuting, long-distance commuting, rural, post-socialist, Central Russia

Article history: Received 15 August 2016; Accepted 2 February 2017; Published 30 June 2017

1. Introduction

With improved transport systems, commuting to work is becoming a widespread practice in European societies (Bäckström, Sandow and Westerlund, 2016; Cassel, Macuchova, Rudholm and Rydell, 2013; Hofmeister and Schneider, 2010; Sandow, 2008, 2011). Commuting is caused by the variations in housing prices, wages and employment opportunities, and acts “as an equilibrating mechanism on the labour and housing markets” (Sandow, 2011, p. 2). Commuting provides access to employment opportunities on a wide geographical labour market without changes in residential arrangements (Lück and Ruppenthal, 2010; Plusnin, Zausaeva, Zhidkevich and Pozanenko, 2013; Sandow, 2011). Therefore, it is often viewed as an individual strategy alternative to migration (Hogarth, 1987; Mkrtchyan, 2010; Nefedova, 2015b; Sandow, 2011).

One specific characteristic of most European countries is their relatively dense urban networks with small distances between settlements, allowing daily short and long-distance commuting. In this context, commuting with a longer period of time spent at the employment site, e.g. weekly or monthly commuting, is a less frequent practice used to cope with employment issues (Sandow, 2011). In Russia, however, with its vast territory, sparse settlement network, poor transport connectivity and, most importantly, a small number of high-wage destinations (Zubarevich, 2013), commuting with longer periods spent at employment locations is quite common (Mkrtchyan, 2010; Nefedova, 2015a, 2015b; Plusnin, Pozanenko and Zhidkevich, 2015; Plusnin et al., 2013; Saxinger, 2015). Furthermore, the practice is explicitly and implicitly encouraged by employers, as such a workforce is often cheaper than the local one (Saxinger, 2015).

Some authors have argued that a significant number of people who are engaged in various forms of commuting in Russia come from peripheral rural areas (Nefedova, 2015a; Plusnin et al., 2013, 2015), as such places face the strongest employment issues (Nefedova, 2008, 2013b). In Central and Eastern European countries, in general, and in Russia in particular, the disadvantages of rural regions established under socialism – such as underdeveloped infrastructure, weak enterprises, and a deficiency of skilled labour forces (Leibert, 2013; Nefedova, 2013b) – have been reinforced under neo-liberal conditions. Currently, many such localities face a lack of job opportunities, which contributes to the selective out-migration of mainly economically active groups of the local population (Hováth, 2008; Kashnitsky and Mkrtchyan, 2014; Wastl-Walter and Váradi, Veider, 2003; Wiest, 2015) and encourages different types of commuting.

a Institute of Geography, Russian Academy of Sciences, Moscow, Russia (*corresponding author: M.Gunko, e-mail: msgunko@igras.ru)
Despite the fact that commuting is widespread in contemporary Russia, the official statistical reports account for it in a poor manner. The existing statistical data do not permit one to confidently define the number of commuters in terms of the distance travelled to work, their commuting rhythms, and the period spent at employment locations – not to mention the individual socioeconomic characteristics of commuters. Research of these noted issues is primary conducted with qualitative methods or is based on simple estimates (Mkrtchyan, 2010; Nefedova, 2015b; Plusnin et al., 2013). Against this background, this contribution seeks to shed more light on the commuting of Russian rural residents. The primary aim is to identify variations in the individual socio-economic characteristics distinguishing commuters and non-commuters in general, as well as between the commuters themselves, depending on the time they spend at their employment locations. Researching the commuting phenomenon in Russia as a part of the wider subject of population mobility, allows us to add to knowledge of the features of polarised spatial development in the post-socialist and especially post-Soviet context. In addition, it provides an interesting insight on labour mobility within a country with a very sparse settlement network, testing the assumptions of contemporary commuting theory.

The analyses used in the research are based on the statistical data from The Population Survey on Employment Issues (ONPPZ, 2013) and the Russian Population Census (VPN, 2010), and our own surveys which targeted adults living in rural peripheral municipal districts¹ of the Kostroma and Tula regions, located in the so-called Nonchernozem² part of Central Russia.

The structure of the paper is as follows: the first section deals with the conceptualisation of commuting; subsequently, rural development and commuting in Central Russia, in particular in the Tula and Kostroma regions, are described; the next section develops the data and methodological approaches, followed by the results of the study; in conclusion, our discussion attempts to overview the issues at hand.

2. Conceptualising commuting

Migration is explained to a great extent in terms of relative opportunity (Gilmartin, 2008), a balance between gaining access to better life opportunities elsewhere and a place-bound stability (Thulin and Vilhelmsen, 2014), and is often triggered by adverse structural conditions (Chen and Rosenthal, 2008). There are many obstacles to changing place of residence, however, including structural barriers (e.g. disparity of housing prices) and subjective reasons, e.g. family and other social ties, place attachment (Barcus and Brunn, 2010; Hogarth, 1987; Massey et al., 1993). Nevertheless, to lead a more or less comfortable life, work is essential as a means to provide financial resources for most people. Therefore, some individuals overcome the impediments to taking employment outside their place of residence through commuting. Usually, a commuter is defined as an individual who crosses the administrative borders of a municipality when travelling to work ( Cassel et al., 2013).

Depending on the distance and time of travel, commuting as a practice may be subdivided into short- and long-distance commuting. In the literature, long-distance commuting is often defined as a door-to-door journey to employment location which takes at least 40 minutes ( Cassel et al., 2013; Sandow and Westin, 2010; Sandow, 2011). Given the distances and the underdeveloped transport infrastructures in Russia, ordinary daily journeys-to-work within metropolitan regions, where daily commuting is most common, take around 1–2 hours (Makhrova, Nefedova, Treivish, 2013). Therefore, a high proportion of commuters in Russia are long-distance commuters if classified by European standards. Long-distance commuting is the outcome of a job search process where longer distance is traded for higher wages (Bäckström et al., 2016; Cassel et al., 2013; Sandow and Westin, 2010; Sandow, 2011; Saxinger, 2015; So, Orazem and Otto, 2001). The impact of the labour market on commuting relates to individual social characteristics, professional skills and education. Numerous authors argue that better-educated, primarily white-collar professionals are more mobile and prone to long-distance commuting ( Cassel et al., 2013; Hogarth, 1987; Sandow, 2008). Their higher income levels reduce the relative travel costs (So et al., 2001). Furthermore, commuting, in general, and long-distance commuting, in particular, are gender imbalanced: on average, men commute to a larger extent than women in terms of the number and length of trips (Sandow and Westin, 2010; Sandow, 2011; Zhidkevich, 2013).

According to the duration of stay at the employment location, commuting may be performed daily or with a longer period of stay at that location, i.e. long-period commuting. Hogarth (1987) and Sandow (2011) argue that such commuting occurs when a person takes employment outside the comfortable temporal travel to work zone; therefore, it is logical to assume that the majority of long-period commutes are carried out over a long or very long distance. Long-period commuting is often viewed as an extreme, problematic strategy (Bykov, 2011; Carrington, Hogg, McIntosh and Scott, 2012; Nefedova, 2015b). This type of commuting implies that a person lives in some type of lodging, for instance, during the week and returns home for the weekend (Hogarth, 1987). Other regular commuting rhythms are also possible, i.e. monthly or seasonally (Plusnin et al., 2013; Saxinger, 2015). Long-period commuting may also be irregular, i.e. commuting with a sporadic rhythm. The phenomenon combining various types of long-period commuting (e.g. weekly, monthly), as well as the “fly-in, fly-out” mode of work (Perry and Rowe, 2015; Langdon, Biggs and Rowland, 2016; Saxinger et al., 2016) has become known in Russia as otkhodnichestvo (or vakhta). The term “otkhodnichestvo” originates from the old Russian word meaning “to depart”, and was first used in the Russian Empire to define peasants who worked outside of their home regions during the cold seasons. This type of labour mobility has been widespread for several hundreds of years and disappeared only in the 1930s after collectivisation (Plusnin et al., 2015, 2013). Since the late 1980s, however, overall population mobility has increased and otkhodnichestvo has again become a

---

¹ The lowest unit of administrative-territorial division in Russia with an urban (county town) or a rural settlement as an administrative centre
² As established in Russian geography, the name for European regions of Russia with relatively low fertile soils and located mainly around and to the north of Moscow metropolitan region.
significant practice, used to ensure the economic welfare of households and to provide workforce for the major cities and remote Northern-Eastern regions which specialise in natural resource extraction (Florinskaya, 2006; Plusnin et al., 2013; Saxinger, 2015; Shabanova, 1992).

3. Rural development and commuting in Central Russia

3.1 Rural development as a prerequisite to commuting in Central Russia

The differential development between dynamic, growing metropolitan areas and rural or old industrial regions experiencing processes of shrinkage and economic decline in many Central and Eastern European countries, including Russia, is often associated with the changed political system (Benedek and Moldavan, 2015; Fischer-Tahir and Naumann, 2013; Kühn, 2015; Lang, 2012). Territorial development is largely inertial, and the present inequalities which formed under socialism became the starting point for shaping inequalities under market conditions (Nefedova, 2008, 2013a; Smith, 1996). The planned economy had created and protected economic structures which are obsolete in the globalised neo-liberal economy. Market liberalisation and the emergence of competition for development resources only reinforced the differentiations in structural conditions and power relations.

The long-term concentration of population and economic activity in a handful of sparse foci in Russia has led to the formation of vast physically inaccessible, economically backward and depopulating areas of the periphery – even in the ‘heart’ of Central Russia (Gritsay, Ioffe and Treivish, 1991; Ioffe, Nefedova and Zaslavsky, 2004, 2006). The socioeconomic development of such territories, which were poorly-equipped with infrastructure and which were uncomfortable for living already in the Soviet times, was based on the idea that human resources are inexhaustible (Nefedova, 2008, 2012). In the second half of the twentieth century this perception was fuelled by the ‘euphoria’ from oil revenues, which allowed the introduction and subsequent maintenance of expensive, unprofitable projects. Population development changed, low fertility rates and migration fostered rural depopulation. In the late Soviet period, crises could be traced to the labour-intensive economic activities located in rural districts (Ioffe, 1990). These unresolved development issues have had significant negative impacts in the post-socialist period. Many rural enterprises did not survive the harsh climate of global competition (Smith and Timar, 2010; Zubarevich, 2013). In the case of Central Russia, agriculture was for a long time the main sphere of employment in rural regions and districts. Unlike Central and Eastern European countries, new spheres of employment, such as services or tourism, hardly emerged in Russia’s almost monofunctional rural areas (Nefedova, 2013). New enterprises in rural areas are still primarily specialised in agricultural activities, and are then located in southern regions which are favourable to agriculture or close to large centres to ensure demand for their produce. In the majority of Nonchernozem regions with less favourable natural conditions, agricultural enterprises balance on the brink of bankruptcy. Some of them try to preserve employment with the payment of minimum wages, which in fact resemble unemployment benefits (Zubarevich, 2013). Therefore, some individuals in these areas end up in a situation of long-lasting unemployment and impoverishment (Nefedova, 2013a). As income sources have changed, households use different strategies to cope with the adverse structural conditions, often increasing the economic gap between them (Wegren, 2014).

Research on mobility in Russia indicates that, since the late 1990s, migration from structurally weak areas has been largely substituted by different types of commuting (Mkrtchyan, 2010). People are less prone to migrate due to the significance of local social ties in the everyday lives of a Russian household, which among other possibilities, promotes informal practices as a source of a household’s income (Round and Williams, 2010), as well as the increasing disparity in housing prices between places of in- and out-migration (Nefedova, 2015b). The unemployment and wage differences between municipalities – which for blue collar workers may be around 3–4 times and for white collar workers up to 10 times, depending on home and employment locations (Nefedova and Treivish, 2014; Saxinger, 2015) – promote the participation of rural residents in commuting to ensure their households’ economic well-being.

3.2 Commuting in the Kostroma and Tula regions

Longstanding field research carried out in the rural peripheries of Nonchernozem regions has indicated that from one-fifth to one-half of officially registered rural dwellers are not engaged in any formal economic activity in their place of residence (Nefedova, 2013; Pokrovsky and Nefedova, 2014). While some constantly work or study in cities, others are regularly or occasionally engaged in various types of commuting. The data from The Population Survey on Employment Issues (ONPPZ, 2013) provide some insights on the number of commuters among the economically active population and the location of their employment. It presents data for three categories of workers: (1) working in the place of residence; (2) working outside the place of residence but within region of residence; and (3) working outside place and region of residence. Such a classification, unfortunately, does not permit us to confidently define daily commuters and otkhodniki among workers. According to the available data (ONPPZ, 2013), the overall number of various types of commuters in the Kostroma region is about 30 thousand people. For those working outside Kostroma region, the primary employment locations are Moscow, Moscow region, Yaroslavl, Saint-Petersburg and Leningrad region (Tab. 1).

Using data from the latest Russian Population Census (VPN, 2010), which also does not differentiate daily commuters and otkhodniki, we have estimated the shares of intraregional commuters and commuters to other regions in the employed population of Kostroma region. These shares were further analysed according to the remoteness of commuters’ place of residence (municipality) to the regional capital, i.e. Kostroma city (Fig. 1). Remoteness is measured as the order of municipality’s neighbourhood to the regional capital, whereby zone 1 are municipalities directly bordering it. As it may be seen, the largest share of commuters working outside Kostroma region is characteristic not for peripheral districts with the strongest employment issues, but for the suburban ones. This finding corresponds to that of McQuaid, Greig, and Adams (2001): job seekers living in remote communities with low population density and higher

3 People engaged in otkhodnichestvo
unemployment levels are less willing to commute far to work. Slightly over 40% of Kostroma suburban residents commute, among them over 50% commute outside Kostroma region. Beyond the suburbs of Kostroma city, the share of those working outside the region does not exceed 7%. A significant proportion of the local rural population is comprised of intra-regional commuters who work not only in Kostroma but also in smaller cities of the region, e.g. Manturovo, Sharya.

A different situation may be viewed in the Tula region, where the estimated number of commuters was about 80 thousand people (ONPPZ, 2013). The vast majority of them work in Moscow and Moscow region (Tab. 1). Research suggests that in the districts of Tula region, bordering the Moscow region, every third person works in Moscow or the Moscow region (Nefedova, Averkieva and Makhrova, 2016). While Tula city attracts commuters primarily from its own suburbs (Fig. 2).

Such variations of commuters’ employment locations between Kostroma and Tula regions may be explained by the fact that Tula region borders the Moscow region, and some enterprises located in that region offer corporate transportation to commuters from nearby districts of Tula region (Gunko, 2015; Nefedova, Averkieva and Makhrova, 2016).

Tab. 1: The destinations of commuters, including out-of-town zone, to other regions from Kostroma and Tula regions
Note: *For Tula region the share of commuters to Saint-Petersburg and Yaroslavl is insignificant and is, therefore, presented in the category ‘other’. Source of data: ONPPZ, 2013

![Fig. 1: The share of intra-regional commuters and commuters to other regions in the employed population of Kostroma region. Source: VPN, 2010; draft by T. Nefedova; design by M. Gunko](image1)

![Fig. 2: The share of intra-regional commuters and commuters to other regions in the employed population of Tula region. Source: VPN, 2010; draft by T. Nefedova; design by M. Gunko](image2)
4. Research design

4.1 Case study: rural municipal districts

The case study districts are located within the Nonchernozem area of Central Russia, which is infamous for the presence of persistent, long-lasting and reinforcing negative effects on the local labour markets after the transition from the planned to a market economy (Gunko, 2015; Nefedova, 2013a; Zubarevich, 2013). Four typical peripheral rural districts were selected: they are peripheral both in terms of spatial (i.e. remote from large cities) and aspatial features, e.g. they are depopulated areas which are heavily subsidised from regional and federal budgets (Copus, 2001; Kühn and Weck, 2012; Plöger and Weck, 2014). Makarievsky (7,616 rural residents) and Shariyinsky districts (9,564 rural residents) are located in Kostroma region; Efremovsky (23,254 rural residents) and Suvorovsky districts (16,823 rural residents) in the Tula region. Giving the large area of the case study rural districts (Makarievsky – 1,066 sq.km, Efremovsky – 1,649 sq.km, Shariyinsky – 3,993 sq.km, Makarievsky – 4,850 sq.km), as well as the low density of paved roads within them (BDPMO, 2015), a car trip from the most remote rural settlements to the districts’ administrative centres often takes over an hour.

Three of the four districts have only one urban settlement (6,695 people in Makariiev, 23,914 in Sharia, 36,161 in Efremov), which serves as their administrative centre. Suvorovsky district has two urban settlements (Suvorov – 17,615 people and Chekalin – 965 people). The population of most rural settlements (from 62% in Efremov district to 82% in Sharia district) does not exceed 50 people. Additionally, long-term rural depopulation (from – 15% in Efremovsky district to – 45% in Makarievsky district in the post-Soviet period) contributes to the relative increase of the smallest rural settlements’ share in the total population.

Employment in the case study rural districts has traditionally been in agriculture, state-financed public services (e.g. education, healthcare) and public administration, accompanied by the emergent small-scale entrepreneurial activities, primarily trade, in the post-socialist period. Along with these activities in the districts of Kostroma region, which is the richest region of Central Russia in terms of timber resources, the harvesting and primary processing of wood is also widespread. The present unemployment rates in the case study rural districts resemble each other and do not exceed 3–5% (BDPMO, 2015). This indicator is measured by the Russian statistical agency, as the share of economically active individuals who have voluntarily registered in the

---

Fig. 3: Rural municipality case study areas. Source: draft by M. Gunko; design by A. Medvedev

---

4 All figures in Section 4.1 are according to BDPMO (2015)
tertiary public unemployment services. Therefore, a number of researchers have argued, based on their own surveys, that the real level of unemployment in Nonchernozem rural districts sometimes exceeds 20% (Aseeva, 2013; Bondarenko and Grudneva, 2010; Gulyaeva and Grigorieva, 2012; Gunoko, 2015).

4.2 Data

Our study analyses the differences between the socio-economic characteristics of daily commuters and otkhodniki, as well as between overall commuters and non-commuters from the rural districts. The target group for the study was defined as adults (between 18–65 years old), residing in Makarievsky, Shariyinsky, Suvorovsky, and Efremovsky rural municipal districts. Data were collected between May and August 2014 in the form of a questionnaire survey. The questionnaires were handed out to randomly selected people who fitted the age criteria and filled out on the spot.

The questionnaire was designed to collect quantitative information regarding type of employment (commuter or non-commuter, daily commuters or otkhodniki) and items relevant for the study socio-economic characteristics of the individual and household. These characteristics included age, gender, family status, type of employment of the spouse (commuter or non-commuter), presence of children in the household, number of children, and the type of transport used to get to work.

Overall, people were open to participation: the non-response rate was around 5%. A total of 257 people were surveyed in the four municipalities. Among the surveyed, 121 were male and 136 female. The median age of the respondents was 44 years. The majority, 227 respondents, were married. 233 respondents had at least one child who was a minor. 92 respondents were commuters, 52 of them – otkhodniki.

4.3 Methodological approach

In our analysis, we dealt primarily with categorical variables. Therefore, multiple logistic regression models were estimated in order to statistically analyse and compare the socio-economic characteristics of commuters and non-commuters (Model 1), as well as daily commuters and otkhodniki in the chosen municipal districts (Model 2). In multiple logistic regression, the dependent variable is binary – coded as 0 and 1. The categorical independent variables in our model were coded into binary dummy variables. The only numeric variable, age, for the ease of the analysis was also coded into a dummy variable, where 0 represented younger (18–40 years old) and 1 older (41–65 years old) working age groups.

Eight of nine independent variables included in the models are informed by the findings of previous commuting studies (e.g. Cassel et al., 2013; McQuaid et al., 2001; Sandow and Westin, 2010; Sandow, 2008). The inclusion of the ninth independent variable, engagement in supplementary economic activity, was dictated by the need to consider the specificity of socio-economic conditions of rural life in post-Soviet Russia. As argued by A. Smith and Timar (2010), there is a broad set of widespread practices undertaken by households in the post-Soviet regions in response to economic marginalisation. The income from performing such activities is often kept ‘off-record’ not to pay taxes and may constitute up to one-third of the family budget (Gunoko, 2015). Therefore, the engagement of the household members in supplementary economic activities in the place of residence may also act as a factor affecting the choice to commute. All variables used in the analysis are summarised in Table 2.

5. Results

Prior to conducting the main analysis, we checked for the effect of location on commuters’ numbers and features. The effect was found to be statistically insignificant, suggesting that there is little variation between the number and socio-economic characteristics of commuters in the four districts under study. Then we estimated two logistic regression models: Model 1, ‘commuters vs. non-commuters’, Model 2, ‘daily commuters vs. otkhodniki’. The output of Model 1 is presented in the Table 3, the reference group is commuters (non-commuters = 0). The output for Model 2 is presented in Table 4, the reference group is otkhodniki (daily commuters = 0).

In Model 1 (commuters vs. non-commuters), of the 9 independent variables, 5 have a significant effect on the dependent variable. In Model 2 (daily commuters...
supplementary economic activities within the settlement of otkhodniki. The first explanation may be that otkhodniks are primarily engaged in supplementary economic activity. The results suggest that having any type of supplementary economic activity which brings income, reduces the likelihood that the respondent is a commuter. Moreover, households of daily commuters are on average more likely to be engaged in such activities than households of otkhodniki. Such results may be interpreted in two ways, and we believe that to a certain degree both explain the obtained result. The first explanation is that otkhodniki themselves simply cannot engage in supplementary economic activities within the settlement of residence due to long periods being away from home, while their spouses are also short on time, being overloaded with household-related work which they have to perform alone. The second explanation may be that otkhodniki are an extreme practice to overcome economic marginalisation, utilised only when all other options in the place of residence have been exhausted. The findings of previous qualitative studies on otkhodniki in Russia implicitly support this hypothesis. The majority of otkhodniki are either manual workers, e.g. construction workers, or people working in the service sector, e.g. security guards, retail store workers, taxi drivers or nannies (Plusnin et al., 2015, 2013; Savoskul, 2013); i.e. they are less legible and are willing to occupy those employment niches which do not require a special education and are generally not taken up by the local population. This is not a surprise, as the main destinations for otkhodniki are the metropolitan regions of Moscow and Saint-Petersburg, as well other large cities where the spheres of highly-qualified employment are more competitive than their low-skilled counterparts. Moreover, having a university degree is negatively associated with the chances that the respondent is a commuter, although it is not a significant factor affecting employment type in both models. This result is contrary to those obtained for Western European countries (Cassel et al., 2013; Hogarth, 1987; Sandow, 2008), where higher educated individuals are more prone to commute and to commute over longer distances.

The second independent variable with a statistically significant effect both in models is the engagement in supplementary economic activity. The results suggest that having any type of supplementary economic activity which brings income, reduces the likelihood that the respondent is a commuter. Moreover, households of daily commuters are on average more likely to be engaged in such activities than households of otkhodniki. Such results may be interpreted in two ways, and we believe that to a certain degree both explain the obtained result. The first explanation is that otkhodniki themselves simply cannot engage in supplementary economic activities within the settlement of residence due to long periods being away from home, while their spouses are also short on time, being overloaded with household-related work which they have to perform alone. The second explanation may be that otkhodniki are an extreme practice to overcome economic marginalisation, utilised only when all other options in the place of residence have been exhausted. The findings of previous qualitative studies on otkhodniki in Russia implicitly support this hypothesis. The majority of otkhodniki are either manual workers, e.g. construction workers, or people working in the service sector, e.g. security guards, retail store workers, taxi drivers or nannies (Plusnin et al., 2015, 2013; Savoskul, 2013); i.e. they are less legible and are willing to occupy those employment niches which do not require a special education and are generally not taken up by the local population. This is not a surprise, as the main destinations for otkhodniki are the metropolitan regions of Moscow and Saint-Petersburg, as well other large cities where the spheres of highly-qualified employment are more competitive than their low-skilled counterparts. Moreover, having a university degree is negatively associated with the chances that the respondent is a commuter, although it is not a significant factor affecting employment type in both models. This result is contrary to those obtained for Western European countries (Cassel et al., 2013; Hogarth, 1987; Sandow, 2008), where higher educated individuals are more prone to commute and to commute over longer distances.

Being married increases the chances that the respondent is a commuter, supporting the previous findings of Bäckström et al. (2016) and Sandow (2011) that the decision to commute

vs. otkhodniki), only two independent variables have a significant effect on the dependent – gender and engagement in supplementary economic activity. This may indicate that there is less difference between the socio-economic characteristic of different types of commuters than between commuters and non-commuters in typical peripheral rural districts of Central Russia.

The first point to note is that gender has a statistically significant effect on the type of employment. On average, male respondents are more prone to commute, and especially to commute over longer distances and with a long period of stay in employment locations. This phenomenon was found in many previous studies (e.g. Sandow and Westin, 2010; Sandow, 2011). It may be explained both by the fact that women are more likely to find employment closer to home (Hanson and Pratt, 1995) and by the uneven share in domestic responsibilities, such as child-care and other household-related work (Hjorthol, 2000).

The second independent variable with a statistically significant effect in both models is the engagement in supplementary economic activity. The results suggest that having any type of supplementary economic activity which brings income, reduces the likelihood that the respondent is a commuter. Moreover, households of daily commuters are on average more likely to be engaged in such activities than households of otkhodniki. Such results may be interpreted in two ways, and we believe that to a certain degree both explain the obtained result. The first explanation is that otkhodniki themselves simply cannot engage in supplementary economic activities within the settlement of residence due to long periods being away from home, while their spouses are also short on time, being overloaded with household-related work which they have to perform alone. The second explanation may be that otkhodniki are an extreme practice to overcome economic marginalisation, utilised only when all other options in the place of residence have been exhausted. The findings of previous qualitative studies on otkhodniki in Russia implicitly support this hypothesis. The majority of otkhodniki are either manual workers, e.g. construction workers, or people working in the service sector, e.g. security guards, retail store workers, taxi drivers or nannies (Plusnin et al., 2015, 2013; Savoskul, 2013); i.e. they are less legible and are willing to occupy those employment niches which do not require a special education and are generally not taken up by the local population. This is not a surprise, as the main destinations for otkhodniki are the metropolitan regions of Moscow and Saint-Petersburg, as well other large cities where the spheres of highly-qualified employment are more competitive than their low-skilled counterparts. Moreover, having a university degree is negatively associated with the chances that the respondent is a commuter, although it is not a significant factor affecting employment type in both models. This result is contrary to those obtained for Western European countries (Cassel et al., 2013; Hogarth, 1987; Sandow, 2008), where higher educated individuals are more prone to commute and to commute over longer distances.

Being married increases the chances that the respondent is a commuter, supporting the previous findings of Bäckström et al. (2016) and Sandow (2011) that the decision to commute

### Table 3. Output for Model 1: ‘commuters vs. non-commuters’ (significance level = 0.05)

|                | Estimate | Std. Error | z value | Pr(>|z|) | Sig. |
|----------------|----------|------------|---------|----------|------|
| (intercept)    | 15.412   | 1455.397   | 0.011   | 0.991    |      |
| gender         | -2.376   | 0.718      | -3.306  | 0.000    | yes  |
| education      | -0.153   | 0.629      | -0.244  | 0.807    | no   |
| family status  | 13.128   | 1455.398   | 0.009   | 0.992    | no   |
| type of spouse’s employment | 1.256 | 0.839 | 1.497 | 0.134 | no |
| children       | 0.211    | 1.183      | 0.177   | 0.880    | no   |
| number of children | 1.053 | 0.163 | 0.906 | 0.365 | no |
| supplementary economic activity | -1.987 | 0.850 | -2.337 | 0.019 | yes |
| age            | 0.512    | 0.648      | 0.791   | 0.428    | no   |
| transportation to work | -0.612 | 0.581 | -1.052 | 0.292 | no |

### Table 4. Output for Model 2: ‘daily commuters vs. otkhodniki’ (significance levels = 0.05)

|                | Estimate | Std. Error | z value | Pr(>|z|) | Sig. |
|----------------|----------|------------|---------|----------|------|
| (intercept)    | -1.556   | 1.533      | -1.015  | 0.310    |      |
| gender         | -2.581   | 0.516      | -4.999  | 0.000    | Yes  |
| education      | -0.806   | 0.453      | -1.779  | 0.075    | No   |
| family status  | 2.523    | 1.106      | 2.280   | 0.022    | Yes  |
| type of spouse’s employment | 0.084 | 0.533 | 0.158 | 0.874 | No |
| children       | 0.221    | 1.173      | 0.189   | 0.850    | No   |
| number of children | 2.314 | 0.499 | 4.632 | 0.000 | Yes |
| supplementary economic activity | -1.298 | 0.433 | -2.992 | 0.002 | Yes |
| age            | -1.782   | 0.450      | -3.959  | 0.000    | Yes  |
| transportation to work | -0.282 | 0.430 | -0.656 | 0.512 | No |
is associated with family status and is carried out more often by households than by individuals. The presence of a minor child in a household also increases the likelihood that the respondent accepts the commuting strategy. As argued by Green (1997), households with children are less prone to migration and in adverse employment conditions substitute it by commuting. Although, in both of our models this factor is not statistically significant, having more than one child, which implies an increase in the minimum required living costs, is a statistically significant factor favouring commuting in Model 1.

Younger people are on average more likely to become commuters (Model 1). Besides the obvious explanation associated with the high labour expenditures of commuting, the obtained regularity could also be connected to the presence of minor children in the household since in Russia, especially in rural regions, the average age of parents having their first child is below 30 years. In Model 2 age has no significant effect on the variables.

6. Discussion and conclusions

The present inequalities as well as the current economic crisis in Russia caused by the decrease in oil prices, sanctions and counter-sanctions, can be projected to an increased spatial socio-economic polarisation. In such situations, rural regions are especially vulnerable, as the high unemployment rates there affect the intensification of migration and commuting. In the two regions analysed in this study, 15–60% of rural dwellers from the different municipalities commute to work. The variations in the share of commuters among the employed population is related to the municipality’s location. Suburban residents were more prone to commute, and more likely to be otkhodniki, than residents of more remote peripheral municipalities who on average commute over shorter distance. The primary commuting destinations from Tula and Kostroma regions include the metropolitan regions and major cities of European Russia, mirroring general migration trends (Kashnitsky and Mkrtchyan, 2014; Zayonchkovskaya and Mkrtchyan, 2009).

Commuting and especially otkhodnichestvo is not just a specific form of labour mobility, but rather a peculiar lifestyle characterised by multi-localities (Plusnin et al., 2015; Saxinger, 2015) and implying a new social status and economic behaviours of those who perform it. Little is still known, however, about the individual socio-economic characteristics defining commuters from the bulk of the population in Russia.

The aim of our study was precisely to analyse the differences between the socio-economic characteristics of rural commuters and non-commuters in contemporary Russia, as well as the difference between the two types of commuters – daily commuters and otkhodniki, drawing on the empirical evidence from rural municipalities in Central Russia. Some variables associated with commuting, including gender, age and the presence of minor children, as defined in previous studies (e.g. Cassel et al., 2013; Sandow, 2011), have been confirmed in this research. One significant difference is in terms of education. Having a university degree has an opposite effect on Russian commuters than on those analysed in previous studies with empirical evidence from European countries (e.g. Cassel et al., 2013; Hogarth, 1987; Sandow, 2008). Russian commuters seem to be less educated than non-commuters. The observed variation, though statistically insignificant, could mean that people with a university degree either migrate or it is easier for them to find employment in the place of residence.

Our analysis suggests that there are in fact more differences between commuters and non-commuters in Russia than between daily commuters and otkhodniki. This may not only indicate that the motives of the latter are the same but also that there are similarities between them in terms of personality. Plusnin et al. (2015) argue that commuters, and especially otkhodniki, are proactive individuals, and this is also supported by Chepurenko (2010) who indicated that many potential entrepreneurs choose to be engaged in daily commuting or otkhodnichestvo, rather than start their own business due to the difficulties of starting and maintaining a business in Russia. Such issues, especially with respect to commuters’ personalities, however, require further research.

Being utilised by specific social groups means commuting on a wide scale is not a universal remedy to cope with employment issues, and it seems to be more a temporal tactic than a sustainable strategy. Even for those currently engaged in commuting and otkhodnichestvo, it is hard to predict how long will they be willing to undertake such work, especially in a situation when economic crisis rapidly alters the conditions of labour markets in the main destinations of employment. Furthermore, places which attract commuters clearly receive a lot of benefits, such as low-cost labour, additional tax revenues and demand for goods and services, while the commuters’ home regions suffer from workforce shortages in the sphere of low-wage employment as well as the reductions of tax and non-tax revenues in municipal budgets. Instead of mitigating adverse employment conditions, commuting contributes to the further deterioration of the situation on the rural labour markets reinforcing the overall spatial socio-economic polarisation of Russian regions. Though, it must be noted, that effects of commuting are far more complex than it may seem at a first glance. The potential positive effects, primarily in the social sphere resulting from high population mobility, are less visible and harder to research. However, they do require close attention in further studies.

Acknowledgement

A special gratitude is expressed to all survey participants. The authors would also like to thank Ilya Kashnitsky and Andrey Medvedev for the help with data visualisation. The research was supported by the Russian Scientific Foundation, grant No. 14-18-00083: “Geography of Recurrent Population Mobility within the Rural–Urban Continuum”.

References

Moskvi. In: Zayonchkovskaya Z. [eds.]: Immigranti v Moskve (pp. 18–44). Moscow, Tri kvadrata.


Please cite this article as:

Fig. 8: Newly established Casino Victory in Popůvky municipality (a suburban area of Brno city), built on the regenerated brownfield site (Photo B. Frantál)

Fig. 9: Las Vegas, one of the most popular world centres of gambling (Photo Z. Szczyrba)

Illustrations related to the paper by D. Fiedor et al.