REGIONAL STATISTICS, 2013, VOL. 3: 3–29 DOI: 10.15196/RS03101

IWAN J. AZIS[a*]

Complex Interplay of Factors in Institutional Model of Decentralization: Theory and Application

Abstract

The Institutional Model of Decentralization (IMD) is elaborated and used to explain two things: first, how the hypothesized improvements in efficiency and growth after decentralization may fail to materialize; second, how the interplay among economic, administrative and institutional factors affect the welfare outcome of decentralization, given the widespread local capture following political decentralization. Rather than exerting direct effects, however, the mechanism is complex, involving intangibles. When applied to actual cases in some regions, a particular method capable of capturing complex inter-relations and quantifying intangibles is therefore used. It is revealed that people's participation plays the most critical role in reducing capture while simultaneously maximizing welfare. As the quality of local leaders is found decisive in influencing the outcome, a typology of leaders is subsequently constructed.
JEL Classification: B52, C93, E02, O17, P36, P46, R13

Keywords:

Introduction

Decentralization is a multi-dimensional phenomenon, encompassing several interconnected aspects. Its theoretical supports originate in the informational advantage and coordination (policy enforcement) capability of local government. Although informational advantage can be secured by adopting a pro-market policy, a market system alone may not be sufficient to establish an effective coordination at the local level unless the decision making is decentralized. A more decentralized system, particularly on fiscal front, is also superior for promoting economic growth (United Nations 1991, Oates 1994, Bruno and Pleskovic 1996).[1] No wonder the World Bank embraced decentralization as one of the major governance reforms on its agenda (World Bank 2000, Burki–Perry–Dillinger1999).

Yet, the experience in many countries shows that the performance after decentralization has not always been consistent with the promise. Growth can be lower, and overall welfare conditions may not improve, if not worsened. Imperfections in local provision and poorly trained local bureaucrats are among the suggested reasons

a) Cornell University 213 West Sibley Hall Ithaca, NY 14853, USA, E-mail: ija1@cornell.edu.

1 Other justifications for decentralization include: raising efficiency through reduced transaction costs, diffusing social and political tensions, strengthening people's participation, and ensuring political and cultural autonomy.

(Prud'homme 1994, Tanzi 2002). Problems due to lack of coordination in extracting bribes at the local level can also lead to 'excess' rent extraction (Shleifer–Vishny 1993), although some argued that corruption should be more difficult to commit under decentralization (Huther–Shah 1998). But the most significant risk of decentralization is the spread of *local capture* especially in regions with high degree of income disparity. Although the verdict regarding the relative proneness of local and national governments is still out (Bardhan–Mookherjee 2005), in general the likelihood of capture by elites is greater at the local than at the national level. The possibility of power sharing between contesting parties is typically smaller at the local than at the national level. Some analysts consequently suggest that the justification for decentralization should be based on the political economy explanation (Besley–Coate 1999), and to be successful decentralization should entail democratic, fiscal, and administrative components (Manor 1999, Binswanger 1999).

There is a large number of studies on fiscal federalism and administrative decentralization (transfer of authority to local governments). Literatures on institutional economics dealing with political decentralization are also numerous. Yet, theoretical and empirical work on how these different dimensions interact to affect growth and welfare is curiously scanty. By using the *Institutional Model of Decentralization* (IMD), in this paper I try to delineate the interactions between institutions, regional growth and a more broadly-defined welfare.[2] The starting point is to delve into the theoretical concept of regional incentives by way of comparing 'rewards' and 'punishment' for adopting growth-enhancing policy. Once the basic concept is established, a more variety of institutional factors to reflect the property of local accountability are introduced, ranging from local capture, people's participation, initial conditions, and the quality of local leaders. Incorporating these factors in the analysis of decentralization constitutes the essence of IMD. The empirical application based on a series of field survey in one of the emerging economies that went through a major decentralization program is discussed in the last Section.

Incentives for Regional Growth: Rewards and Punishment

The basic premise of decentralization-growth nexus is that, local governments are more efficient at providing infrastructure and public services compared to the higher levels of government (Oates 1972). Greater efficiency is thus at the center of the relation between decentralization and higher rates of economic growth (Martinez-Vazquez–McNab 2003). But overall efficiency is not always aligned with private efficiency. With enhanced authority after decentralization, local leaders may put private benefits above social benefits, depending on how they perceive the implication on their probability of staying

2 *Institutional Model of Decentralization* (IMD) was first explained in Azis (2008) and elaborated further in Azis (2010). One of the consequential issues emerged in IMD is the relation between policy and institutions. Negative welfare effect can be the result of wrong policies, but it can also be the product of right policies with wrong institutions. In some cases, policy matters more than institution (Henry–Miller 2008), in others institution matters more (Rodrik–Subramania–Trebbi 2002, Easterly–Levine 2002). Referring to the case in Sub-Saharan Africa, Sachs (2003) argues that institution matters but not for everything.

in power. When private benefit rules, incentives to foster growth become secondary, and when private motives produce a detrimental impact on regional resources, growth falters.

Let p^g = probability that local government stays in power if it fosters growth; and p^c = probability that local government stays in power if it kills growth by intensifying local capture. The latter could happen because local governments have had few incentives either to resist capture or to rein in competition for rents (e.g., Bardhan and Mookherjee, 2005; and Shleifer & Treisman, 2005 for the case of Russia). Denote C for the benefits accrued to local officials through local capture; and R^r for regional-own revenues, the size of which is determined by the local rates that include both tax rates and other revenue collection rates, θ, and regional output Y^r. The share of central government revenues (from additional growth) going to local governments is denoted by α. Thus, $\alpha.t.Y$ is the actual revenue received by local governments where $t.Y$ is central government's total revenue. How much local government will value growth is therefore proportional to $\alpha.t.Y$. Central government can use α as the "carrot" in promoting regional growth.

Define $PROB = p^g / p^c$, the value of which depends on whether local officials are appointed (by the center) or locally elected. If they are appointed, then the center can presumably choose $PROB$ freely and make it as high as it wants. If they are elected locally, the outcome depends on the ability of central government to affect the outcome of the election, e.g., not endorsing specific candidates. Thus, the center can use $PROB$ as the "stick." If, however, the center has little control over the election outcome, and capture becomes an important factor, $PROB$ may be less than unity, i.e., local government may be more likely reelected if it kills growth than if it fosters it. Local governments chooses growth only if the incentive is higher than the incentive to obtain private benefits from local capture:[3]

$$p^g.(\alpha.t.Y + R^r) > p^c.C \qquad (1)$$

or

$$PROB.(\alpha.t.Y + \theta.Y^r) > C \qquad (2)$$

Another way to put it, local governments are more likely to choose growth under the following conditions: stronger "stick" (higher $PROB$), larger "carrot" (higher α), higher national growth potential (higher Y), more effective generation of national tax revenues (higher t), and higher local-own revenues $\theta.Y^r$. On the other hand, one can also focus on the efforts to lower private benefits from local capture, C, by attacking the negative factors such as corruption, weak legal system, and ineffective law enforcement.

The above formula provides a way of identifying a set of policies without singling out "non-economic factors" often cited just for convenience. In reality, however, not all countries can alter policies for either historical or political reasons. In such a case, the focus can be directed towards exploring new instruments to complement the existing

3 Social benefits (e.g., enhanced growth and welfare) and private benefits (e.g., resources appropriated for private use) can be generated simultaneously. Put in the context of motivation and incentive to obtain benefits, however, social and private benefits can be in a competitive mode. That is, local leaders' incentive to obtain private benefits from local capture may either forgo or exceed the incentive to obtain social benefits. Or, the reverse may hold as in inequality (2): i.e., local leaders' incentive to obtain social benefits (e.g., from enhanced growth and welfare) exceeds the incentive to obtain private benefits.

ones. Take the case of α . It is important to distinguish between ex-ante and ex-post α . If central and regional governments can commit to a tax sharing schedule or to a particular arrangement of center-local transfer, the two will be the same. But if central government is broke and desperately needs funds to keep down its deficit, the ex-ante α may be higher than the ex-post α . The one which is relevant to local governments' decisions is the ex-post α . If a proportion of national revenues has been fixed, a new fraction of revenues, say λ , can be introduced. The point is, central government can control a policy instrument that will function as the "carrot":

$$PROB.[(\lambda(Y^r)+\alpha).t.Y+\theta.Y^r]>C \qquad (3)$$

One could erroneously imply that by raising θ , *ceteris-paribus*, growth incentives could be enhanced. But higher θ can deter investment and growth of Y^r by discouraging investors to come and do business in the region (see also footnote 4). Thus, the level of Y^r can be inversely related to the size of θ .

Assuming no income leakage, consider the following:

$$Y^r = Y^r(K^r, L^r, N^r) \qquad (4)$$

where K^r, L^r, and N^r are capital, labor and other inputs, respectively. Decomposing the regional capital stock into: (a) initial stock adjusted by the depreciation rate $\overline{K_0}^r(1-\delta)$; and (b) the regional investment flow ΔK^r , and considering that ΔK^r is inversely related to the regional tax and other revenue rates θ ; that is, $\Delta K^r = f(\theta)$ where $\partial f(\theta)/\partial\theta < 0$, the total regional-own revenue is:

$$R^r = \theta.Y^r\{[\overline{K_0}^r(1-\delta)+f(\theta)], L^r, N^r\} \qquad (5)$$

A fundamental growth equation is thus obtained by combining (3) and (5):

$$PROB.\{(\lambda(Y^r)+\alpha).t.Y+\theta.Y^r[(\overline{K_0}^r(1-\delta)+f(\theta)), L^r, N^r]\}>C \qquad (6)$$[4]

Thus, central government can use the "stick" ($PROB$ and C) and the "carrot" (λ , α) to foster growth. Note also from (6) that a growth-oriented strategy at the national level (higher Y) can help fulfill the inequality. Absent of a stick-and-carrot system, the incentives for local leaders to obtain private benefits through local capture can be greater than the incentives to foster growth.

But the ultimate goal of regional development goes beyond just growth. A broadly-defined welfare is more appropriate to use. The problem is, once a multi-dimensional goal is considered, various trade-offs emerge, where different institutional factors play different roles in affecting welfare. A further complication appears with respect to the complex relations between local capture and the goal, where institutions including capture interact among themselves as well as with welfare. To establish such a more

4 Two important points worth to note. Since increased capital is often accompanied by increased imports (income leakage), the true R^r is likely lower than what is stated in (5). Notice also that if local governments fervently want to raise θ, regional-own revenues may at first increase. As regional investment begins to be affected adversely by higher θ, total revenues will decline. Thus, the policy choice concerning θ depends on the initial condition. If the current θ is so high that it lies to the right of optimal θ (defined as θ that gives the highest level of R^r), raising it further will kill growth. Reducing θ under such circumstances will raise not only the regional-own revenues but also investment flows and hence growth.

realistic system where institutions are endogenous, however, requires a model framework beyond just the fundamental growth equation as in (6). This task is taken next.

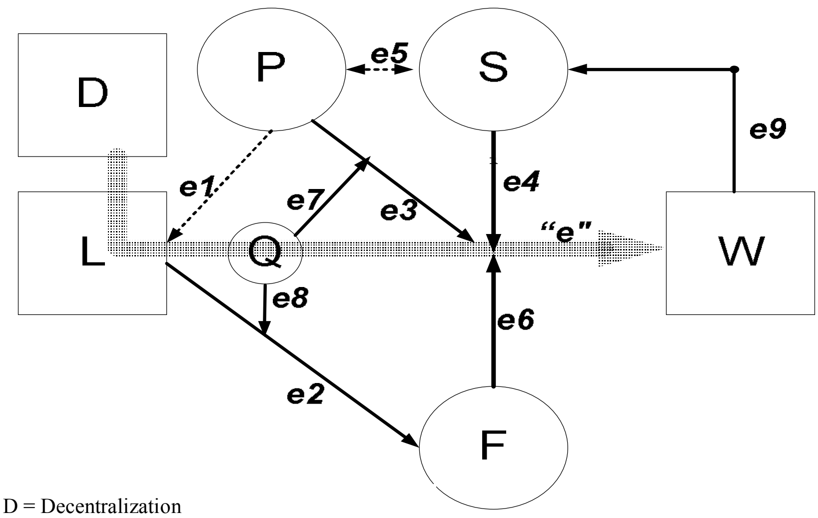
Endogenous Institutions and Regional Welfare

Theoretical Framework

Rules, organizations, beliefs, internalized norms, and implied regularity of behaviors which constitute "institution" define the incentive structure of societies and economies (North 1993, 1995). The system is equilibrium if the implied regularity of behaviors to follow the rules are best-responses to the beliefs and internalized norms that are formed by the implied regularity of behaviors (Greif 2006).[5] Many policies may fail to achieve their objectives because the institution in which these policies or rules were elements of was not in equilibrium.[6] It is therefore important to understand how various institutions interact among themselves and how the interactions influence the welfare outcome.

Figure 1

How Local Capture Affects Welfare: Full Framework



D = Decentralization
L = Local capture
W = Welfare effect of decentralization
S = Initial welfare condition
P = People's participation
Q = Quality of leader
F = size of local budget

5 Persistence or inertia, institutional path dependence, or steady-state equilibrium in institutional setting are among the terms used to describe the study of endogenous institutional change.

6 Equilibrium in a more practical term means that there is no individual or a group of individuals that has an incentive to deviate from an agreement or what is previously agreed. In other words, an institution that is in equilibrium consists of rules in which targeted individuals have incentives to follow the rules. Thus, a law regulating issuance of driving licenses may not be effective because bribing public officials renders it more profitable and time-saving.

A framework capturing the interplay of institutional factors is shown in Figure 1. Decentralization policy (D) with direct election for local leaders generates ‘local capture’ (L). However, the effect of L on local welfare (W) varies: in some regions the effect is positive (“positive local capture”), in others it results in a “negative local capture.” Among various factors that determine the effect of L on W, three stand out: initial welfare condition (S), people’s participation (P), and size of local budget (F). Note that the resulting W determines the subsequent level of initial condition S, that is, the steady-state level of initial welfare is influenced by any perturbations in the system that lead to changes in W, hence S in the subsequent period. The implied mechanisms also explicate the persistent gap between poor and rich regions observed in many countries.

The key question is, how the interplay of all these factors influences the way local capture influences welfare. The extent and severity of accountability problems and local capture depend on the following factors: (1) Pre-existing distribution of power at the local level, e.g., allocation of social and economic power within communities; (2) Lobby and campaign contributions by wealthier groups; (3) Fairness and regularity of elections; and (4) Transparency in local decision-making processes. Establishing these conditions may require institutional and bureaucratic reforms, yet it is precisely this type of reform that is most difficult to conduct. Overcoming institutional factors is always more difficult than choosing the policy itself. It is complicated, involving a strong path-dependence, and often frustrating. Absence of this reform, however, a higher local capture almost certainly produces lower benefits of decentralization.

Literature on institutional perspectives stresses the importance of *participatory* process. The degree of political participation differs between countries and regions. One of the most determining factors is the initial welfare condition or social structure represented among others by human development index (HDI) and the level of poverty and income inequality. Greater inequality and larger proportion of the poor imply a smaller fraction of informed voters or lower political awareness, i.e., upward mobility at lower end tends to raise political awareness more significantly than at higher end (concavity). When awareness is low, critical voices and the process of check-and-balance are constrained. This can limit the quality of public services and the welfare outcome of decentralization in general (Azis 2010).

The above are all associated with the *quality* factors. Each of them can be adversely affected by the intensity of local capture. While quality is important, however, the number and size of activities (e.g., public services) also influence the overall performance. The *quantity* that local government can generate depends not only on the budget size but also the management of it.[7] Ironically, under some circumstances the size of the budget can be positively affected by local capture, that is, if local leader is of Type-A (to be explained later), and local elites are powerful and wealthy. Under such circumstances, local policy makers can operate using resources in excess of the official

7 Revenue decentralization and central-local financial transfer without clear expenditure assignment are likely inconsistent with “money follow function” principle. They are not welfare enhancing, especially when the capacity of budget management is limited and are prone to corruption and overprovision.

budget. Thus, given *quality* factors, greater local capture can still be welfare-improving (relation e2 in Figure 1).

To reiterate the key problem, it is important to understand how the spread of local capture (L) during election determines the welfare outcome of decentralization (W). More particularly, how the effect is influenced by the extent of people's participation (P), the initial level of welfare (S), and the size of local budget (F). As argued earlier, the extent of participation is influenced by the initial welfare through informed voters and high political awareness.

Based on the analysis of implied behaviors in a coordination game (see Azis 2010 and Azis–Wiharja 2009), the role of leaders' quality (Q) is critical. However, in reality the effect of L on W associated with Q can be ambiguous; some leaders are able to take advantage of capture to augment local budget (Type-A leader), others may fail to do so (Type-B); see relation e8.[8] Also, in some cases local leaders are effective in motivating local citizens to participate in most development programs, while in others this may not be the case (relation e7).

The dynamics is captured among others through relation e9. For example, low initial welfare (S) as a result of negative local capture will negatively affect *W* and *S* in the subsequent period. Through relation e5, this may be associated with a low level of participation (P), creating a persistent evolution of low-welfare states and low-quality institutions (endogenous institutions). The possibility that local capture can generate positive welfare effects provides a more complex yet useful analysis with direct policy implications. It can be shown, for example, that policies to enable greater participation are superior to others because the welfare effect is higher given a level of capture.

Local capture *L*, participation *P*, and initial welfare *S* (poverty and inequality) represent the *quality* component of institutions. The size of local budget *F*, on the other hand, represents the *quantity*. Consider the following general form of welfare function:

$$W(.) = W(L, S, P, F) \quad (7)$$

Under a standard condition,

$$\partial W(.)/\partial L < 0; \qquad \partial W(.)/\partial S > 0; \qquad \partial W(.)/\partial P > 0; \qquad \partial W(.)/\partial F > 0.$$

Since F is affected by L, decomposing (7) into quality and quantity components:

$$W(.) = H(L, S, P).(F(L) \quad (8)$$

where $\partial H(.)/\partial L$ and $\partial F(L)/\partial L$ are marginal quality and quantity, respectively. Taking the derivative of (8):

$$\partial W(.)/\partial L = \partial F/\partial L \, H(.) + \partial H(.)/\partial L \, F(L) \quad (9)$$

8 It is wrong to suggest that the central and the provincial governments should retreat into minimalist role. On the contrary, they should play an activist role in conducting the necessary reforms. It is also the responsibility of the center to facilitate institutional supports for a successful decentralization. These include supplying technical services toward building local capacity, promoting mobilization of people in local participatory development, helping to set quality standards, auditing and evaluation, providing supra-local support to local finance (including being responsible in the coordination efforts to optimize externalities across regions), and investing, when necessary jointly with local government, in infrastructure.

In most cases, $\partial H(.)/\partial L < 0$. However, a good quality leader capable of motivating participation (relation e7) may generate $\partial H(.)/\partial L > 0$, making the sign of $\partial H(.)/\partial L$ indeterminate.

As discussed earlier, on the quantity side the effect of local capture can be negative or positive depending on the type of local leader (relation e2). If local leader is of a favorable type (Type-A), e.g., motivated to foster regional welfare by augmenting the size of local budget available for development, the first term of equation 9 can be positive. Otherwise, it will be negative (Type-B). Thus, from quantity and quality perspectives the net effect of rising local capture on welfare depends on the quality of local leader (Q).

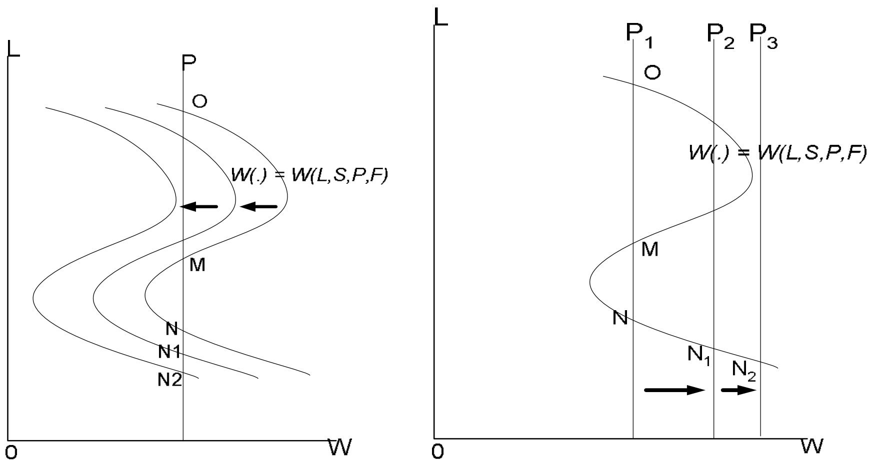
While the signs of $\partial H(.)/\partial L$ and $\partial F(L)/\partial L$ are uncertain, however, somewhere in between there exists a critical value of *H(.)* and *F(.)* such that the effect of rising local capture leads to

$$\partial W(.)/\partial L > 0.$$

When this occurs, the system produces a backward-bending curve that generates multiple equilibria as shown in Figure 2.[9]

Figure 2

Local Capture and Welfare: Alternative Strategies



The goal is either to raise *W* given local capture *L*, or, minimize local capture *L* given *W*. The latter is equivalent to finding lowest *L* along the vertical line *P*. Reducing income

9 Note that participation (P) is independent of capture. Many studies have revealed that participation is influenced by factors such as socio-culturally prescribed family (household heads, spouses, age range), and gender roles (married woman with children); see Beard (2005). While defining participation is not easy, at least the following elements should be entailed: representation, empowerment, benefits for all, and poverty reduction (Blair 2000).

inequality and poverty will facilitate such a goal since shifting the bending curve leftward will guarantee a new equilibrium with lower intensity of local capture (e.g., *N1* and *N2* in the left panel of Figure 2).[10]

An alternative strategy to insure a low capture is to raise participation; this is depicted by a rightward shift of the vertical line *P* in the right panel of Figure 2. This turns out superior as it produces not only lower capture supported by greater accountability, but also higher welfare.

To complement the above analysis of quality component, a corresponding typology of local leaders is constructed, in which alternative behaviors of leaders with respect to quantity component (local budget) is taken into account (see Table 1). Although the main sources of regional revenues are central grant and local taxes, local leaders may find additional sources from local capture, L [0,1], where L=0 means no local capture and L=1 indicates that the region is fully captured ("owned" by elites). The presence of local capture can result in an increase, a decrease, or an unchanged local budget, depending on how local leaders use the additional resources for. Type-A leader will take advantage of the capture to increase the amount of local resources beyond the official budget. A "Complete progress" is achieved when participation and/or initial welfare condition is high, and at the same time local leader is of Type-A. If $\partial H(.)/\partial L < 0$, even with a type-A leader the expected outcome is "Incomplete progress," confirming the key role of participation and initial conditions. Type-B leader leaves the local budget unchanged for any degree of local capture, while Type-C leader has a strong tendency to appropriate local funds either for private benefits or to exercise "return-the-favor" behavior. Either way, it reduces the amount of local resources available for development purposes.

Table 1

Typology of Local Leader and Decentralization Outcome

	$\partial F(L)/\partial L > 0$ **Type-A Leader**	$\partial F(L)/\partial L = 0$ **Type-B Leader**	$\partial F(L)/\partial L < 0$ **Type-C Leader**
$\partial H(.)/\partial L > 0$ **High participation and/or low inequality/poverty**	Complete progress	Propitious	Stagnant
$\partial H(.)/\partial L < 0$ **Low participation and/or high inequality/poverty**	Incomplete progress	Deviating	Deteriorating

10 Recall that lower inequality and poverty tend to raise political awareness that can reduce the intensity of local capture.

To reflect the dynamic implications of the above theoretical exposition, specific functions with assigned numerical values are needed. As shown in the next Sub-section, to understand better the interactions among institutional factors and how the interplay can influence welfare, a series of sensitivity analysis are conducted using the example of a specific welfare function.

Welfare Function: An Example

Reiterating equation 8, the welfare function is decomposed into quantity and quality components. $F_t(L)$ denotes the quantity component under different local leader types (t=A,B,C), where $F_A'(L)>0$, $F_B'(L)=0$, and $F_C'(L)<0$. Ideally, the entire amount of funds acquired through local capture is added to the local budget, i.e., local budget equals to central grant plus local taxes plus the funds acquired from local capture. The value of F(L) is indexed as the fraction of how the value of real amount of local budget deviates from its ideal amount. Suppose different types of local leader takes the following functions:

Type-A: $F_A(L) = c_1 + c_2 L^{c_3}$ where $0<c_1, c_2, c_3 < 1$ and $0 < c_1+c_2<1$ (10)

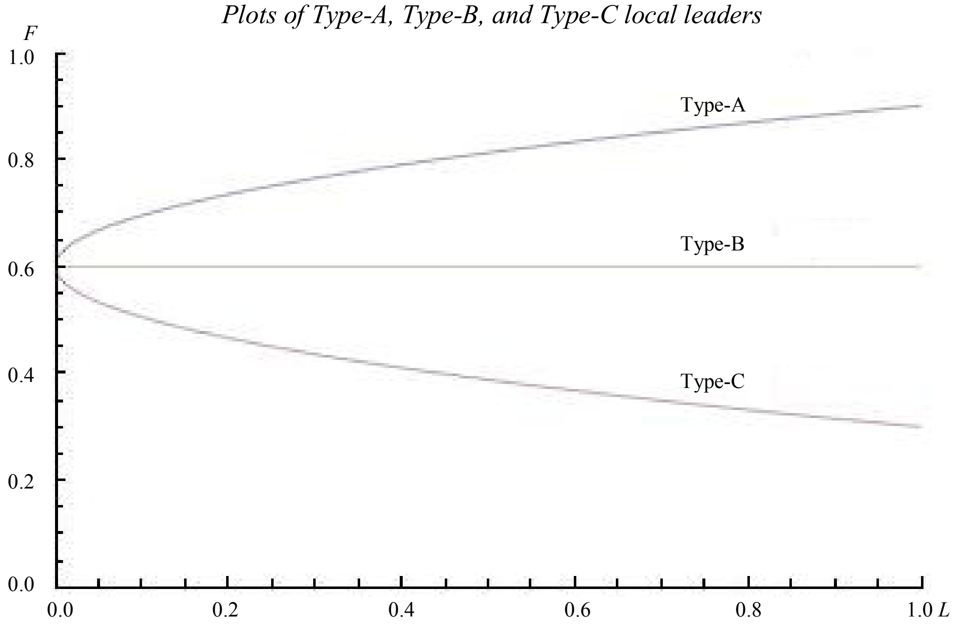
Type-B: $F_B(L) = c_1$ where $0<c_1 < 1$ (11)

Type-C: $F_C(L) = c_1 - c_2 L^{c_3}$ where $0<c_1, c_2, c_3 < 1$ and $0 < c_1-c_2<1$ (12)

Consider Type A: $F(L)=0.6+0.3\sqrt{L}$; Type B: $F(L)=0.6$; Type C: $F(L)=0.6-0.3\sqrt{L}$. Plots of Type-A, Type-B, and Type-C local leaders are illustrated in Figure 3.

Figure 3

Plots of Type-A, Type-B, and Type-C local leaders



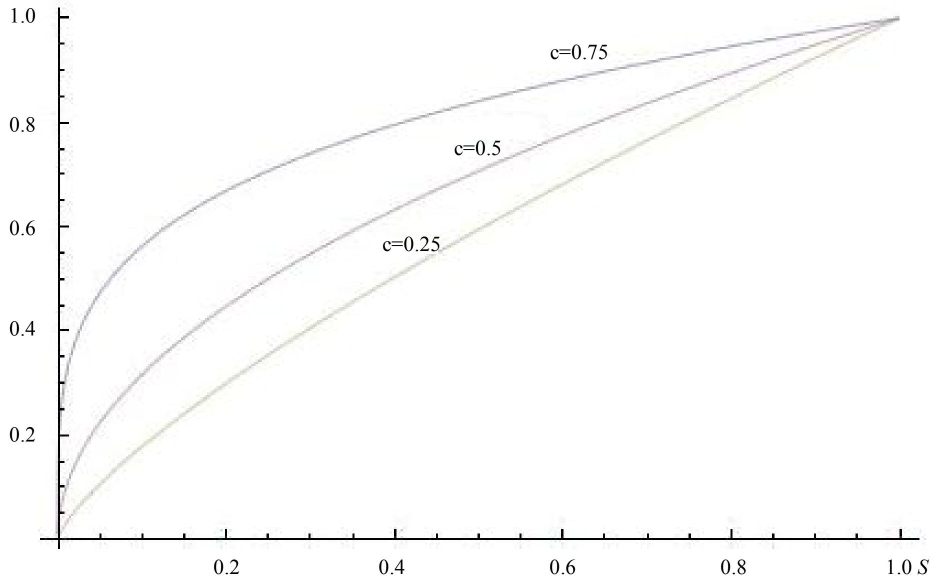
On the quality component, participation is one of the critical factors. Since lower initial welfare (labeled S in Figure 1) implies a smaller fraction of informed voters or lower political awareness, and hence lower participation (P) is an increasing function of initial welfare (S). Moreover, an increase of initial welfare at a lower level tends to raise participation more significantly than at a higher level (concavity). Given population, the degree of people's participation is P $\in$ [0,1]. Furthermore, the initial welfare is indexed such that S $\in$ [0,1] where S=0 implies zero welfare and S=1 indicates perfect welfare. An example of participation function is as follows

$$P(S)=S^c \qquad (13)$$

where 0<c<1 indicates the speed of increase in participation as S increases. The participation curves for some values of c are shown in Figure 4.

Figure 4

Participation and Initial Welfare



Given (13), the quality component of welfare H(P,L,S) can be reduced to H(P,L). Indexing quality factors, H $\in$ [0,1] it is specified that H=0 means no participation for any values of L, and H=1 indicates full participation. Consider the quality function of the following form

$$H(P,L) = P^{b_1} a \frac{L(L-b_2)^2}{b_3} \qquad (14)$$

Parameter a (0<a<1) denotes the region's business climate, where larger value reflects a more conducive business environment, and parameter b1 denotes an index of income inequality (e.g. GINI index). Parameter b2 measures the sensitivity of welfare on income inequality (b1), while b3 measures the efficiency and effectiveness of local government's management. The latter is associated with administrative decentralization, which is one of the three dimensions of decentralization (Filippetti and Sacchi, 2013).

Assigning the following values to the above equation, a=0.1, b1=0.5, b2=0.7, and b3=0.25, and participation is P=0.3, if equations 10, 11 and 12 take the following forms

Type-A: $F_A(L) = 0.6 + 0.3\sqrt{L}$ (10a)

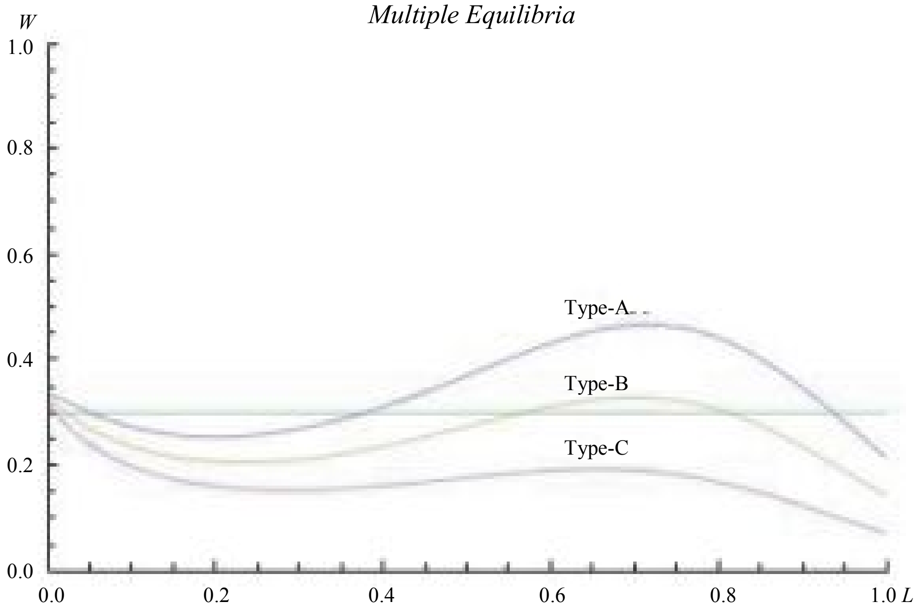
Type-B: $F_B(L) = 0.6$ (11a)

Type-C: $F_C(L) = 0.6 - 0.3\sqrt{L}$ (12a)

the plot of welfare function along L □ [0,1] can be derived as shown in Figure 5. The shape of the curve is very similar to that in Figure 2, implying multiple equilibria, but different shapes obviously can be generated using different forms of welfare function and different values of parameters.

Figure 5

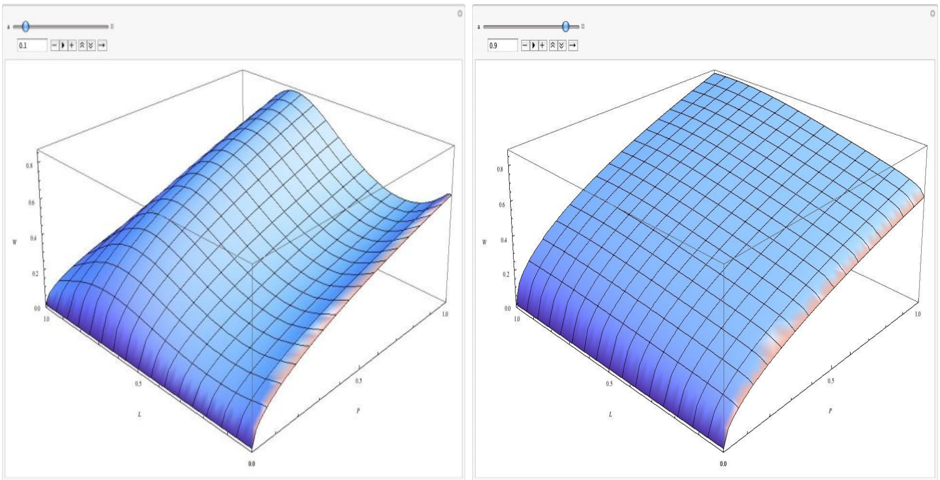
Multiple Equilibria



To the extent that the post-decentralization performance depends on the region's ability to attract business activity, even with high local capture if the size of parameter a is large and the local leader is of Type-A, welfare tends to improve. Only at an extremely high level of capture the welfare starts to decline, albeit slightly. Such a pattern is independent of the level of participation (P), and to some extent also independent of the business climate. Of course, higher participation is likely to generate greater welfare (left panel of Figure 6).

Figure 6

Sensitivity Analysis: Business Climate



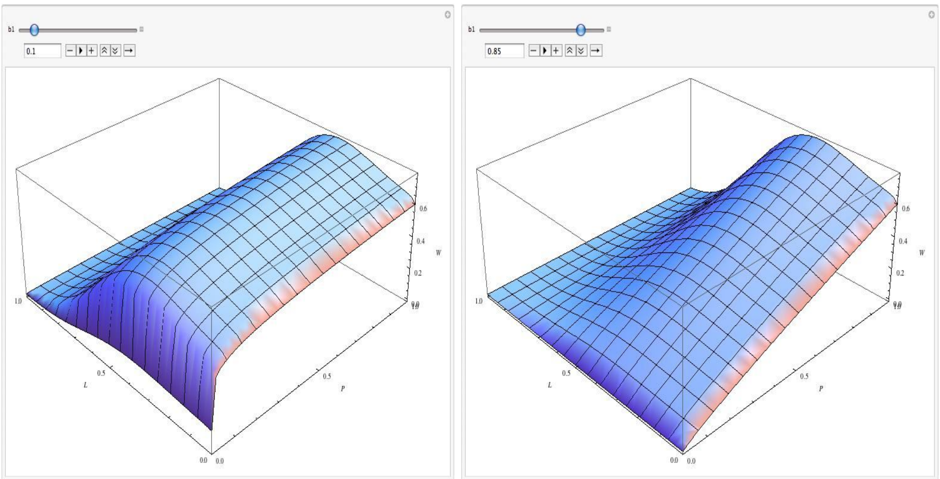
If the region fails to create a favorable business climate, on the other hand, at a low range of capture welfare tends to decline as the degree of capture increases. The relation turns complimentary (higher capture produces higher welfare) when local capture is relatively high but not extremely high. This is due to the fact that the “quantity” factors in equation 8 become more dominant, such that for Type-A leader the funds and other resources acquired from the capture tend to augment local budget. This pattern is not affected by the level of participation (P). The right panel of Figure 6 shows that along the P-axis higher participation produces higher welfare.

The ability of local leaders to influence people’s participation depends on the prevailing social and political structure. In the model, the income inequality (GINI index, where $0 < b1 < 1$) is used as a proxy for such a structure. To evaluate the role of this index in the capture-welfare nexus, one has to consider the sensitivity of the highest achievable welfare to the degree of local capture. There are two possible scenarios: highest welfare can be achieved with lower capture, and highest welfare can be achieved only if local capture is high. Obviously, the former is more desirable, at least in a moral sense. Using b2 to reflect this condition, low b2 is more ideal than high b2 ($0 < b2 <1$).

In an ideal situation (low b2), when income inequality is low, a higher welfare can still be achieved even with growing local capture. But this is only up to a certain point ($L = 0.5$). Beyond that point, local leaders cannot afford to maintain the level of welfare, presumably because they are compelled to return the favor of local elites who supported their candidacy. Under this scenario, a small increase of participation can significantly improve local welfare. As participation continues to grow, however, improvements taper off (concavity). While a similar pattern is also observed in a high inequality environment (high b1), welfare improvements are more sensitive to the interplay of capture and participation. Only with a higher participation the same level of capture can result in a higher welfare (compare the two panels in Figures 7).

Figure 7

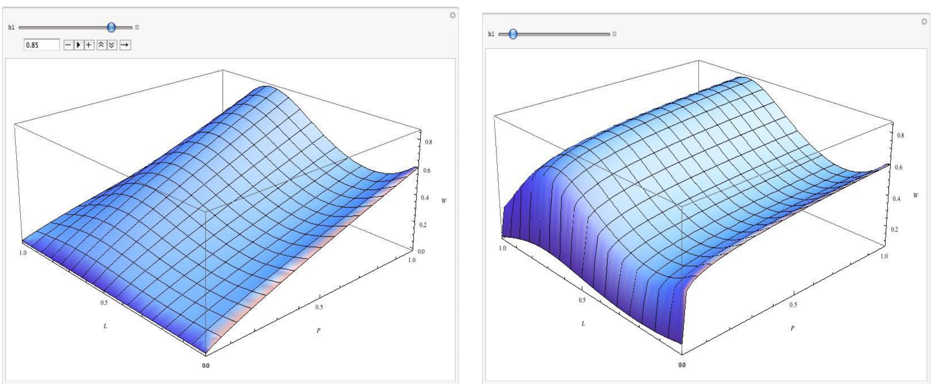
Sensitivity Analysis: Welfare is Less Sensitive to Social Structure (Income Inequality)



Consider the case where the social structure plays a very significant role in determining welfare (high b2). When income inequality is low (small b1), welfare may first increase with growing capture until it reaches the peak. As local capture becomes more widespread, welfare starts to fall. A highest welfare is thus achieved when local capture is limited. This applies irrespective of the level of participation (right panel of Figure 8). On the other hand, if income inequality is high, welfare tends to be lower given the whole range of participation. At a certain level of capture, welfare improvements will take place only if the level of participation increases. Like in the previous low b2 scenario, here the interplay of capture and participation clearly has an important role in determining the level of welfare; more so than in the case of low inequality (compare the two panels in Figure 8).

Figure 8

Sensitivity Analysis: Welfare is Sensitive to Social Structure (Income Inequality)

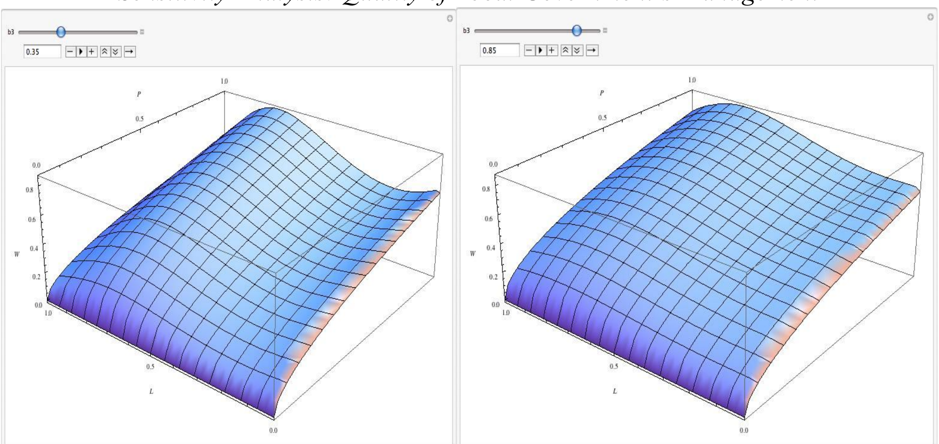


The impact of administrative decentralization is best explained by varying the size of b3 to reflect different management quality of local government. Given a level of participation, when local capture increases welfare can move in either direction depending on the extent to which local leaders are capable of managing their tasks efficiently (getting things done) and effectively (getting things done to worthwhile effect).[11] The closer b3 to unity ($0 < b3 < 1$) the better is the quality of local management.

Depicted in the left panel of Figure 9, when b3 is relatively low (set at 0.35), and local capture is also relatively small, within a certain range welfare and capture move in the opposite direction: higher capture causes welfare to decrease, albeit slightly. As capture reaches a very high level, the relation between the two becomes complimentary. While a similar pattern is observed if the quality of local government's management is good, the welfare outcome is generally better (right panel of Figure 9).

Figure 9

Sensitivity Analysis: Quality of Local Government's Management



A Case Study

Results of applying the above IMD are discussed in this Section. The approach taken was by conducting field survey to measure the intensity of the complex interrelations among institutional factors depicted in Figure 1. The ultimate purpose of the survey is to generate a consistent ranking of the three factors (participation, P; local budget, F; and initial conditions, Q) in IMD by treating the complex interplay of different factors. Given the complexity and intangible factors involved, the *Analytic Hierarchy Process* (AHP) and the *Analytic Network Process* (ANP) were used, the detailed of which is discussed in the Appendix. The field survey was conducted in seven regions throughout Indonesia over the period of 2008-2009.[12] One way to structure the relations among institutional

11 Busy and seemingly efficient operations do not always produce the best outcome. Less-busy ones may indicate that the local government is in control and know what to do to provide the best services to the people.

12 Although at the beginning popular demand for decentralization was not strong enough to push for an immediate change, for purely political reasons the Indonesian government proceeded with a big-bang decentralization in 1999 (became

factors and the welfare effect of decentralization is by placing all the relevant factors in a hierarchy, where the highest level (objectives) determines welfare criteria, and under each criterion the importance of relevant institutional factors is determined.[13] When feedback influences are considered, where factors in each level in the hierarchy can influence and be influenced by factors in other levels, a network is formed. Both the hierarchy and the network used in the survey are shown in the Appendix.

Table 2

Results of Field Survey: For Group in Hierarchy Model

Super Decision				Expert choice	
PALU					
Name	Ideals	Normals	Raw		Overall Inconsistency = .25
1People's Particip	1	0.364146	0.604568	1People's Particip	.364
2Initial condition	0.966851	0.352075	0.584528	2Initial condition	.352
3Available Budget	0.779296	0.283778	0.471138	3Available Budget	.284
JAMBI					
Name	Ideals	Normals	Raw		Overall Inconsistency = .07
1People's Particip	0.909111	0.401597	0.642985	1People's Particip	.402
2Initial condition	0.354626	0.156655	0.250816	2Initial condition	.157
3Available Budget	1	0.441747	0.707268	3Available Budget	.442
MATARAM					
Name	Ideals	Normals	Raw		Overall Inconsistency = .04
1People's Particip	1	0.601918	1	1People's Particip	.602
2Initial condition	0.144856	0.087192	0.144856	2Initial condition	.087
3Available Budget	0.516498	0.31089	0.516498	3Available Budget	.311
MALANG					
Name	Ideals	Normals	Raw		Overall Inconsistency = .03
1People's Particip	1	0.52247	0.999987	1People's Particip	.522
2Initial condition	0.523064	0.273285	0.523057	2Initial condition	.273
3Available Budget	0.390922	0.204245	0.390917	3Available Budget	.204
BANJARMASIN					
Name	Ideals	Normals	Raw		Overall Inconsistency = .02
1People's Particip	0.692824	0.289886	0.557808	1People's Particip	.319
2Initial condition	0.697165	0.291702	0.561303	2Initial condition	.276
3Available Budget	1	0.418412	0.805122	3Available Budget	.405
BANDUNG					
Name	Ideals	Normals	Raw		Overall Inconsistency = .11
1People's Particip	1	0.527833	1	1People's Particip	.528
2Initial condition	0.264558	0.139642	0.264558	2Initial condition	.140
3Available Budget	0.629981	0.332525	0.629981	3Available Budget	.333
SEMARANG					
Name	Ideals	Normals	Raw		Overall Inconsistency = .07
1People's Particip	1	0.426469	0.713323	1People's Particip	.433
2Initial condition	0.455387	0.194208	0.324838	2Initial condition	.253
3Available Budget	0.889448	0.379322	0.634464	3Available Budget	.314

As displayed in Table 2, with the exception of two regions, *Jambi* and *Banjarmasin*, results from the group survey indicate that people's participation (P) is generally ranked highest among factors that determine the welfare effect of local capture. Notice that for *Palu*, the least developed among seven regions, the second most important factor after participation is not the availability of local budget (F), but instead the initial welfare condition (S). This is consistent with the premise that the persistence of poor region is

operational in 2001). The mixed results in terms of post-decentralization performance make Indonesia suitable for model validation. The country managed to avoid a chaotic situation despite the abrupt change, and some regions have enjoyed the benefits of the policy. Yet, the resulting outcome in other regions has been disappointing, not according to what the theory suggests.

13 The AHP uses relative measurements (ratio scales) derived from paired comparisons. Ratio scales are a fundamental kind of number amenable to performing the basic arithmetic operations of addition and subtraction within the same scale, multiplication and division of different scales, and combining the two operations by meaningfully weighting and adding different scales to obtain a unidimensional scale. Hence they are very useful not only for capturing perceptions towards welfare criteria and institutional factors, but also for synthesizing the priority results that requires some arithmetic operations.

caused by the low initial condition (path dependence). Similarly, for the relatively well-to-do region, *Malang*, the second most important factor is also the initial welfare condition. While *Malang* is already developed (higher S), all factors including participation tend to result in a "positive local capture," reinforcing the region's welfare condition. Perceptions of respondents in that region corroborate the hypothesis.

Table 3

Results of Field Survey: For Group, Individuals, and Combined in Hierarchy Model

Hierarchy						
Group Survey						
Summarized result of the group survey in 7 regions based on Hierarchy model (Using Super Decision and Expert Choice software)						
Super Decision					**Expert choice**	
Name	Rank	Ideals	Normals	Raw	Rank	
1People's Participation	1	0.93609511	0.436124866	0.76595211	1	0.443099225
2Initial condition	3	0.41671289	0.194145702	0.340972015	3	0.20007053
3Available Budget	2	0.70685427	0.329322192	0.578377832	2	0.319149594
Individual Survey						
Summarized result of the individual survey in 4 regions based on Hierarchy model (Using Super Decision software)						
Name	Rank	Ideals	Normals	Raw		
1People's Participation	1	0.73421686	0.377562735	0.691382905		
2Initial condition	3	0.40464311	0.208083135	0.381036434		
3Available Budget	2	0.61913245	0.318381897	0.58301257		
Combination of Group and Individual suverys						
Summerized result of the group and individual survery based on Hierarchy model (Using Super Decision)						
Name	Rank	Ideals	Normals	Raw		
1People's Participation	1	0.82903366	0.405788734	0.727712989		
2Initial condition	3	0.41063366	0.200993648	0.360448		
3Available Budget	2	0.66154094	0.323805844	0.580690577		

All seven regions combined, the ranking shows that participation is indeed the most critical factor (the weight being .436 using *Super Decision*, and .443 using *Expert Choice*), followed by the size of local budget (.329 and .319, respectively); see Table 3.[14] The ranking remains the same for individuals category, the geometric means of which are shown in the second (middle) part of Table 3. When the results of group and individuals survey are combined, the weights for participation and local budget are .406 and .324, respectively.

14 Super Decision is an experimental software capable of calculating the super-matrix operations involved in the ANP-type of network. By removing the feedback components, it can produce a priority ranking similar to that based on AHP-type of hierarchy. Expert Choice, on the other hand, is specifically designed for AHP.

Table 4

Sensitivity Analysis: Removing One Region at a Time for Group and Individuals in Hierarchy Model

Hierarchy								
Sensitivity Analysis(Individual and Group) by excluding region by region								
1.Palu								
	Individual				Group			
Name					Rank	Ideals	Normals	Raw
1People's Participati					1	0.9258	0.4494	0.7968
2Initial condition					3	0.3622	0.1758	0.3117
3Available Budget					2	0.6955	0.3376	0.5985
2.Jambi								
	Individual				Group			
Name	Rank	Ideals	Normals	Raw	Rank	Ideals	Normals	Raw
1People's Participati	1	0.813	0.432	0.784	1	0.9407	0.4422	0.7886
2Initial condition	3	0.35	0.186	0.337	3	0.4281	0.2012	0.3589
3Available Budget	2	0.549	0.292	0.529	2	0.6671	0.3136	0.5593
3.Mataram								
	Individual				Group			
Name	Rank	Ideals	Normals	Raw	Rank	Ideals	Normals	Raw
1People's Participati	1	0.686	0.354	0.642	1	0.9258	0.4133	0.7327
2Initial condition	3	0.407	0.21	0.381	3	0.497	0.2219	0.3933
3Available Budget	2	0.637	0.328	0.595	2	0.7448	0.3325	0.5894
4.Malang								
	Individual				Group			
Name					Rank	Ideals	Normals	Raw
1People's Participati					1	0.9258	0.4232	0.7327
2Initial condition					3	0.4012	0.1834	0.3175
3Available Budget					2	0.7802	0.3566	0.6174
5.Banjarmasin								
	Individual				Group			
Name	Rank	Ideals	Normals	Raw	Rank	Ideals	Normals	Raw
1People's Participati	1	0.697	0.364	0.659	1	0.9842	0.4668	0.8075
2Initial condition	3	0.43	0.224	0.406	3	0.3825	0.1814	0.3138
3Available Budget	2	0.593	0.309	0.56	2	0.6671	0.3164	0.5474
6.Bandung								
	Individual				Group			
Name	Rank	Ideals	Normals	Raw	Rank	Ideals	Normals	Raw
1People's Participati	1	0.748	0.366	0.69	1	0.9258	0.4225	0.7327
2Initial condition	3	0.438	0.214	0.404	3	0.4495	0.2051	0.3557
3Available Budget	2	0.709	0.347	0.655	2	0.7205	0.3288	0.5702
7.Semarang								
	Individual				Group			
Name					Rank	Ideals	Normals	Raw
1People's Participati					1	0.9258	0.4378	0.7751
2Initial condition					3	0.4106	0.1941	0.3437
3Available Budget					2	0.6803	0.3217	0.5695

The robustness of the results is tested by conducting two types of sensitivity analysis: dynamic analysis, and removing one region at a time. The first type is done for each set of questionnaire, the results of which indicate that the most sensitive factor for poverty is participation, for inequality is the initial welfare condition, and for growth and HDI is the size of local budget. Thus, if local development needs to focus more on poverty alleviation, efforts have to be made to raise people's participation. Field observations corroborate such finding: the welfare effects of decentralization with local capture in regions where people are more politically aware and actively participate in various local development programs tend to be more positive. The second sensitivity analysis reveals that in all cases the superiority of participation (P) continues to hold (Table 4). The survey results are therefore fairly robust.

Table 5

Results of Field Survey: For Individuals in Network Model

Network(Individual's geometric mean)

JAMBI				
ANP				
Name	Rank	Ideals	Normals	Raw
1People's Participation	3	0.470977795	0.224316683	0.428350904
2Initial condition	2	0.638308832	0.304012909	0.580537309
3Available Budget	1	0.896066878	0.426777944	0.814966396

MATARAM				
ANP				
Name	Rank	Ideals	Normals	Raw
1People's Participation	1	0.796528566	0.352012708	0.777336689
2Initial condition	3	0.671627639	0.300617182	0.657223917
3Available Budget	2	0.716403781	0.32065884	0.701039823

BANDUNG				
ANP				
Name	Rank	Ideals	Normals	Raw
1People's Participation	1	0.667566395	0.337707543	0.667566395
2Initial condition	3	0.423094747	0.214034549	0.423094747
3Available Budget	2	0.666475413	0.337156056	0.666475413

Individual survey				
Summarized result of the survey in 3 regions based on network model (Using Super Decision software)				
Name	Rank	Ideals	Normals	Raw
1People's Participation	2	0.630326285	0.29875847	0.605760718
2Initial condition	3	0.566064191	0.269440723	0.544494766
3Available Budget	1	0.753518886	0.35866752	0.724806658

Unlike the case of hierarchy model, results from the network model with feedback effects are mixed. For example, in *Mataram* and *Bandung* the geometric means point to participation being the most important one, but for *Jambi* the size of local budget is the most important factor determining the welfare effect of decentralization (Table 5). Yet, when the three cases are combined, the geometric means put the size of local budget in the highest rank, followed by participation.

It is important to note, however, that during the field survey the network model was applied only to individuals, not groups, and only limited number of region was covered. Nonetheless, the fact that people's participation (P) and the size of budget (F) stand out as important factors in explaining the welfare effect of decentralization confirms the IMD's conjecture that quality and quantity factors jointly play an important role in the analysis of decentralization.

Conclusions

The focus of this paper is on how the interrelations among different dimensions of decentralization – fiscal, administrative and institutional – affect regional growth and welfare. By using the *Institutional Model of Decentralization* (IMD), it is first shown that when private benefits to local leaders exceeds social benefits, regional growth after decentralization may falter. That is, post-decentralization efficiency often predicted by the theory stays as unrealized hypothesis. With widespread local capture occurring in many countries following political decentralization, a next question of interest is: what mechanism can explain the effect of local capture on decentralization outcome?

By expanding the objective from growth to a more broadly-defined welfare, it is shown that the interactions among initial condition, effective size of budget, people's participation, and quality of local leaders affect the way local capture influences regional welfare. The interactions are complex, non-linear, and may generate multiple equilibria. As the quality of local leaders including their capacity to motivate participation determines the outcome, a typology of leaders is constructed. Different types of self-reinforcement factors dictate leader's behavior that can reinforce different outcome, establishing the evolution between welfare and institution (implying endogenous institutions). Based on the sensitivity analysis, it is also conjectured that the nature and intensity of each factor's role in different countries vary depending on the country's social, political and economic structure.

Applied to the case of seven regions in Indonesia, the theoretical prediction of IMD is substantiated. On the group category, given local capture people's participation is generally ranked the highest among factors that govern the welfare effect of decentralization. It plays the most critical role in reducing the capture while simultaneously maximizing the welfare The persistence of initial condition is also verified by the fact that it is ranked high, only in two extreme cases it is ranked the second, i.e., in the least developed (*Palu*) and very well-to-do region (*Malang*). In the latter, all factors including participation result in a "positive local capture," reinforcing the region's welfare condition. The survey also reveals that regions identified as "deteriorating" and having low participation tend to be poor. And they are persistently so. Not only that this corroborates the critical role of initial condition, but it also suggests that in a democratic system like Indonesia, decentralization is welfare-enhancing only for regions under "complete" progress, not for all regions.

Given that a multiple equilibria scenario after decentralization is common, not unique for Indonesia, an institutional reform is warranted. But the reform ought to take account the potential interactions among institutional factors and characteristics as described in IMD, not just aiming at minimizing transaction costs.

Appendix: Brief Explanations of AHP and ANP

Let A_1, A_2, A_3, ..., A_n be n elements in a matrix within a hierarchy. The pairwise comparisons on pairs of elements (A_i, A_j) are represented by an *n-by-n* matrix **A** = (a_{ij}), where i,j = 1, 2, 3,....., n. Define a set of numerical weights w_1, w_2, w_3, ..., w_n that reflects the recorded comparisons,

$$
\begin{array}{c c}
 & \begin{array}{ccccc} A_1 & A_2 & & & A_n \end{array} \\
\begin{array}{c} A_1 \\ \\ A= \\ \\ A_n \end{array} &
\begin{bmatrix}
w_1/w_1 & w_1/w_2 & \cdots\cdots & & w_1/w_n \\
 & & & & \\
.. & .. & .. & .. & .. \\
 & & & & \\
w_n/w_1 & w_n/w_2 & \cdots\cdots & & w_n/w
\end{bmatrix}
\end{array}
$$

The scales used in the pairwise comparisons in AHP are based on Saaty's scaling system (Saaty, 1996), i.e., from 1 to 9. Since every row is a constant multiple of the first row, A has a unit rank. By multiplying A with the vector of weights w,

$$Aw = nw$$

To recover the scale from the matrix ratios, the following system ought to be solved:

$$(A-nI)w = 0$$

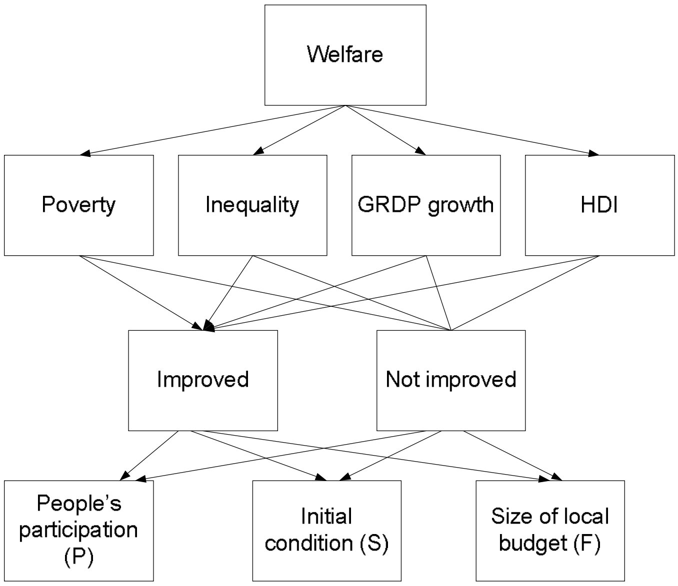
Clearly, a nontrivial solution can be obtained if and only if det(A-nI) vanishes, i.e., the *characteristic equation* of A. Hence, n is an *eigenvalue* and w is an *eigenvector,* of A**.** Given that A has a unit rank, all its eigenvalues except one are zero. Thus, the *trace* of A is equal to ***n*.** If each entry in A is denoted by a_{ij}, then $a_{ij} = 1/a_{ji}$ (reciprocal property) holds, and so does $a_{jk} = a_{ik} / a_{ij}$ (consistency property). By definition, $a_{ii} = a_{jj} = 1$, that is, when comparing two same elements. Therefore, if we are to rank n number of elements, i.e., A is of the size n-by-n, the required number of inputs from the paired comparisons is less than n^2; it is equal to only the number of entries of the sub-diagona part of A. Hence, if there are three elements in a particular level of a hierarchy, only three pairwise comparisons are required. In general, however, the precise value of w_i/w_j is hardly known simply because the pairwise comparisons are only an estimate, suggesting that there are some perturbations. While the reciprocal property still holds, the consistency property does not. By taking the largest eigenvalue denoted by $\overline{\lambda}_{max}$,

$$A^p \cdot w^p = \overline{\lambda}_{max} \cdot w^p$$

where A^p is the actual, or the given, matrix (perturbed from matrix A). Solving the above gives a consistent matrix whose entries are w_i/w_j; it is a consistent estimate of A, although A^p itself does not need to be consistent. A^p will be consistent if and only if $\overline{\lambda}_{max} = n$. As long as the precise value of w_j/w_i cannot be given, which is common in real situation due to the bias in the comparisons, $\overline{\lambda}_{max}$ is always greater than or equal to n; hence, a measure of consistency can be derived based on the deviation of $\overline{\lambda}_{max}$ from n. When more than two elements are compared, the notion of consistency can be associated with *transitivity* condition: if $A_1 \succ A_2$ and $A_2 \succ A_3$, then $A_1 \succ A_3$. It should be clear that in solving for w, the *transitivity* assumption is not strictly required; the inputted comparisons do not have to reflect a full consistency. Yet, as shown above, the resulting matrix and the corresponding vector remain consistent. It is this consistent vector w that reflects the priority ranking of the elements in each level. Hence, in a standard hierarchy with three levels (goals, criteria, and alternatives), the elements in each level are pairwise compared with respect to elements in the level above it, and the resulting vector at the bottom level reflects the priority ranking of the alternatives.

Figure A1

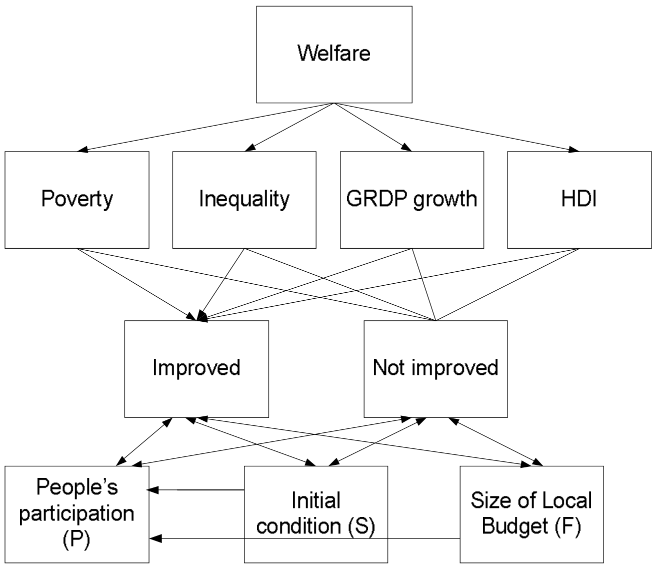
Role of Institutional Factors in Decentralization: Hierarchy Model



To test the model framework for decentralization using AHP, the hierarchy shown in Figure A1 is used. Maximizing local welfare is the goal, and four indicators are identified: poverty, inequality, local output growth (GRDP growth), and human development index (HDI). These indicators are to be prioritized by using pairwise comparison matrix from which the eigen-vector reflecting the consistent ranking is derived. After clarifying whether in the respective region each of the indicators has improved or not, the next step is to rank the importance of the determinants of those welfare indicators. As discussed before, of all possible determinants three stand out: people's participation (P), initial condition (S), and size of local budget (F). By taking into account the consistent ranking of welfare indicators and institutional factors, the overall results can be synthesized. In some regions, a modified structure of the ANP-type of network is used in which feedback effects capturing the performance of each factor (improved or not improved) are identified. Unlike AHP, a network model recognizes two-way dependence relationships that exist among variables). With feedback, the alternatives can depend on the criteria as in a hierarchy but they may also depend on each other. The criteria themselves can depend on the alternatives and on each other. Hence, it involves a network rather than a hierarchy. With such a feature, the results are more stable because one considers the influence on and survival in the face of other influences.

Figure A2

Role of Institutional Factors in Decentralization: Network Model



The network framework used in the survey is shown in Figure A2. Notice that the 2-way arrows indicate the feedback effects between the bottom level and the level above it, and between some elements in the same level (size of local budget and initial welfare condition influence participation).

While in a hierarchy-based model a set of pairwise comparison matrices are used, the presence of feedback influences in a network model requires a *supermatrix* that contains a set of sub—matrices. This supermatrix should capture the influence of elements in a network on other elements in that network. Denoting a cluster by C_h, $h = 1, \ldots\ldots m,$ and assuming that it has n_h elements $e_{h1}, e_{h2}, e_{h3} \ldots\ldots, e_{hmh}$, Figure A3 shows the supermatrix of such a hierarchy:

Figure A3

Supermatrix of a Hierarchy

$$W = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & 0 \\ W_{21} & 0 & \cdots & 0 & 0 & 0 \\ 0 & W_{32} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & W_{n-1,\, n-2} & 0 & 0 \\ 0 & 0 & \cdots & 0 & W_{n,\, n-1} & I \end{bmatrix}$$

Figure A4

Supermatrix of a Holarchy

$$W = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & W_{1,n} \\ W_{21} & 0 & \cdots & 0 & 0 & 0 \\ 0 & W_{32} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & W_{n-1,\, n-2} & 0 & 0 \\ 0 & 0 & \cdots & 0 & W_{n,\, n-1} & 0 \end{bmatrix}$$

When the bottom level affects the top level of the hierarchy, a form of network known as *holarchy* is formed, the supermatrix of which will look like the one displayed in Figure A4. Notice that the entry in the last row and column of the supermatrix in Figure A3 is the identity matrix I corresponding to a loop at the bottom level of the hierarchy. This is a necessary aspect of a hierarchy viewed in the context of supermatrix. On the other hand, the entry in the first row and last column of a holarchy in Figure A4 is nonzero, indicating that the top level depends on the bottom level. The entries of sub-matrices in W_{ij} are the ratio scales derived from paired comparisons performed on the elements within the clusters themselves according to their influence on each element in another cluster (outer dependence) or elements in their own cluster (inner dependence). The resulting *unweighted supermatrix* is then transformed into a matrix each of whose columns sums to unity to generate a stochastic supermatrix. The derived weights are used to weight the elements of the corresponding column blocks (cluster) of the supermatrix, resulting in a *weighted supermatrix* which is also stochastic. The stochastic nature is required for the reasons described below. The typical entry of Figure A5 supermatrix is shown in Figure A6.

Figure A5

Supermatrix of a Network

$$
W = \begin{array}{c c} & \begin{array}{c c c c} C_1 & C_2 & \cdots & C_N \\ e_{11} e_{12} \cdots e_{1n_1} & e_{21} e_{22} \cdots e_{2n_2} & & e_{N1} e_{N2} \cdots e_{Nn_N} \end{array} \\ \begin{array}{c c} C_1 & \begin{array}{c} e_{11} \\ e_{12} \\ \vdots \\ e_{1n_1} \end{array} \\ C_2 & \begin{array}{c} e_{21} \\ e_{22} \\ \vdots \\ e_{2n_2} \end{array} \\ \vdots & \\ C_N & \begin{array}{c} e_{N1} \\ e_{N2} \\ \vdots \\ e_{Nn_N} \end{array} \end{array} & \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & \cdots & W_{2N} \\ \vdots & \vdots & \cdots & \vdots \\ W_{N1} & W_{N2} & \cdots & W_{NN} \end{bmatrix} \end{array}
$$

Figure A6

Entry in the Supermatrix of a Network

$$
W_{ij} = \begin{bmatrix} W_{i1}^{(j_1)} & W_{i1}^{(j_2)} & \cdots & W_{i1}^{(jn_j)} \\ W_{i2}^{(j_1)} & W_{i2}^{(j_2)} & \cdots & W_{i2}^{(jn_j)} \\ \vdots & \vdots & & \vdots \\ W_{in_i}^{(j_1)} & W_{in_i}^{(j_2)} & \cdots & W_{in_i}^{(jn_j)} \end{bmatrix}
$$

Since an element can influence the second element directly and indirectly through its influence on some third element and then by the influence of the latter on the second, every such possibility of a third element must be considered. This is captured by squaring the weighted matrix. But the third element also influences the fourth, which in turn influences the second. These influences can be obtained from the cubic power of the weighted supermatrix. As the process is performed continuously, one will have an infinite sequence of influence matrices denoted by W^k, $k = 1,2 \ldots\ldots\ldots$ The question is, if one takes the limit of the average of a sequence of N of these powers of the supermatrix, will the result converge, and, is the limit unique? It has been shown that such a limit exists given the stochastic nature of the weighted supermatrix (Saaty, 2001). There are 3 cases to consider in deriving W^k: (1) $\lambda_{max} = 1$ is a simple root and there are no other roots of unity in which case given the nonnegative matrix W is *primitive*, we have $\lim_{k\to\infty} W^k = we^T$, implying that it is sufficient to raise the primitive stochastic matrix W to large powers to yield the limit outcome; (2) there are other roots of unity that cause cycling, in which case Cesaro sum is applied; and (3) $\lambda_{max} = 1$ is a multiple root, in which case the Sylvester's formula with $\lambda_{max} = 1$ is applied. For further details, see Saaty (2001) and

Azis (2009). In practice, however, one simply needs to raise the stochastic supermatrix to large powers to read off the final priorities in which all the columns of the matrix are identical and each gives the relative priorities of the elements from which the priorities of the elements in each cluster are normalized to one. The powers of the supermatrix do not converge unless it is stochastic, because then its largest eigenvalue is one. When a convergence is failed to achieve (a cyclic case) the average of the successive matrices of the entire cycle gives the final priorities (Cesaro sum), in which the limit cycles in blocks and the different limits are summed and averaged and again normalized to one for each cluster. At any rate, raising the stochastic supermatrix to large powers gives what is known as *limiting supermatrix*. Hence, there are 3 supermatrices to be used: (1) the original unweighted supermatrix of column eigenvectors obtained from pair wise comparison matrices of elements; (2) the weighted supermatrix in which each block of column eigenvectors belonging to a cluster is weighted by the priority of influence of that cluster, rendering the weighted supermatrix column stochastic; and (3) the limiting supermatrix obtained by raising the weighted supermatrix to large powers. To apply the network model, our survey team used *Super Decision* software, and for the hierarchy model the team used *Super Decision* and *Expert Choice*.

REFERENCES

Azis, Iwan J. (2009): Crisis, Complexity, and Conflict, Emerald, United Kingdom.

_________, (2008): "Institutional Constraints and Multiple Equilibria in Decentralization," Review of Urban and Regional Development Studies, Vol. 20, No. 1, March, p. 22–33.

_____________, (2010): "Endogenous Institution in Decentralization," Peace Economics, Peace Science and Public Policy, Volume 16, Issue 2.

______________., Maria Wiharja (2010): Theory of Endogenous Institutions and Evidence from an In-depth Field Study in Indonesia, Economic & Finance Indonesia, Vol. 58. 3., p. 309–334.

Bardhan, Pranab. – Dilip Mookherjee (2005): "Decentralization, Corruption and Government Accountability: An Overview," Handbook of Economic Corruption edited by Susan Rose-Ackerman, Edward Elgar, 2006.

Beard, Victoria A. (2005): "Individual Determinants of Participation in Community Development in Indonesia," Environment and Planning C:Government and Policy. Vol 23. p 21–39.

Besley, Timothy – Stephen Coate (1999): "The Public Critique of Welfare Economics: An Exploration," NBER Working Papers 7083, National Bureau of Economic Research, Inc.

Binswanger, Hans (1999): "Technical Consultation on the Decentralization of Rural Development," Proceedings of FAO conference, Rome, 16-18 December 1997, Rome: FAO.

Blair, Harry (2000): "Participation and Accountability at the Periphery: Democratic Local Governance in Six Countries," World Development, Vol 28, No 1, pp 21–39.

Bruno, M. – Pleskovic, B. (eds.) (1996): Annual World Bank conference on development economics. Washington, DC: World Bank.

Burki, S. J. – Perry, G. – Dillinger, W. (eds.) (1999): Beyond the center. Decentralizing the state. Washington D.C.: The World Bank.

Easterly, William – Ross, Levine (2002): "Tropics, Germs, and Crops: How Endowments Influence Economic Development," NBER Working Papers 9106, National Bureau of Economic Research, Inc.

Eggertsson, Thrainn (1996): "No Experiments, Monumental Disasters: Why It Took a Thousand Years To Develop a Specialized Fishing Industry in Iceland," Journal of Economic Behavior and Organization 30:1, pp. 1–23.

Filippetti, Andrea – Agnese, Sacchi (2013): "Varieties of decentralization, institutional complementarities, and economic growth: evidence in OECD countries," National Research Council of Italy and Birkbeck University of London (UK), unpublished.

Greif, Avner (2006): Institutions and the Path To The Modern Economy, Cambridge University Press, N.Y.

Henry, Peter Blair. – Conrad Miller (2008): Institutions vs. Policies: A Tale of Two Islands, NBER Working Paper No. 14604, December.

Huther, Jeff. – Shah, Anwar (1998): "Applying a simple measure of good governance to the debate on fiscal decentralization," Policy Research Working Paper, Series 1894, The World Bank.

Londregan, John Benedict. – Keith T. Poole (1996): "Does High Income Promote Democracy? World Politics – Vol. 49, No. 1, October.

Manor, J. (1999): "Explaining the last wave of decentralizations", The Political Economy of Democratic Decentralization. World Bank, Washington DC.

Martinez-Vazquez, J. – McNab, R. M. (2003): Fiscal decentralization and economic growth. World development 31, 1597–1616.

North, Douglass C. (1993): Prize Lecture, The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel. Lecture to the memory of Alfred Nobel, December 9.

_______________, (2005): Understanding the process of economic change. Princeton University Press, Princeton.

Oates, W.E. (1972): Fiscal Federalism New York. Harcourt Brace Jonanovitch.

___________(1994): Federalism and government finance, in Quigley, J.M. and Smolenskey, E. (eds.) Modern Public Finance, 125-151. Cambridge, Ma.: Harvard University Press.

Prud'homme, Remy (1994): " On the Dangers of Decentralization." Policy Research Working Paper, No. 1252, Washington, DC: World Bank.

Rodrik, Dani. – Arvind, Subramanian – Francesco, Trebbi (2002): "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development," NBER Working Paper No. 9305, October.

Sachs, Jeffrey (2003): "Institutions Matter,but Not for Everything," IMF Finance and Development, Vol. 40, No. 2, June 2003.

Saaty, T. L. (1996): Fundamentals of decision making and priority theory with the analytic hierarchy process. Pittsburgh, PA: RWS Publication.

__________ (2001): Decision making with dependence and feedback: The analytic network process. Pittsburgh: RWS Publication.

Shleifer, Andrei. – Vishny, Robert W. (1993): "Corruption," The Quarterly Journal of Economics, MIT Press, vol. 108(3), pages 599–617, August.

_____________. – Daniel, Treisman (2005): A Normal Country: Russia After Communism, Journal of Economic Perspectives, Vol 19, No 1, pp 151–174.

Tanzi, Vito (2002): "Pitfalls of Decentralization." In Managing Fiscal Decentralization. E. Ahmad and V. Tanzi, eds., 17–30: Routledge.

United Nations (1991): Fiscal decentralisation and the mobilisation and use of national resources for development: Issues, experiences and policies in the ESCAP region. New York: United Nations.

World Bank (2000): The World Development Report 1999/2000. Washington, D. C.: The World Bank and Oxford: Oxford University Press.

REGIONAL STATISTICS, 2013, VOL. 3: 30–40 DOI: 10.15196/RS03102

GABRIELL MÁTÉ[a] – ANDRÁS KOVÁCS[b] – ZOLTÁN NÉDA[c]

Hierarchical Settlement Networks

Abstract

A network representation is introduced for visualizing hierarchical region structures on various spatial scales. The method is based on a spring-block model approach borrowed from physics and it was previously used with success to detect regions in any geographical space. We illustrate here our network construction method for the case of Transylvania, USA and Hungary.

Keywords: settlement networks, spring-block model, USA, Transylvania, Hungary.

Introduction

It is commonly believed that the core of any complex system is a network (Barabási 2005). The underlying network structure carries useful information for understanding the evolution and measurable characteristics of large systems. Networks have proven that they are useful for studying biological, social and economic systems (Barabási 2012).

A network approach might be helpful also in visualizing and understanding the complex hierarchical inter-relationship between settlements in a given geographic region. The method for such an approach is straightforward, assuming that one is able to detect hierarchical region-like structures and region centers on different spatial scales. The hierarchical space-division at various scale, could be synthesized in a compact and visually interpretable manner by drawing up such a tree-like network topology. Here, based on our previously elaborated region detection algorithm (Máté–Néda–Benedek 2011), we present a method for the construction of such a hierarchical connection network.

This method can be helpful in the spatial analysis of any kind of territorial data. The complexity of socio-economic processes with spatial dimension implies that natural space delimitation cannot be possible (Dusek 2004), the determination of territorial units (regions) depends on the examined phenomenon and its geographical characteristics (Haggett 2001). Researching the spatial flow of goods, services and information one can border regions with different spatial extent in different hierarchical structures. Thus it has got high importance in bordering to clarify the regional structures on micro, mezzo and macro regional level, based on different socio-economic data.

a Institute for Theoretical Physics, Heidelberg University, Philosophenweg 16 D-69120 Heidelberg, Germany E-mail: G.Mate@tphys.Uni-Heidelberg.DE

b Edutus College, Department of International Business, 1114 Budapest, Villányi út 11–13., Hungary, E-mail: kovacs.andras @edutus.hu

c Department of Physics, Babes-Bolyai University, str. Kogalniceanu 1, R0-400084 Cluj-Napoca, , Romania, E-mail: zneda@phys.ubbcluj.ro

Our method enables us to border regions on different spatial levels by using different kinds of data. This flexibility may contribute to a more efficient regional research in bordering and in the horizontal and vertical division of socio-economic and geographical space (Nemes Nagy 2009).

This paper can be divided into three main parts. First we will introduce the methodological background and its geographical relevance, in the second part we introduce three examples where spatial data are analyzed and visualized (examples are from the USA, Transylvania and Hungary). In the last part we would like to point out the socio-economic relevancies of this method in spatial analysis and regional science.

Method

Preliminaries

Spring-block models have been used widely to model various phenomena related to relaxation, avalanche-like dynamics or self-organized criticality (Járai-Szabó–Néda 2012, Kovács–Néda 2007). The model considers a system of blocks that are interconnected by springs in a lattice like topology. The blocks can slide on a surface and apart of the elastic forces experienced from their neighbors a friction force is damping their free movement on the surface. The friction forces are usually velocity dependent. In the simplest version there is a dynamic friction force that is independent of the sliding block's velocity and a static friction force (acting when the block is in rest relative to the surface), which has a larger value than the dynamic one. Originally, the model was considered in one-dimension by Burridge and Knopoff in order to model the power-law like distribution of the earth-quakes magnitude (Burridge–Knopoff 1967). The model was then extended to two-dimension by Olami, Feder and Christensen (1992). One of the most spectacular application of this model is it's success in modeling fracture and fragmentation patterns resulting from quasi-static drying of granular materials in contact with a frictional substrate (Leung–Néda 2010). The success of reproducing this relaxation dynamics in drying granular materials, motivated the use of the spring-block model for hierarchical region detection. Regions can be imagined as resulting from a fragmentation dynamics in a geographic space. Once the interconnection topology of the settlements is defined and the strength of the interconnecting forces are revealed, a simple relaxation dynamics will lead us to a hierarchical grouping. This hierarchical region-detection method was used in (Máté–Néda–Benedek 2011). Here, for the sake of an easier understanding of the network construction method, we now briefly review the main elements in our spring-block approach for region detection. For a more detailed description we suggest the reader to consult (Máté–Néda–Benedek 2011).

The spring-block model for region detection

Settlements are regarded as *blocks* with masses m_i proportional with their population (W_i). The blocks are displaced on a two-dimensional plane, so that their relative coordinates are determined by the original GPS coordintes of the settlements. On each block "*i*" a friction force is acting on behalf of the surface. For the sake of simplicity we consider here an over-damped motion of the blocks. The value of the force acting on block "*i*" is defined in analogy with classical mechanics:

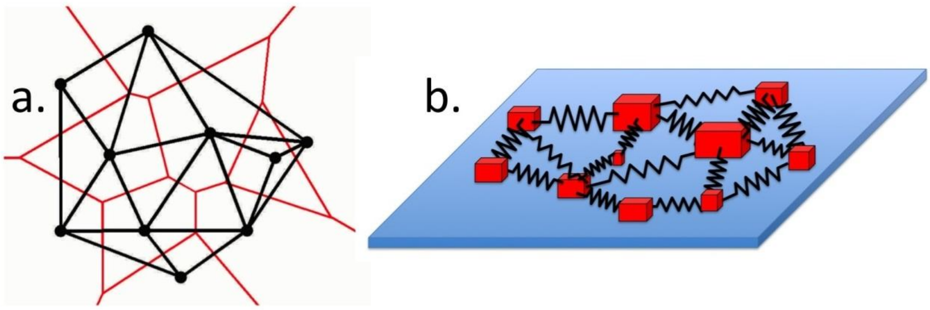
$$F_i^f = m_i g \mu \qquad (1)$$

Here g is a kind of gravitational constant, taken as $g=1$ (defining the units for forces), and μ is the static friction coefficient.

The blocks are interconnected by springs. In order to determine the interconnection topology we use a Delanuay triangulation (Okabe–Boots–Sugihara–Chiu 2000). The Delanuay triangulation performed on the point-like coordinates of the settlements will yield the first-order neighbors (Figure 1a). Springs are placed between these first-order neighbors, and as a result of this, a two-dimensional spring-block lattice will form (Figure 1b). Relaxation is performed on this spring-block lattice.

Figure 1

Construction of the spring-block lattice



a) Delauney traingulation for detecting the neighbors of the settlements. Black dots represent the position of the settlements, red polygons are the Voronoi cells around them and the black lines illustrate the Delauney triangulation. We connect by springs the points connected by these edges.

b) A two-dimensional spring-block system. Red blocks represent settlements of different sizes and the springs connect the neighboring blocks detected through the Delanuey triangulation.

Springs are modeling the socio-economic connections between the neighboring settlements. Settlements that are closer to each other should attract each other in a stronger manner and settlements that have a correlated tendency of evolution for some of their relevant socio-economic indicators should also attract each other in a strong manner. Taking into account these obvious facts, the spring-force between settlement "*i*" and "*j*" is quantified as:

$$\vec{F}_{ij}^s = -\frac{k_{ij}}{d_{ij}} \vec{\varepsilon}_{ij} \qquad (2)$$

In the above equation k_{ij} denotes the connectivity strength between settlements "*i*" and "*j*", d_{ij} is the distance between them and $\vec{\varepsilon}_{ij}$ is a unit vector pointing from on settlement to the other. In contrast with many earlier spring-blocks models, here the spring-forces are not linear. The tension in the spring is increasing while the distance between the settlements gets smaller. Another difference relative to the classical spring-block models is that the here the springs cannot break, and there are able to resist any tension.

The crucial point in defining our spring-block model is to quantify the k_{ij} *connectivity strength*. In order to achieve this, we consider a relevant socio-economic parameter, $q_i(t)$, which is available on settlement level for a relatively long time-period, t. This can be for example the population of the settlement, the GDP per inhabitant or yearly average tax data per inhabitant. Furthermore, we suppose that we do have access to this data with a uniform time-sampling interval, taken here as unity. We can define in such manner for each settlement the relative change of this quantity for a unit time:

$$r_i(t) = \frac{q_i(t+1) - q_i(t)}{q_i(t)} \qquad (3)$$

The main hypothesis for the used spring-block approach is that we assume that for the relevantly connected settlements the time-variation of the r_i relative changes has to be strongly correlated. This means that we can quantify the k_{ij} connection strengths as the time-like correlations between the r_i quantities. Using Pearson correlations, we assume thus:

$$k_{ij} = \left| \frac{\langle r_i(t) r_j(t) \rangle_t - \langle r_i(t) \rangle_t \langle r_j(t) \rangle_t}{\sigma[r_i(t)]\sigma[r_j(t)]} \right| \qquad (4)$$

Here, $\langle \ \rangle_t$ stands for an average taken over time, and σ is the standard deviation:

$$\sigma[r_i(t)] = \sqrt{\langle r_i^2(t) \rangle_t - \langle r_i(t) \rangle_t^2} \qquad (5)$$

Please note the absolute value on the right side of equation (4). This is introduced due to the fact k_{ij} must be positive, and both a high correlation and anti-correlation indicates a strong connection between settlements "*i*" and "*j*".

The Dynamics

Once the main elements of the model, and their relation to the relevant socio-economic and geographic data are clarified, we sketch the dynamics that leads to the hierarchical territory division.

1. Initially the settlements are placed on a plane according to their GPS coordinates. The m_i mass of each block is fixed according to the population of the corresponding settlement ($m_i = W_i$). Neighbors are defined with the Delanuay triangulation and springs are inserted between them. The d_{ij} distances between the neighboring settlements and their k_{ij} connectivity strength is calculated and fixed for the whole duration of the simulation.
2. The resultant elastic force acting on each block is computed: $\vec{F}_i^s = \sum_{j(i)} \vec{F}_{ij}^s$. In the above equation we denoted by $j(i)$ the neighbours of block i.
3. The friction coefficient is fixed to a large value, so that no resultant force acting on the blocks exceeds the friction force: $\forall i \rightarrow \left|\vec{F}_i^s\right| < F_i^f$
4. We decrease the value of μ, until the first block is allowed to slide, i.e. the magnitude of the resultant tension force is bigger than the static friction force acting on the block ($\left|\vec{F}_i^s\right| > F_i^f$).

5. The block "i" in one simulation step will be moved by a distance *h* in the direction of the F_i^s force. The *h* value is chosen to be much smaller than the smallest initial distance between any two settlements. As a result of this slip, if any two blocks come closer to each other than a predefined d_{min}value, these blocks are united in one cluster. This cluster will be a new block in the system, inheriting the mass and links of both blocks and in the next simulation steps will represent all the blocks that are glued together in it as result of this coalescence process.
6. After each move the distances among the moving blocks and it's neighbors are recalculated.
7. In the next simulation step the moving block and it's neighboring blocks are checked whether they satisfy the slipping condition. If yes they are moved by a distance *h* in the direction of the resultant forces. This simulation step is repeated until no block can move. i.e. $\forall i \rightarrow \left|F_i^s\right| < F_i^f$
8. The movement of blocks will continue in the consecutive simulation steps, until the condition $\left|F_i^s\right| > F_i^f$ is satisfied for the new position.
9. When no more blocks can move, the μ friction coefficient is lowered until the first block satisfies the slipping condition again. The simulation proceeds from here by repeating the steps from 5.
10. The dynamics ends when all blocks have collapsed in one.

The above defined model and dynamics has only two freely adjustable parameters. One is the *h* distance of the slip in one simulation step and the other is the d_{min} distance for the collapse of two blocks. If we take them small enough, (as described earlier), their value will not influence the hierarchical collapsing dynamics and the method will become free of adjustable parameters.

Hierarchical territory division

The above algorithm will hierarchically group the elements respecting their connection strength with neighbors and their initial spatiality. Smaller blocks will slide towards large ones, since the friction forces acting on the latter ones are larger. These large settlements will than become centers attracting other blocks. It is reasonable to assume that the above presented spring-block relaxation dynamics will reveal regions at different scales, as the system hierarchically collapse in larger and larger clusters. Keeping track of all these collapsing events, and by projecting backward in time the collapsing order, one can define a hierarchical clustering of the settlements for the investigated geographical region.

The above presented method and model was implemented in an interactive JAVA application (Máté–Néda–Benedek 2009), which allows the user to follow visually the whole clustering process up to the end when all blocks concur in one. The program memorizes all the intermediary situations and can interactively visualize the settlement partitions corresponding to them. One can freely play with this model by visiting the website dedicated to it (Máté–Néda–Benedek 2009).

The hierarchical settlement connection network

The collapsing dynamics obtained in the spring-block model can be used to draw a hierarchical network structure for the settlements. In order to construct this network our hypothesis is that each cluster is dominated by the largest settlement inside it. This means that when two clusters collapse we assign a link between the dominating settlements. In order to visualize different levels of hierarchy and to obtain a visually interpretable network structure we draw lines with different colors or gray-shades according to the collapsing stage. One possibility is to use the whole color palette from blue to red according to the simulation time (step) of the time moment for the collapse. Blue lines will represent the initial stage of the collapse, and dark red lines the latest stages when only few clusters are present in the dynamics. In such manner interpreting the resulted weighted graphs has to be done by taking into account the color code illustrated on the attached legend. Another possibility is to use only warm colors, and indicate the collapse stage with a color shifting from red to yellow. This convention generates more user-friendly networks, but it offers less resolution for illustrating different stages of the collapse. It is also possible to use gray-shades instead of colors. In such case the darker levels will correspond to earlier collapse time and lighter lines will indicate connections made on bigger clusters level. We will illustrate in our examples all these methods.

Data

In order to obtain the connectivity strength (k_{ij}) between the settlements we need a uniformly interpreted long-term dataset for some relevant socio-economic parameter ($q_i(t)$). The most convenient data in such sense is the population W_i of the settlements. One might assume that the variations in the population data will reflect both cultural and economic aspects of the settlements, so it seems appropriate for a well-balanced and relevant socio-economic indicator. It also has the advantage that in general long-term census data is easily accessible for all geographic territories. We have gathered freely accessible data on the Internet for Transylvania on settlement level (Varga 2007) (11 census data between 1850–2002) and for USA on county level (University of Virginia Library n.y) (last 5 census). Due to the courtesy of Professor József Nemes Nagy, we obtained excellent data on settlement level for Hungary (11 census data from 1870 to 1970). For the case of Hungary Prof. Nemes Nagy provided us also the average local income tax paid per inhabitant for 20 consecutive years (data between 1990–2009). This data offered the possibility to use also different socio-economic indicators for determining the k_{ij} connectivity strength.

In order to initially position the blocks (settlements) on the simulation plane and to determine their neighbors, we also needed the GPS coordinates of the settlements. This coordinates are usually available in an electronic format. For Transylvania we got them from (Astroforum n.y) and in the case of USA the data for the county center is available at (Comcast n.y.). In the case of Hungary the data was provided by Prof. Nemes Nagy.

Once all these data are electronically available, the spring-block model can be initialized and it's hierarchical collapse will reveal the relevant network structure.

Results

An interactive version of the program running both on Windows, Linux and MacOsX operation systems can be downloaded from (Máté–Néda–Benedek 2009). Once the Java Runtime Environment (JVE) is installed on the computer and correctly running, one can test the program. In the online version one will be able to select the studied territory (Transylvania, USA or Hungary), the type of data used for the connectivity (population, tax) and the visualization method. It is possible to visualize the "regions" obtained on different hierarchical levels by rewinding the collapsing scenario. An alternative to this is to plot the hierarchical network structure that is intended to synthesize visually the region structures obtained at various collapsing stages. Results obtained for Transylvania, USA and Hungary are given in Figures 2, 3 and 4.

For Transylvania we draw the edges using a color code extended from blue to dark red (Figure 2). Someone who is familiar with the map of Transylvania will immediately spot out on this graph as local hubs (gathering the tree-like connection topology in their neighborhoods) the main region capitals (Cluj-Napoca, Timisoara, Sibiu, Brasov, Oradea, Targu Mures and Baia Mare). In order to illustrate the topology of the network structure on lower lever, we have magnified the Banat region, with the obvious center in Timisoara. This blow up indicates the hierarchical space division and the lower level geographical regions in Banat.

For USA we have used a gray-scale code to illustrate the hierarchical network structure on the county level (Figure 3). Two regions (one in the neighborhood of Memphis, Tenesse, and one in the neighborhood of Salt Lake City, Utah) are magnified to show the fine structure.

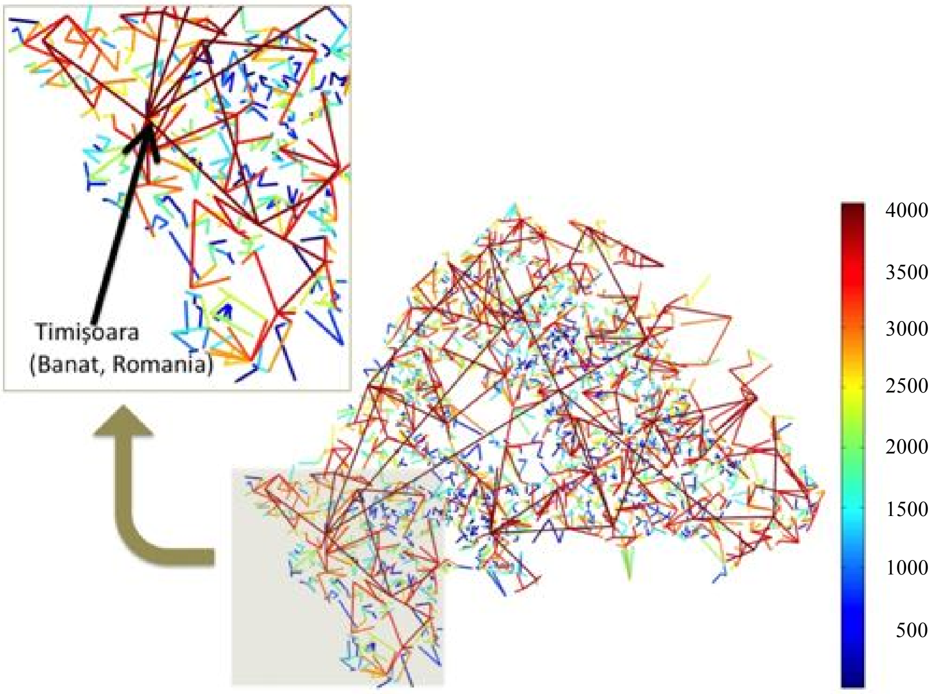
For Hungary (Figure 4) we use yet another representation, by considering only warm colors ranging from red to yellow. Here we compare the settlement network structures obtained using two very different datasets. One network is based on long-term population census (from 1870 to 1970) and the other one is obtained by using data for 20 consecutive years (from 1990 to 2009) of income tax per inhabitant. Although the topology is statistically very similar the two magnified regions show that the obtained network structures are quite different. This illustrates well that the hierarchical networks structures constructed in such manner are rather sensitive to the used data and the spanned time-period.

After the visualization of the results the socio-economic implications will be taken into account. The most important advantage of the applied method lies in the followings:

– The relation between time processes and spatial distribution can be seen in a single map (see methodology). The spatial analysis and visualization of time-based data (time series) is possible with this method. It is based on long-term socio-economic data, and the method helps to discover the space-time relations by different variables. It can serve a basis in bordering of regions, gravity zones of city regions, etc., because the strengths and directions of linkages can be seen in a single map. These maps show the effect of geographical distance on regional structures.

Figure 2

The settlement connection network for Transylvania (Romania) obtained by using long-term population census data.



The color of the lines indicates the simulation time when the collapse in the relaxation dynamics occurred. Colder colors are for earlier stages of the dynamics and warmer colors are for the later stages, as it is indicated in the attached legend. The Banat region, with the obvious center in Timisoara is magnified for giving a better visualization of the local nature of the obtained settlement network.

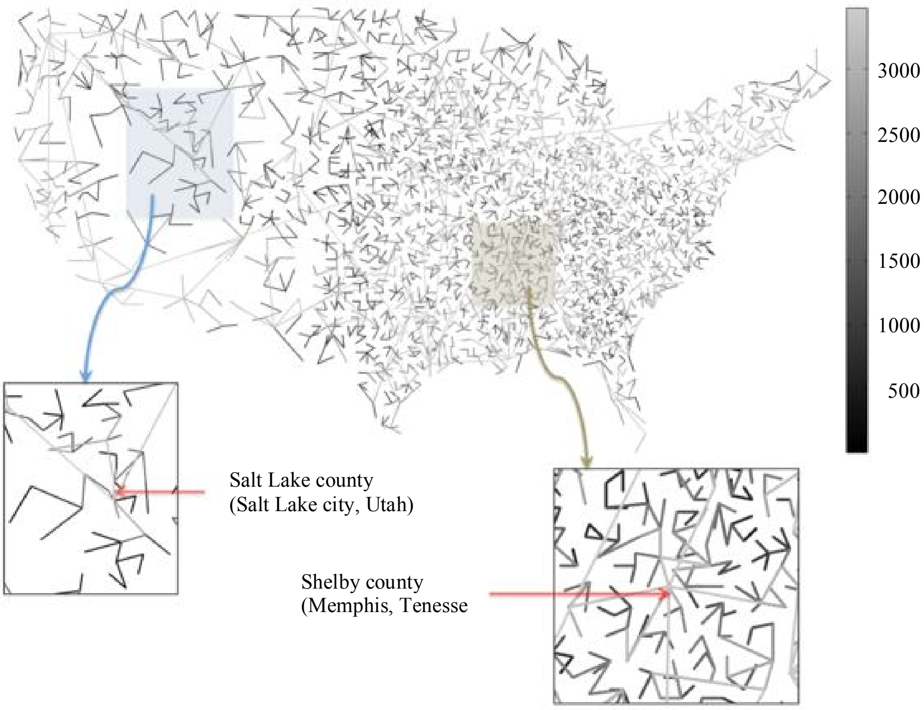
– The level of spatial structure can be determined by the researcher during the visualization process. During the modeling procedure the researcher has the opportunity to determine the desired number of formulating regions. It means that one can draft a multilevel regional structure by stopping the simulation in different stages. In the Transylvanian example a solid vertical structure of micro, mezzo and macro regions can be observed (See the different colors and patterns.).
– The method is suitable for comparing of many different kinds of data with socio-economic relevance. For horizontal and vertical division of socio-economic space strong spatial links and long-term processes are needed. This method allows us to compare the spatial linkages and strength in case of different variables. In the Hungarian example (where we had two variables to work with) we have the opportunity to compare the similarities and differences in spatial links by using population and income data. As one can see in Fig 4, there are considerable differences in the case of the two variables. The maps on population and income have got partly different topology. The map on the census data shows stronger ties

on micro-regional scale than the map on per capita income does. It may come from the different time dimensions of data, from the diverse behavior of social and economic processes, etc.

– The original settlement (or regional) structure of the researched area is depicted in the visualized network. The applied method uses GPS-data in the visualization process, so not only the strength of the link between the neighboring settlements can be identified, but the real settlement network as well. In the maps of the USA and Hungary differences in density and spatial distribution of settlement are easy to identify. In case of Hungary the small village areas of Somogy, Baranya and Cserehát can be distinguished from the region of Alföld.

Figure 3

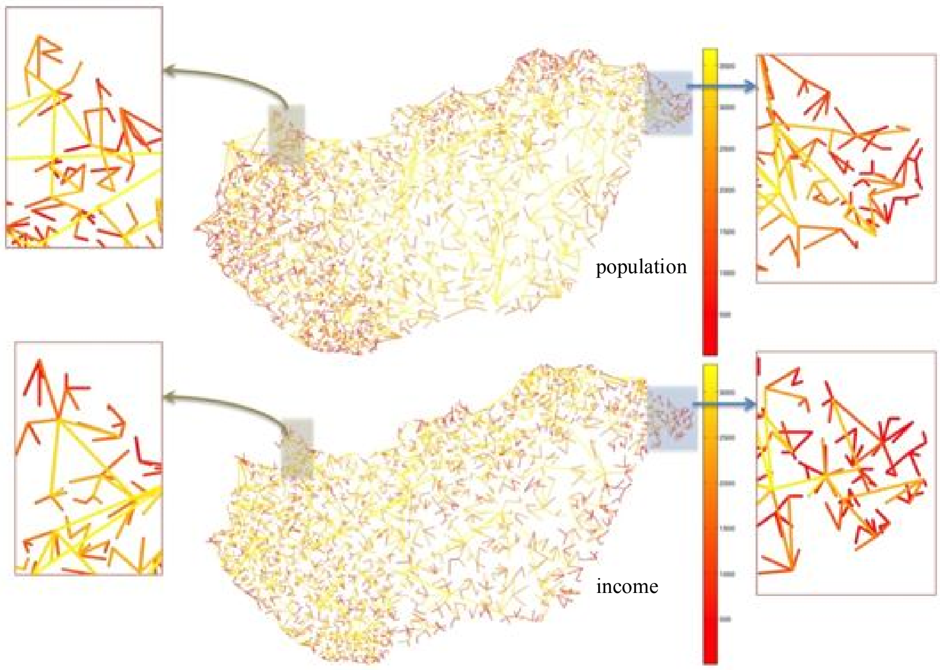
The county connection network revealed for USA by using population census data



A gray-scale code (see the attached legend) indicates the simulation step when the collapse occurred in the relaxation dynamics. Two rectangular regions are magnified for exemplifying the local connectivity topology.

Figure 4

The settlement connection network for Hungary constructed by using long-term population census data (top) and average income tax data for 20 consecutive years (bottom)



The color code indicates the simulation step when the collapse occurred in the relaxation dynamics. Lighter colors (in the direction of yellow) indicate later stages, as it is illustrated on the attached legends. Two regions are magnified. Although the topology looks statistically similar, the networks obtained with the different input data are very different.

Conclusions

The hierarchical settlement connection networks plotted in the present study offers an intuitive and visually appealing presentation for the complex inter-relationship between the settlements. Due to the used "link" definition, the network has a tree-like structure, and no loops are present. Settlements are connected only through the region centers revealed at different spatial scales. These centers act as hubs of different sizes. The shortest distance (minimal number of links) between any two nodes (settlements) will indicate how strongly these settlements are connected. The hierarchy levels are illustrated by using different color codes. Results obtained for the case of Transylvania, USA and Hungary were used to exemplify the method.

Our approach and the applied techniques may help the regional research in several fields. In regional and spatial planning it can contribute to the planning, realization and control phases for determining an optimal regional structure (in advance), to identify regions with special characteristics (backward areas, development axes, city regions, etc.) or to measure and visualize the changes in settlement networks. In business it can contribute to the spatial analysis of market data like spatial relations in revenues, profit, costs, etc., to identify spatial distribution and concentration of business partners (customers, suppliers) or business network elements like market centers, industrial parks, clusters; or to find hierarchies in business activities (competitors, partners).

We believe that such representations could help in designing socio-economic structures that would serve a better and more optimal territorial organization and business activity.

REFERENCES

Barabási, A. L. (2005): Taming Complexity *Nature Physics* 1:68-70

Barabási, A. L. (2012): The network take over *Nature Physics* 8: 14-16

Máté, G. – Néda, Z. – Benedek, J. (2011): Spring-block models reveals region-like structures *PLOS One* 6: 16518

Dusek; T (2004): *A területi elemzések alapjai* [Introduction to spatial analysis] Eötvös Lóránd University. Budapest.

Haggett; P. (2001): *Geography: A global synthesis* Pearson Education, Harlow.

Nemes Nagy, J. (2009): *Terek, helyek, régiók* [Spaces, Places, Regions] Akadémiai Kiadó, Budapest.

Járai-Szabó, F. – Néda, Z. (2012): Earthquake model describes traffic jams caused by imperfect driving styles *Physica A* 391 (22): 5727 – 5738.

Kovács, K. – Néda, Z. (2007): Disorder-driven phase transition in a spring-block type magnetization model *Physics Letters A* 361 (1-2): 18–23

Burridge, R. – Knopoff; L. (1967): Model and theoretical seismicity *Bulletin of Seismological Society of America* 57 (3): 341–371.

Olami, Z. – Feder, J.S. – Christensen, K. (1992): Self-organized criticality in a continuous, nonconservative cellular automaton modelling earthquakes *Physical Review Letters* 68 (8): 1244–1247.

Leung, K. T. – Néda, Z. (2010): Criticality and pattern formation in fracture by residual stresses *Physical Review E* 82 (4): 046118.

Okabe, A. – Boots, B. – Sugihara, K. – Chiu, S.N. (2000): *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams* Wiley, Chichester.

Máté, G. – Néda, Z. – Benedek; J. (2009): Region detection homepage http://atom.ubbcluj.ro/~gmate/regions. Accessed 2012 Jun 3.

Varga, E. A. (2007): *Erdély etnikai és felekezeti statisztikája* [Ethnical and denominational statistics for Transylvania, Census data for 1850-2002] http://www.kia.hu/konyvtar/erdely/erd2002.htm. Accessed 2012 Jun 3.

University of Virginia Library (n.y): *Historical census browser* http://fisher.lib. virginia.edu/collections/stats/histcensus/. Accessed 2011 Jan 6.

Astroforum (n.y): *Coordonate de localită□i din România* [Coordinates of Romanian settlements] http:// www.astroforum.ro/index.php?showtopic=90. Accessed 2011 Jan 6.

Comcast (n.y.): *Coordinates of counties in USA* http://home.comcast.net/~rickking04/ gis/COUNTYSPC.XLS. Accessed 2012 Jun 3.

Acknowledgment

This research was supported in the framework of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program – Elaborating and operating an inland student and researcher personal support system' project. The project was subsidized by the European Union and the State of Hungary. We thank Prof. Jozsef Nemes Nagy for the provided data.

REGIONAL STATISTICS, 2013, VOL. 3: 41–56 DOI: 10.15196/RS03103

ÁRON KINCSES[a)] – GÉZA TÓTH[b)]

Investigation of spatial structure: modelling approach*

Abstract

In this paper the authors wish to introduce an application of the gravitational model through a concrete example. In their investigation the gravitational model was transformed to analyse the impact of accessibility in a way, that not only the size of gravitational forces but their direction can also be measured. Displacements were illustrated by a bi-dimensional regression, which gives a new perspective to the investigation of the Hungarian spatial structure.

Keywords: gravitational model, bi-dimensional regression, accessibility, spatial structure, Hungary.

Introduction

The overall goal of modelling is to simplify reality, actual processes and interactions and on the basis of the obtained data to draw conclusions and make forecasts. Models based on gravitational analogy are the tools of spatial interactions of classical regional analyses. They were first applied in the 19th century (Carey 1858, Reilly 1929, Stewart 1948, Converse 1949, Zipf 1949, Dodd 1950, Hammer–Ikle 1957, etc.).

The application of the geographical gravity is confirmed by the theory of experience according to which (just as in time) the things that are closer to each other in space are more related than distant things. This is called the "first law of geography" (Tobler 1970).

There are two basic areas of the application of gravitational models based on physical analogy: the spatial flow analysis (Filippo, S. et al. 2012, Martinez et. al. 2003), and the demarcation of catchment areas (Vieira et al. 1993, Mate et al. 2011). The potential models based on gravitational analogy are the most important groups of accessible models. In general, it can be stated that they are accessible approaches according to which models show potential benefits of the region compared to other regions where the benefits are quantified (Spiekermann–Wegener–Schürmann 1997).

The use of accessible models in transport-geographical studies is very common. However, when models are used, it is not entirely clear what is actually modelled;

a) Hungarian Central Statistical Office, 1024 Budapest, Keleti Károly út 5-7., Hungary, E-mail: aron.kincses@ksh.hu
b) Hungarian Central Statistical Office, 1024 Budapest, Keleti Károly út 5-7., Hungary, E-mail: geza.toth@ksh.hu

* In this paper the authors wish to introduce an application of the gravitational model through a concrete example. In their investigation the gravitational model was transformed to analyse the impact of accessibility in a way, that not only the size of gravitational forces but their direction can also be measured. Displacements were illustrated by a bi-dimensional regression, which gives a new perspective to the investigation of the Hungarian spatial structure.

because of their complexity their interpretations may be difficult (Kincses–Tóth 2011). It should be stressed that accessibility has no universally accepted definition; in empirical studies different methodological background indicators are used (see Geurs–van Wee 2004, Gutierrez 2001, Tschopp et al. 2003, Willigers–Floor–van Wee 2007). The gravitational theory is a theory of contact, which examines the territorial interaction between two or more points in a similar way as correlations are analysed in the law of gravitation in physics. According to Dusek (2003), despite the analogy, there are significant differences between gravitational models used by social sciences and the law of gravitation used in physics. It is worth bearing in mind that "the gravitational model is not based on the gravitational law”. It is a fundamental statement based on the experience of undeniable statistical character that takes into consideration spatial phenomena. According to this statement, phenomena interact with each other. The phenomena, which are closer to each other in space, are more related than distant phenomena (Dusek 2003, p. 45.).

There are a number of differences between the law and the model. In this study, we wish to highlight a new point of view. As a consequence of the spatial interaction, classical gravitational potential models show the magnitude of potential at spatial points. Regarding the law of gravity in physics, the direction of forces cannot be evaded. In our approach each unit area is assigned an attraction direction. That is, in the case of the gravity model (although such spaces are free of vortex) the space is characterized using vectors.

Method

The universal gravitational law, Newton's gravitational law, states (1686) that any two point-like bodies mutually attract each other by a force, whose magnitude is directly proportional to the product of the bodies’ weight and is inversely proportional to the square of the distance between them (Budo 1970) (1st formula):

$$F = \gamma \cdot \frac{m_1 \cdot m_2}{r^2} \qquad (1)$$

where γ is the proportional factor, the gravitational constant (independent of time and place).

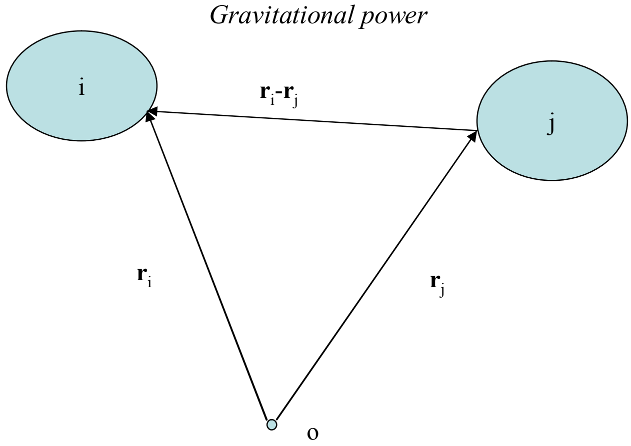
If r indicates the radius vector drawn from the mass point number 2 to a mass point number 1, then r/r is the unit vector drawn from 1 towards 2, thus, the impact of gravitational force on mass point 1 from mass point 2 is in (Equation 2) (Figure 1), which reads

$$\vec{F}_{1,2} = -\gamma \cdot \frac{m_1 \cdot m_2}{r^2} \cdot \frac{\vec{r}}{r} \qquad (2)$$

A gravitational field is set if the gradient (K) can be specified by a direction and magnitude in each point of the range. Since K is a vector quantity, three numbers (two in plane) are required to be known at each point, for example the right angle components of gradient K_x, K_y, K_z, which are functions of the site. However, many fields, including the gravitational field, can be characterized in a more simple way. They can be expressed by a single scalar, the so-called potential function, instead of three values.

Figure 1

Gravitational power



The impact of force on i by j: $\vec{F}_{ij} = -\gamma \cdot \frac{m_i \cdot m_j}{r^2} \cdot \frac{\vec{r}_i - \vec{r}_j}{r}$

The potential is in a similar relation with gradient as work and potential energy with force.

Taking advantage of this, gravitational models are also applied in most social sciences where space is usually described by a single scalar function, (Kincses–Tóth 2011), whereas in gravitational law vectors characterizing the space are of great importance. The primary reason for this is that arithmetic operations calculated with numbers are easier to handle than with vectors. Perhaps, we could say that by working with potentials we can avoid calculation problems in problem solving. The potential completely characterizes the whirl-free gravity gravitational field, because there is a definite relationship between the field strength and the potential:

$$\vec{K} = -gradU \Rightarrow \qquad K_x = -\frac{\partial U}{\partial x}; \quad K_y = -\frac{\partial U}{\partial y}. \qquad (3)$$

In other words, the potential (as mathematical functional) is the negative gradient of field strength. Various types of potentials and models, which are different from the ones directly *based* by the gravitational analogy, but in this case, force effects among space power sources are quite different. In fact, these models differ from each other since the attractive forces remain above a predetermined limit value and within set distances.

The force in a general form is:

$$\left|\vec{F}\right| = C\frac{m_1^{\alpha} m_2^{\beta}}{r^{\gamma}},$$

where *C, α, β,* γ are constants (Barthélemy 2011).

However, how they describe actual power relations between social masses is another question.

Although potential models often characterize concentration on focal points of areas and spatial structures, they fail to provide information in which direction and with how much force the social attributes of other areas attract each delimited area unit. Thus, we attempt to use vectors in order to show the direction where the Hungarian micro regions[3] tend to attract micro regions (LAU1) in the economical space compared to their actual geographical position. This analysis can demonstrate the most important centres of attraction, or discrepancies, and the differences between the gravitational orientations of micro-regions can be displayed on a map after the evaluation of data from 2000, 2005 and 2010 has been performed. In the study the geometric centres of specific micro-regions were the co-ordinates of Hungarian micro-regions, which were determined in the EOV co-ordinate system[4] by (Geographical Information System) GIS software.

Our goal can be reached by using Equation (3) to potentials or directly with the help of forces. We chose the latter one.

In the conventional gravity model (Stewart, 1948) D_{ij} is the "demographic force" between i and j where W_i and W_j are the population size of the settlements (regions), d_{ij} is the distance between i and j, and finally, g is the empirical constant (Equation 5).

$$D_{ij} = g \cdot \left(\frac{W_i \cdot W_j}{d_{ij}^2}\right) \quad (5)$$

In this study, first the W_i and W_j weight factors represent personal income, which is the gross income serving as basis of the personal income tax, dij is the actual distance between i and j regional centres measured on road by a minute (regardless of the traffic conditions and only the maximum speed depending on the road type is taken into consideration). The personal income is the best variable to represent the Hungarian spatial development structure. Later the weight factors will be represented by other variables, which are to model the underlying causes of changes in the development.

By generalizing the aforementioned formula, we write the following equation (Equation 6 and 7):

$$D_{ij} = \left|\vec{D}_{ij}\right| = \frac{W_i \cdot W_j}{d_{ij}^c}$$

(6), (7)

$$\vec{D}_{ij} = -\frac{W_i \cdot W_j}{d_{ij}^{C+1}} \cdot \vec{d}_{ij}$$

where W_i and W_j are the masses, d_{ij} is the distance between them, c is a constant, which is the change of the intensity of inter-regional relations as a function of distance. As the exponent, c, increases and the intensity of inter-regional relations becomes more sensitive to distance, this significance of masses gradually decreases (Dusek 2003). The minus sign express mathematically, that the masses attract each other (see Figure 1).

With the extension of the above Equation we cannot only measure the strength of the force between the two regions, but its direction as well.

3 A micro-region code is the code of a (non- administrative) breakdown covering the whole area of Hungary, which is based on real connections between settlements in terms of work, residence, transport, medium-level provision (education, health care, trade) etc. Through their connections settlements are attracted by one or more central settlements in the system of statistical micro-regions.

4 The EOV is a plane projection system used uniformly for the Hungarian civilian base maps and, in general, for spatial informatics. Geometric classification: Conformal cylindrical projection in transversal position.

While performing calculations, it is worth dividing the vectors into x and y components and summarize them separately. To calculate the magnitude of this effect (vertical and horizontal forces of components) the following Equations are required (Equation 8 and 9), which follow from (6):

$$D_{ij}^{X} = -\frac{W_i \cdot W_j}{d_{ij}^{c+1}} \cdot (x_i - xj) \qquad (8)$$

$$D_{ij}^{Y} = -\frac{W_i \cdot W_j}{d_{ij}^{c+1}} \cdot (y_i - y_j) \qquad (9)$$

where x_i, x_j, y_i, y_j are the coordinates of the i and j regions.

However, if we perform the calculation on all unit areas involved, we will know in which direction their forces exactly act and how strongly they affect the given unit area (Equation 10).

$$D_{i}^{X} = -\sum_{j=1}^{n} \frac{W_i \cdot W_j}{d_{ij}^{c+1}} \cdot (x_i - x_j) \qquad (9)$$

$$D_{i}^{Y} = -\sum_{j=1}^{n} \frac{W_i \cdot W_j}{d_{ij}^{c+1}} \cdot (y_i - y_j)$$

It should be noted that while in potential models, the results are modified by the introduction of "self-potential", in the examination of forces we disregard the introduction of "self- forces".

Thus, it is possible to determine the magnitude and direction of force in which other areas affect each territorial unity. The direction of the vector, which is assigned to the region, determines the attraction direction of other unit areas, while the length of the vector is proportional to magnitude of force. For the sake of mapping and illustration, we transformed the received forces into shifts proportional to them in the following way (Equation 11 and 12):

$$x_i^{\text{mod}} = x_i + \left(D_i^X * \frac{x^{\max}}{x^{\min}} * k \frac{1}{\frac{D^{X\max}}{D^{X\min}}} \right) \qquad (11)$$

$$y_i^{\text{mod}} = y_i + \left(D_i^Y * \frac{y^{\max}}{y^{\min}} * k \frac{1}{\frac{D^{Y\max}}{D^{Y\min}}} \right) \qquad (12)$$

where X_i^{mod} and Y_i^{mod} are the coordinates of new points modified by the gravitational force, x and y are the coordinates of the original point set, the extreme values of which are x^{max}, y^{max}, a x^{min}, y^{min}, D_i is the force along the x and y axes, k is a constant and in this case it is 0.5. This has the effect of normalizing the data magnitudes.

We assume that in our model the amount of interactions between the „masses" is the same as in Equation 7, and based on the superposition principle, it can be calculated for a given region by Equation 10. The new model cannot directly be compared with transport-geographical data, but the results compared with traffic data in potential models to verify our model (Kincses–Tóth 2011).

Our model is a kind of complement to the potential models that ensures a deeper insight into them. In the following sections of this study we intend to communicate some significant results of this model.

Application of two-dimensional regression

The point set obtained by the gravitational calculation (W_i is the gross income serving as basis of the personal income tax, population, tax payers, dwellings built, local taxes, number of active corporations and unincorporated enterprises of the i micro-region in Hungary, d_{ij} is the distance between i and j micro-regions), is worth comparing with the baseline point set, that is, with the actual real-world geographic coordinates and examining how the space is changed and distorted by the field of force. The comparison, of course, can be done by a simple cartographic representation, but with such a large number of points, it is not really promising good results. It is much better to use a two-dimensional regression.

The two-dimensional regression is one the methods of comparing partial shapes. The comparison is possible only if one of the point coordinates in the coordinate systems differing from each other is transformed to another coordinate system by an appropriate rate of displacement, rotation and scaling. Thus, it is possible to determine the degree of local and global similarities of shapes as well as their differences that are based on the unique and aggregated differences between the points of the shapes transformed into a common coordinate system. The method was developed by Tobler, who published a study describing this procedure in 1994 after the precedents of the 60s and 70s (Tobler, 1961, 1965, 1978, 1994). There are many examples using this procedure, which are not necessarily motivated by geographic issues (Kare-Samal–Marx 2010, Symington–Brunsdon–Charlton 2002, Nakaya 2010).

As for the equation relating to the calculation of the Euclidean version, see Tobler 1994, Friedmann–Kohler 2003, Dusek 2011.

Table 1

The equation of the two dimensional regression of Euclidean

1. Equation of the regression	$\begin{pmatrix} A' \\ B' \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \beta_1 & -\beta_2 \\ \beta_2 & \beta_1 \end{pmatrix} * \begin{pmatrix} X \\ Y \end{pmatrix}$
2. Scale difference	$\Phi = \sqrt{\beta_1^2 + \beta_2^2}$
3. Rotation	$\Theta = \tan^{-1}\left(\frac{\beta_2}{\beta_1}\right)$
4.Calculation of β_1	$\beta_1 = \frac{\sum(a_i - \bar{a})*(x_i - \bar{x}) + \sum(b_i - \bar{b})*(y_i - \bar{y})}{\sum(x_i - \bar{x})^2 + \sum(y_i - \bar{y})^2}$
5.Calculation of β_2	$\beta_2 = \frac{\sum(b_i - \bar{b})*(x_i - \bar{x}) - \sum(a_i - \bar{a})*(y_i - \bar{y})}{\sum(x_i - \bar{x})^2 + \sum(y - \bar{y})^2}$
6. Horizontal shift	$\alpha_1 = \bar{a} - \beta_1 * \bar{x} + \beta_2 * \bar{y}$
7. Vertical shift	$\alpha_2 = \bar{b} - \beta_2 * \bar{x} - \beta_1 * \bar{y}$
8. Correlation based on error terms	$r = \sqrt{1 - \frac{\sum\left[(a_i - a_i')^2 + (b_i - b_i')^2\right]}{\sum\left[(a_i - \bar{a})^2 + (b_i - \bar{b})^2\right]}}$
9. Breakdown of the square sum of the difference	$\sum\left[(a_i - \bar{a})^2 + (b_i - \bar{b})^2\right] = \sum\left[(a_i' - \bar{a})^2 + (b_i' - \bar{b})^2\right] + \sum\left[(a_i - a_i')^2 + (b_i - b_i')^2\right]$ SST=SSR+SSE
10. Calculation of A'	$A' = \alpha_1 + \beta_1(X) - \beta_2(Y)$
11. Calculation of B'	$B' = \alpha_2 + \beta_2(X) + \beta_1(Y)$

Source: Tobler (1994) and Friedman, Kohler (2003) based on Dusek 2011 p.14.

Where *x* and *y* are the coordinates of independent shapes, *a* and *b* are the coordinates of dependent shapes, *a'* and *b'* represent the coordinates of dependent shapes in the system of independent shapes. α_1 determines the measure of horizontal shift, while α_2 determines the measure of vertical shift. β_1 and β_2 are the scalar difference and (Φ) and (Θ) determine the angle of shifting.

SST is the total square sum of difference. SSR is the square sum of difference explained by regression. SSE is the square sum of difference not explained by the regression (residual). Further details about the background of the two-dimensional regression can be seen in Dusek (2011) p. 14–15.

Table 2

Bi-dimensional regression between the gravitational and the geographical space under personal income tax

Year	r	α_1	α_2	β_1	β_2	Φ	Θ
2000	0,942	6304,48	2017,44	0,99	0,00	0,99	0,00
2005	0,942	6030,56	2012,23	0,99	0,00	0,99	0,00
2010	0,941	8026,79	2632,29	0,99	0,00	0,99	0,00

Year	SST, %	SSR, %	SSE, %
2000	100,00	98,73	1,27
2005	100,00	98,74	1,26
2010	100,00	98,69	1,31

Source: own calculation.

Our results show that there is a strong relation between the two point systems; the transformed version from the original point set can be obtained without using rotation (Θ = 0). Essential ratio difference between the two shapes was not observed. Comparing the obtained results, it is obvious that the set of points behaves like a single-centre mid-point similarity, when it is diminished. This means that only the attractive force of Budapest can be determined at a national level.

Table 3

Bi-dimensional regression between the gravitational and the geographical space by population

Year	r	α1	α2	β1	β2	Φ	Θ
2000	0,937	17095	5208	0,97	0,00	0,99	0,00
2005	0,938	15536	4808	0,98	0,00	0,98	0,00
2010	0,938	13723	4303	0,98	0,00	0,98	0,00

Year	SST, %	SSR, %	SSE, %
2000	100,00	98,48	1,51
2005	100,00	98,53	1,47
2010	100,00	98,57	1,43

Table 4

Bi-dimensional regression between the gravitational and the geographical space by the number of tax payers

Year	r	α1	α2	β1	β2	Φ	Θ
2000	0,937	14390	4476	0,98	0,00	0,98	0,00
2005	0,937	14199	4458	0,98	0,00	0,98	0,00
2010	0,938	14245	4492	0,98	0,00	0,98	0,00

Year	SST (%)	SSR (%)	SSE (%)
2000	100,00	98,53	1,47
2005	100,00	98,53	1,47
2010	100,00	98,54	1,46

Table 5

Bi-dimensional regression between the gravitational and the geographical space by the number of dwellings built

Year	r	α1	α2	β1	β2	Φ	Θ
2000	0,940	8387	2743	0,99	0,00	0,99	0,00
2005	0,940	7035	2314	0,99	0,00	0,99	0,00
2010	0,941	4531	1524	0,99	0,00	0,99	0,00

Year	SST, %	SSR, %	SSE, %
2000	100,00	98,66	1,34
2005	100,00	98,64	1,36
2010	100,00	98,69	1,31

Table 6

Bi-dimensional regression between the gravitational and the geographical space by local taxes

Year	r	α1	α2	β1	β2	Φ	Θ
2000	0,945	2333	774	1,00	0,00	1,00	0,00
2005	0,944	2816	924	1,00	0,00	1,00	0,00
2010	0,945	2679	883	1,00	0,00	1,00	0,00

Year	SST, %	SSR, %	SSE, %
2000	100,00	98,84	1,16
2005	100,00	98,82	1,18
2010	100,00	98,84	1,16

Table 7

Bi-dimensional regression between the gravitational and the geographical space by the number of active enterprises

Year	r	α1	α2	β1	β2	Φ	Θ
2000	0,941	7870	2534	0,99	0,00	0,99	0,00
2005	0,942	6984	2272	0,99	0,00	0,99	0,00
2010	0,942	6265	2052	0,99	0,00	0,99	0,00

Year	SST, %	SSR, %	SSE, %
2000	100,00	98,69	1,31
2005	100,00	98,72	1,28
2010	100,00	98,74	1,26

The observed pairs of points are strongly correlated with each other. The analysis of local taxes showed the strongest correlation. The values of beta1 may provide the most important relationship. These unambiguously indicate an east-west divide in Hungary.

Map display and direction analyses

The aforementioned statement can be illustrated by a map display of a two-dimensional regression. The Darcy program can be used in the application (http://www.spatial-modelling.info/Darcy-2-module-decomparaison).

The square grid attached onto the shape-dependent coordinate system and its interpolated modified position further generalizes the information received from the participating points.

The arrows in Figures 2 and 3 show the directions of the shifts, while the colouring illustrates the type of distortion. The warm colours express the divergent forces of the area, which are considered to be the most important gravitational displacements.

The areas illustrated in green and its shades represent just the opposite, namely the most important nodes of gravity.

The data in Table 2 shows that the space shaped by the gravitational model causes only a slight distortion compared to the geographic space. The magnitude of vertical and horizontal displacements increased slightly in 2010.

Practically, the maps produced by Darcy software verify this (Figure 2 and Figure 3). It can be seen that the capital of Hungary is Hungary's main centre of gravity, the centre towards which the largest power is attracted. The regional centres like Győr, Pécs, Szeged, Debrecen are also gravity nodes. The national role of regional centres is weak. In the area of Budapest a gravity fault line emerges.

The reason for this phenomenon is that the Hungarian capital attracts all the micro-regions, while very weak forces are applied to Budapest compared to its mass.

The map also illustrates that the regular force fields are the major transport corridors, namely they are slightly distorted due to highways. Between 2000 and 2010, the significance of green-marked gravitational nodes increased. The comparison of the two maps clearly shows the intensification of regional differences.

Figure 2

Visualizing the gravity field on the basis of incomes, 2000

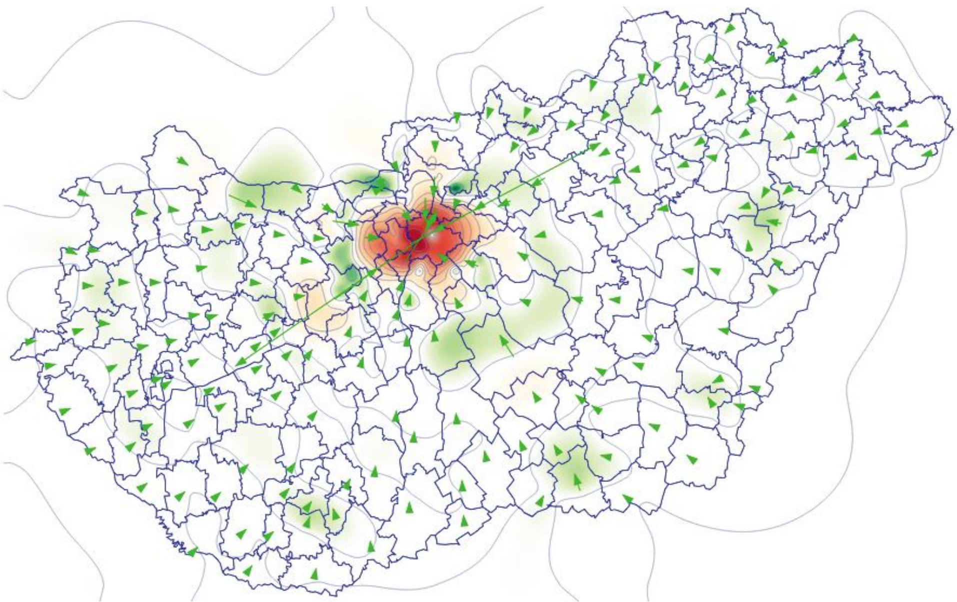
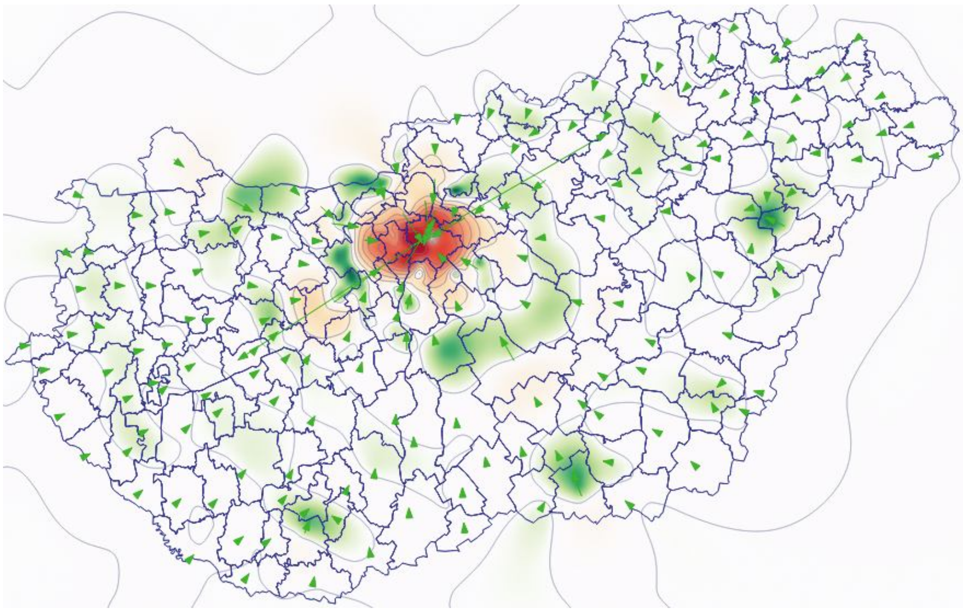


Figure 3

Visualizing the gravity field on the basis of incomes, 2010



We try to analyze what socio-economic factors resulted in the development conditions of 2010 through using the above introduced indicators to model how the gravity field changed. The population number in the micro regions of the western section of the Budapest agglomeration sharply fell clearly indicating a gravity breaking point due to the reversal of previously prevailing suburbanization processes. There are similar depressed areas in the micro regions of Eger and Mezőkövesd. The relationship of the micro regions of Debrecen and Balmazújváros showed the strongest gravity node. In contrast with this, there are only smaller sub centres at the south western tip of Lake Balaton and at the node determined by the micro regions of Esztergom, Tatabánya and Bicske.

Concerning the number of taxpayers, Budapest has a clear 'node position' and there is another, related, gravity centre in the south western tip of Lake Balaton. There are smaller quasi sub centres, on the one hand, around some regional centres (Győr, Debrecen) and, on the other hand, in some more dynamic micro regions (Tatabánya). The most significant divide is between the node in the surroundings of Lake Balaton and the micro regions lying south west of that. In this respect, there are relatively few significant 'fault lines' in the gravity field.

Figure 4

Changes in the gravity field based on population figures, 2000–2010

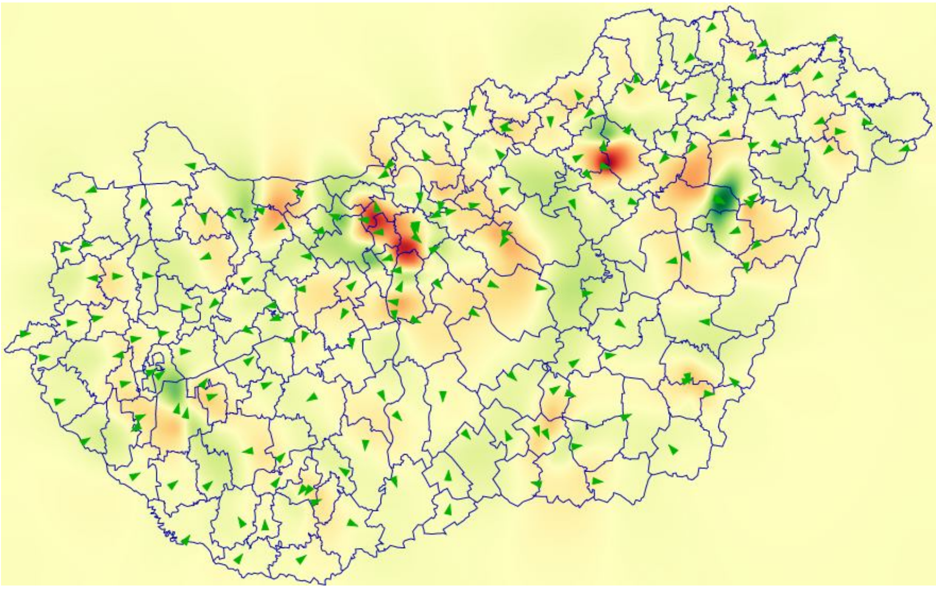
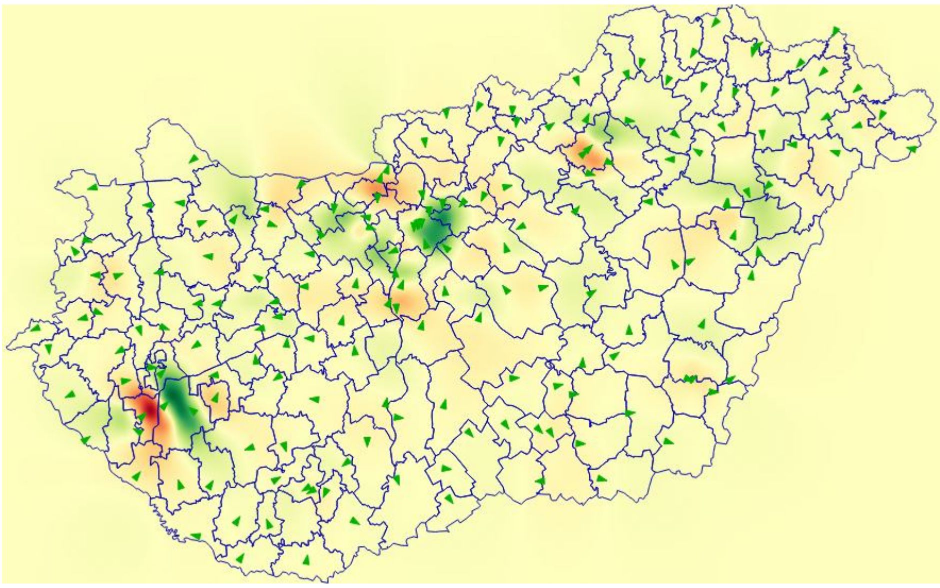


Figure 5

Changes in the gravity field based on the number of tax payers, 2000–2010

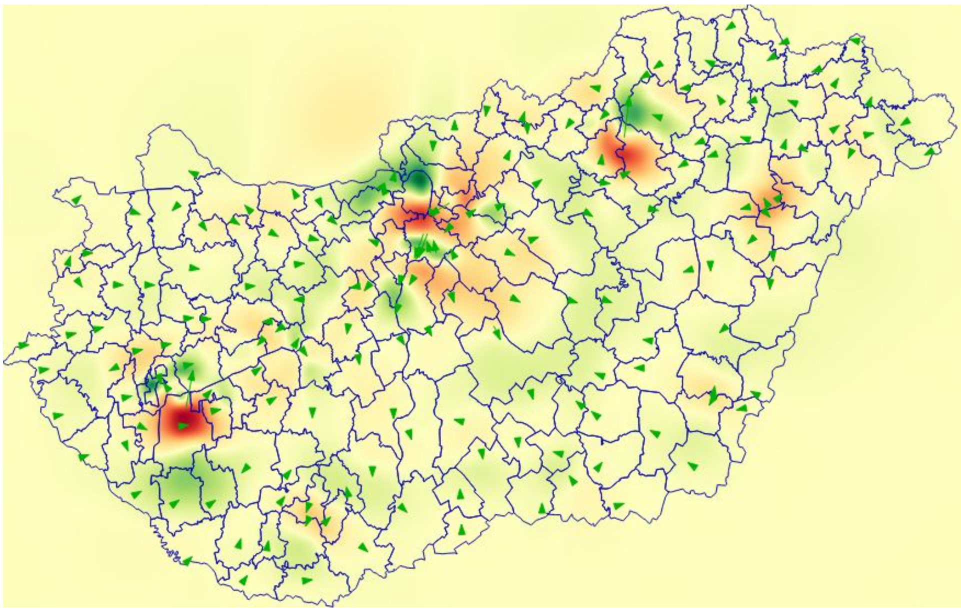


In comparison with the previous ones, a slightly different picture emerges in the number of dwellings built. On the one hand there are several gravity centres and breaking points. A significant centre can be seen in the northern part of the agglomeration of

Budapest. Based on the breaking point lying south of this centre we can say that the rise in the number of dwellings built was implemented at the expense of the capital city. It reflects the constantly changing processes of suburbanization. There is a gravity centre at the regional hub of Miskolc, as well as south of Budapest and in the surroundings of the western tip of Lake Balaton (micro regions of Hévíz and Keszthely as well as the micro-region of Nagyatád). The reasons may be different in case of these micro-regions, i.e. tourism plays a determining role in the former ones, while in case of the latter one – as we will show among active enterprises – it results from the favourable effects of industry.

Figure 6

Changes in the gravity field based on the number of dwellings built, 2000–2010



Maybe the gravity map of local taxes shows the most varied patterns. In this respect, Budapest and the regional centres can be identified as the most important nodes, but the micro regions of serious industrial base, e.g. those of Esztergom and Székesfehérvár deserve attention as well. The most significant gravity 'fault line' is between these micro regions and the agglomeration of Budapest. The previously mentioned micro region of Hévíz is another gravity node in this respect.

Figure 7

Changes in the gravity field based on local taxes, 2000–2010

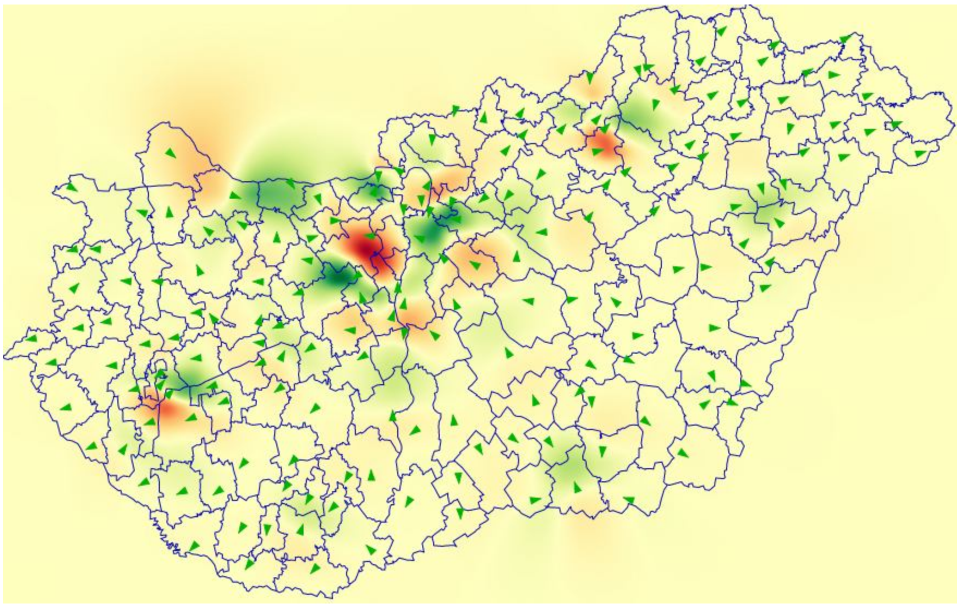
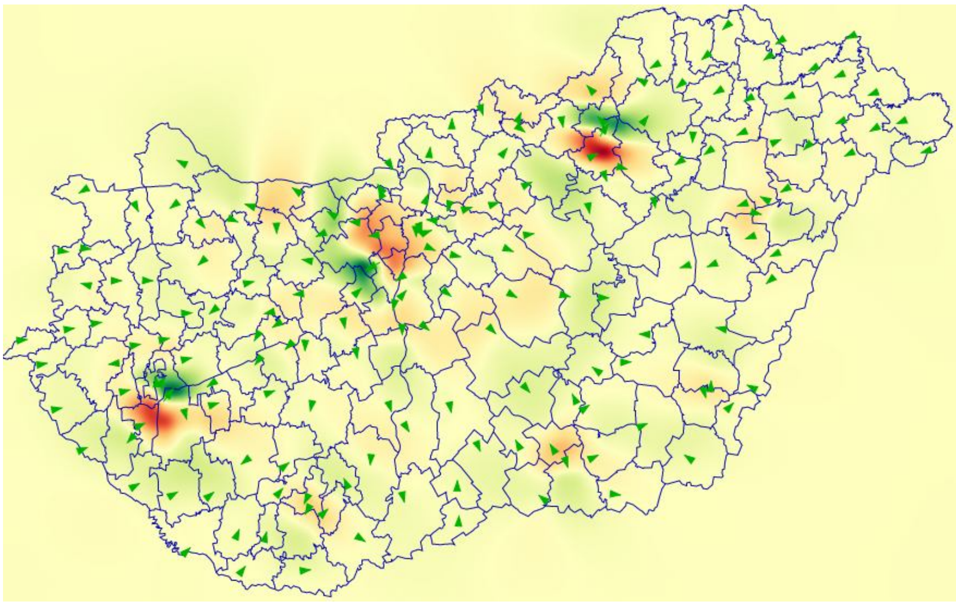


Figure 8

Changes in the gravity field based on the number of active enterprises, 2000–2010



As for the number of active enterprises, the micro regions of Hévíz, Székesfehérvár, Miskolc and Eger play a central role. There are another, smaller centres of local importance, e.g. the micro region of Nagyatád with the surrounding areas, which is separated by a sharp breaking point from the micro regions of Keszthely and Hévíz.

Summary

In our study we made an attempt to introduce the potential and unexplored areas of gravitational models and problems of their interpretation by expanding and extending the methodology. The forces applied were illustrated by using the income tax base as weights in micro-regions of Hungary. On the basis of the model, the result in line with the experience illustrated that Budapest has no counterweight in Hungary and the local central areas are weak. However, the presence of stable local centres is detectable in Debrecen, Miskolc, Nyíregyháza, Szeged and Pécs micro-regions.

There are several indicators to highlight and thoroughly analyze what socio-economic processes are behind a basic change in the spatial structure.

REFERENCES

Budó, Á. (1970). *Experimental Physics I.* National Textbook Press, Budapest.

Carey, H. C. (1858). *Principles of Social Science*. J.B. Lippincott & Co., New York.

Converse P.D. (1949). New Laws of Retail Gravitation. *Journal of Marketing* 14 3, 379–384.

Dodd S. G. (1950). The Interactance Hypothesis. A Gravity Model Fitting Physical Masses and Human Groups. *American Sociological Review* 15 2, 245–256.

Dusek, T. (2003). A gravitációs modell és a gravitációs törvény összehasonlítása. *Tér és Társadalom* 17 (1), 41–57.

Dusek. T. (2011). Kétdimenziós regresszió a területi kutatásokban. Területi Statisztika 51 1, 11–22.

Filippo, S., Maritan, A. and Néda, Z (2012). Continuum approach for a class of mobility models, [physics.soc-ph] 19 Jun 2012, http://arxiv.org/pdf/1206.4359.pdf

Friedman, A., Kohler, B. (2003). Bidimensional Regression: Assessing the Configural Similarity and Accuracy of Cognitive Maps and Other Two-Dimensional Data Sets. *Psychological Methods*, 8 4, 468–491.

Geurs, KT., van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography* 12 2, 127–140.

Gutierrez, J. (2001). Location, economic potential and daily accessibility: an analysis of the accessibility impact of the high-speed line Madrid–Barcelona–French border. *Journal of Transport Geography* 9 4, 229–242.

Hammer C., Ikle, F.C. (1957). Intercity Telephone and Airline Traffic Related to Distance and the Propensity to Interact. *Sociometry* 20 4, 306–316.

Kare,S., Samal,A. and Marx, D. (2010). Using bidimensional regression to assess face similarity.. *Machine Vision and Applications* 21 3, 261–274.

Kincses, Á., Tóth, G. (2011). A potenciálmodellek geometriája. *Területi Statisztika* 51 1, 23–37.

Barthélemy, M. (2011). Spatial network. *Physics Reports*, 499: 1-101 http://arxiv.org/abs/1010.0302.

Vieira,M., Giovani L., Vasconcelos, and Sidney R. Nagel (1993). Dynamics of spring-block models: Tuning to criticality Physical Review E. 47 4., R2221–R2224

Martinez-Zarzoso, I., Nowak-Lehmann, F. (2003). Augmented gravity model. An empirical application to Mercosur-European Union trade flows. *Journal of Applied Economics* 6 2, 291–316.

Mate, G., Neda, Z. and Benedek, J. (2011). Spring-Block model reveals region-like structures. *PLoS ONE* 6(2): e16518. doi:10.1371/journal.pone.0016518.

Nakaya, T. (2010). Statistical Inferences in Bidimensional Regression Models. *Geographical Analysis* 29 2, 169–186.

Reilly, W. J. (1929). *Methods for the Study of Retail Relationships*. University of Texas Bulletin, No. 2944

Schürmann, C., Spiekermann, K. and Wegener, M. (1997). *Accessibility Indicators*. Berichte aus dem Institüt für Raumplanung 39, Dortmund, IRPUD

Stewart J. Q. (1948). Demographic Gravitation: Evidence and Applications. *Sociometry* 11 1-2, 31–58.

Symington, A., Charlton, M.E. and Brunsdon, C.F. (2002). Using bidimensional regression to explore map lineage *Computers, Environment and Urban Systems* 26 2, 201–218.

Tobler, W. R. (1961). *Map Transformations of Geographic Space*. PhD dissertation, Seattle, University of Washington.

Tobler, W. R. (1965). Computation of the Correspondence of Geographical Patterns. *Papers of the Regional Science Association*. 15 1, 131–139.

Tobler, W. R. (1970). A Computer Model Simulating Urban Growth in the Detroit Region. *Economic Geography*. 46 2, 234–240.

Tobler, W. R. (1978). Comparisons of Plane Forms. *Geographical Analysis* 10 2, 154–162.

Tobler, W. R. (1994). Bidimensional Regression. *Geographical Analysis* 26. 3, 187–212.

Tobler, W. R. (2004). On the First Law of Geography: A Reply *Annals of the Association of American Geographers* 94 2, 304–310.

Tschopp, M., Fröhlich, P., Keller, P. and Axhausen, K.W. (2003). *Accessibility, Spatial Organisation and Demography in Switzerland through 1920 to 2000: First Results*. 3rd Swiss Transport Research Conference Monte Verità / Ascona, March 19-21. 2003.

Willigers, J., Floor, H. and van Wee, B. (2007). Accessibility indicators for locational choice of offices: An application to the intraregional distributive effects of high-speed rail in the Netherlands. *Environment Planning. A* 39 9, 2086–2098.

Zipf G. K. (1949). *Human Behaviour and the Principle of Least Effort: An Introduction to Human Ecology*. Addison-Wesley Press, Cambridge.

REGIONAL STATISTICS, 2013, VOL. 3: 57–78 DOI: 10.15196/RS03104

CSILLA OBÁDOVICS[a)]

Regional level analysis of the population by real data and projections in Hungary

Abstract

Population projections of László Hablicsek (1953–2010) have major scientific contribution to understanding the demographic processes of Hungary; its social and economic usefulness cannot be questioned. Although, trends of the past ten years indicate the need for some corrections. The present analysis shows that the population estimation was very close to reality in the past ten years. There was no significant difference between the estimated and real data; only last year's data show some alteration. However, the projections for smaller territorial scales are in need for serious corrections. In recent years, regional disparities in the country rather increased. The migration towards large cities, moving from undeveloped areas to more developed ones raise dissonance among regions. The population concentration proved to be stronger than the forecast, and the analysis showed, that the regional concentration of the population as well as the loss of population has further reinforced the unfavourable position of underdeveloped regions at micro-regional level. Immigration from abroad further enhances regional differences, since the immigrants presumably do not prefer rural, disadvantaged areas, they most probably favour economically developed micro-regions which are close to big cities. The observed changes in the socio-economic environment require corrections in the demographic projections, which is a natural need of the professional audience as well as actors in social policy.

Keywords: long-range population projection, demographic trends, economic development.

László Hablicsek, expert of population projections in the HCSO Demographic Research Institute, was the author of numerous demographic studies and the leader of many researches. He is credited with the long-range population projection and estimation in Hungary. He performed the population estimation on the basis of the 2001 census data and projected the changes in the population number by different projection variants until 2050. Later, in 2006, he updated and corrected the model on the basis of the 2005 micro-census. He applied the Cohort component method[2], in the course of which he took into account demographic phenomena directly influencing the population number, such as fertility, mortality and migration.

a) Faculty of Economics, University of West Hungary, H-9400 Sopron, Erzsébet út 9, Hungary, E-mail: csilla.obad@gmail.com

2 For more details about the method see H. Richter Mária (2002).

His projections made in the past decade will still be in use for long. Hablicsek established the database for regional population, qualification and activity/inactivity projections, the so-called PQW (population, qualification, workforce) system. This is built up by the projections of the following factors: population, educational attainment, economically active population, employed, jobseekers, economically inactive population, full-time students, people rearing child(ren), pensioners, dependents and other inactive people.

The present study aims at reviewing the results of the Hungarian population projection procedures in commemoration of the recently passed away scientific researcher. On the basis of some analyses, studies, research reports and articles, as well as the data of HCSO, the study investigates how the population of Hungary changed in the last forty years and then compares the real data with the estimated ones on regional, county and, finally, on micro-regional level.

In addition to the country-level population projections, the study analyses the relation between micro-regional demographic inequalities and socio-economic development with the help of multivariable statistical methods using the data of micro-regional population projections as well.

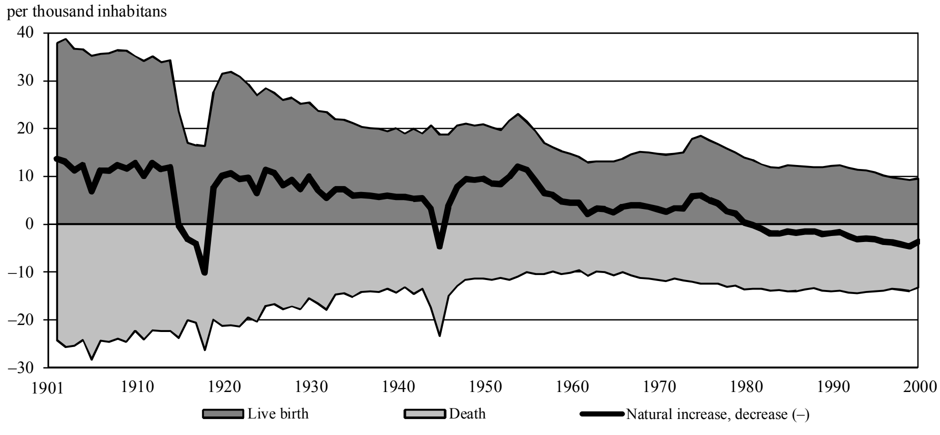
Population projection variants on country level

Population projection is in fact an estimation method, in the course of which the number of births, the number of deaths, migration processes and the changes in life expectancy at birth are taken into account.

The population of Hungary increased until 1980, then, a slow decrease started. The population number grew by 4% between 1970 and 1980, while it fell by 4% between 1980 and 1990. Figure 1 presents the changes in the number of births and deaths as well as the difference between them. The figure well demonstrates that the natural increase of the population turned to natural decrease in 1980.

Figure 1

Birth and death rates between 1901 and 2000



Source: nepszamlalas2001.hu

László Hablicsek prepared the projections in different kinds of variants. The variants were prepared as a function of demographic processes taking into account their different changes in the future. He made the estimations for genders, age cohorts, educational attainment, economic activity, but, in the practice, there are examples for projecting the number of the elderly, the pensioners, or even the number of the school-age population.[3]

In line with the different trends in the number of children, life expectancy at birth and balance of migration[4], the Hablicsek projection was prepared in the following variants: (1) baseline variant, (2) young variant, (3) old variant, (4) low and (5) high variant and (6) European variant. This latter variant shows demographic trends targeted on a shorter run, thus, it is not covered by the present study (Hablicsek 2009). The basic characteristics of indicators relating to each projection variant are included in Tables 1 and 2, while Figure 2 compares the population estimations calculated on the basis of the different variants on country level.

Table 1

Components of population projection variants

Projection variants	Averag number of children	Average age of women at childbirth	Life expectancy	Balance of international migration
Baseline	medium	medium	medium	medium
Old	low	low	high	low
Young	high	high	low	high
Low	low	low	low	low
High	high	high	high	high

Source: Hablicsek (2009).

The baseline variant seems to represent the real future. The low variant corresponds to the pessimistic, while the high variant to the optimistic estimation. The young and the old projections show only slight deviation from the baseline variant. The difference between the young and the old variants is also only half million persons (Hablicsek 2009).

Table 2

Characteristics of population projection variants

Category	Average number of children, child	Average age of women at childbirth, years	Life expectancy men/women, years	Balance of international migration, persons
Low	1.3	29	72.6/80.8	8 000
Medium	1.5	31	75.3/83.0	15 000
High	1.8	33	78.0/85.2	22 000

Source: Hablicsek (2009).

In the baseline variant, medium number of children (fertility rate grows from 1.3), medium life expectancy (75.5 years for men, 82.5 years for women) and medium migration, i.e. an annual surplus of 14–15 thousand persons on the long run are assumed.

3,3 http://fogalomtar.eski.hu/index.php/Népesség-előreszámítás.

In the higher version of the baseline variant, an annual international migration surplus of 30 thousand persons is assumed on the long run along with low life expectancy, which is however slightly higher than at present, and medium number of children.

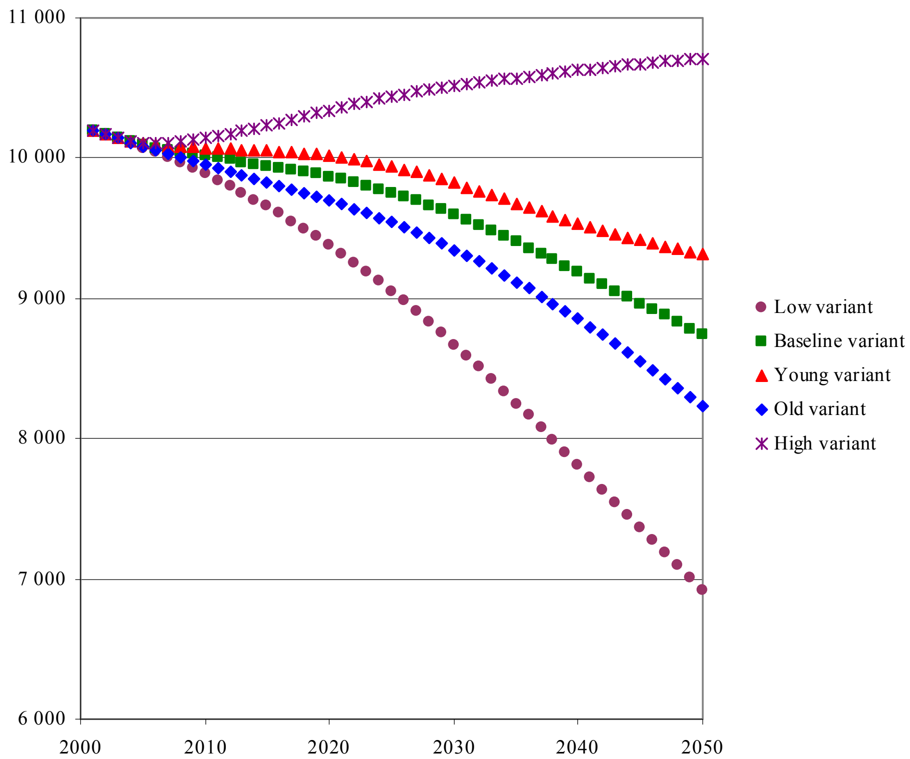
In the low version of the baseline variant, the number of children is also medium, life expectancy is high (82 years for men and 88 years for women) as opposed to the low balance of migration. (Due to the emigration/further migration of foreigners having immigrated to Hungary, the surplus of the migration balance decreases to 7 thousand persons (Polónyi–Tímár 2006).)

In the young variant, the number of children is high (fertility rate increases from 1.3 to 1.6 and then to 1.9), life expectancy is low, while migration is high. Finally, in case of the old variant, each component is reversed, i.e. low number of children (fertility rate remains 1.3), high life expectancy and low migration are assumed.

According to the estimation of Hablicsek (2001), if the baseline variant was associated with a positive balance of migration of 40 thousand persons every year (47 thousand immigrants, 7 thousand emigrants every year), the population number would steadily remain 10 million even along with a relatively low fertility and a slowly increasing life expectancy.

Figure 2

Change in the number of population according to the different projection variants



Sources: Own calculation based on the data of HCSO.

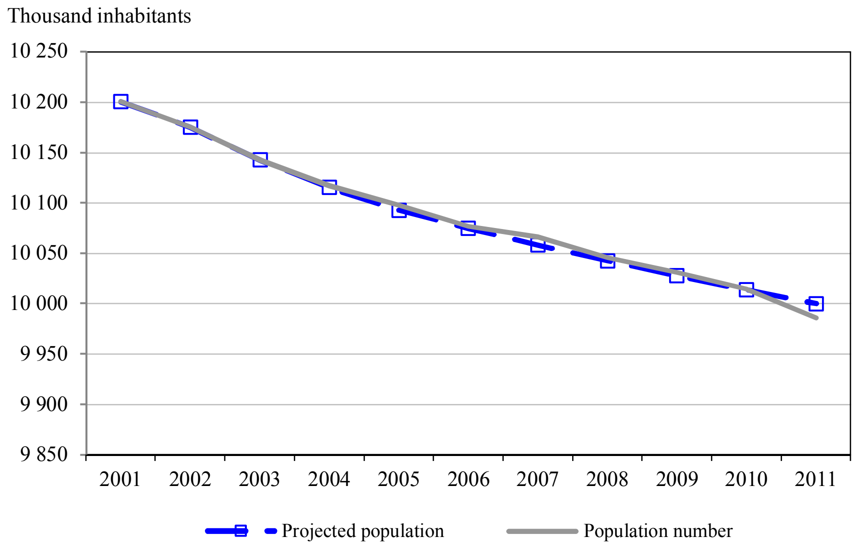
The long-range projection of Hablicsek analyzed in my present study is based on the census 2001. The 10 years since that time is not a really long period in this context, however, it is worth examining whether there is any difference, and if yes, how large the difference is between the projection and the real population data.

Out of the different population projection variants, the baseline variant approximated reality the most accurately for the past ten years. However, it could not be foreseen in 2001 that, by the end of the decade, the economic crisis would exert its effect on demographic processes as well, so a slightly larger difference could be observed between the projection and the real data of the last year (Figure 3).

In addition to the economic crisis, modifications in social allowances and changes in tax laws and family support could also influence our indicators of natural population movement. We have to be of course aware that measures aiming at increasing childbearing intention may change the trend towards another variant, or unfavourable economic processes may result in approaching just the old variant. All these effects are of course only probable, and there are only speculations about their extent.

Figure 3

Population number and projected population number between 2001 and 2011 (baseline variant)



Sources: Own calculation based on the data of HCSO.

It is well demonstrated by the figure that the values of estimated and real population numbers are very close to each other on the time-base. When subjecting the two curves to statistical control as well, the results can be considered very good. (The relative dispersion of the residua is 5.5%.) Except for three data points, each estimated value is within the confidence interval, the data of 2007 showed somewhat lower population number than the real data, while the estimation for 2011 was somewhat higher than the real data, i.e. the real data of these two years showed an upward and a downward deviation, respectively from the estimations. This deviation is presumably not the result of internal coherences in demographic processes, but it can be much more attributed to the social policy, the support system and the economic changes.

Regional projections

In regional breakdown, the results are not nearly so clear. Demographic processes show great diversity just as socio-economic processes. For that very reason, population projection made only on country level is not sufficient, it is necessary to present the differences within the country as well. The regional, county-level and micro-regional projections prepared by László Hablicsek also come from this recognition making a good step forward in demographic projections and in demonstrating the connections between social and economic policies.

According to projections, Hungary will probably record an 8.3% population loss by 2021 compared to 1980. A special attention is paid to regions diverging from the national trend in negative or positive direction (Table 3).

Table 3

Population projection and change in population size between 1980 and 2021 by regions

Year	Southern Great Plain	Southern Trans-danubia	Northern Great Plain	Northern Hungary	Central Trans-danubia	Central Hungary	Western Trans-danubia	Country
1980[a)]	1 464 658	1 059 160	1 590 901	1 400 079	1 120 956	3 033 056	1 040 653	10 709 463
1990[a)]	1 397 627	1 015 783	1 546 612	1 323 508	1 110 302	2 966 523	1 014 468	10 374 823
2001[a)]	1 380 383	997 668	1 563 709	1 302 833	1 116 721	2 831 095	1 007 860	10 200 269
2006[b)]	1 347 294	970 700	1 533 162	1 261 489	1 108 124	2 855 670	1 000 142	10 076 581
2011[c)]	1 320 040	943 002	1 503 758	1 221 183	1 098 640	2 917 461	993 030	9 997 114
2016[c)]	1 292 479	915 471	1 473 083	1 182 618	1 085 648	2 974 491	982 664	9 906 454
2021[c)]	1 268 030	889 153	1 446 569	1 148 000	1 072 055	3 026 871	972 378	9 823 056
Population number (previous year=100%)								
1980								
1990	95.4	95.9	97.2	94.5	99.0	97.8	97.5	96.9
2001	98.8	98.2	101.1	98.4	100.6	95.4	99.3	98.3
2006	97.6	97.3	98.0	96.8	99.2	100.9	99.2	98.8
2011	98.0	97.1	98.1	96.8	99.1	102.2	99.3	99.2
2016	97.9	97.1	98.0	96.8	98.8	102.0	99.0	99.1
2021	98.1	97.1	98.2	97.1	98.7	101.8	99.0	99.2
Population number (1980=100%)								
1980	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1990	95.4	95.9	97.2	94.5	99.0	97.8	97.5	96.9
2001	94.2	94.2	98.3	93.1	99.6	93.3	96.8	95.2
2006	92.0	91.6	96.4	90.1	98.9	94.2	96.1	94.1
2011	90.1	89.0	94.5	87.2	98.0	96.2	95.4	93.3
2016	88.2	86.4	92.6	84.5	96.9	98.1	94.4	92.5
2021	86.6	83.9	90.9	82.0	95.6	99.8	93.4	91.7

Source: Own calculation based on VÁTI TEIR.

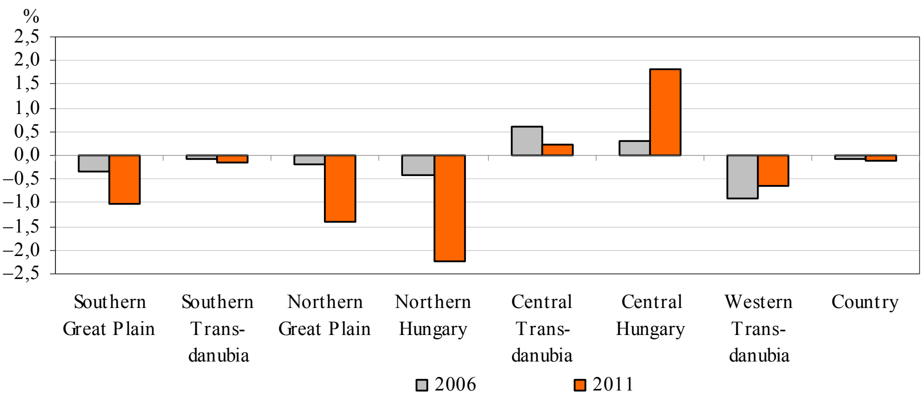
a) Real data. b) Further calculation. c) Projection.

The most drastic population loss is expected in Northern Hungary and Southern Transdanubia. The population of Southern Great Plain will decrease more than in the whole country, and its annual decline will slightly exceed the national average. A positive deviation can be observed in case of Central Hungary and Central Transdanubia. In these two regions, stagnation or even a population increase could be observed in some periods. Central Hungary is the only region of Hungary, where, compared to 1980, the population number practically will not change by 2021, and since the new projection in 2006, the population has been increasing. It must be mentioned however, that this region comprises Budapest as well. Despite the considerable population loss in Budapest, a positive change is probable in the total population size of the region, which can be attributed to the very significant development, the population attracting ability, the favourable age-structure and the high fertility rate of the agglomeration. Although there was some fluctuation at the beginning of the observed period, between 1990 and 2001, the population loss was here the largest in the country, but this negative peak was largely offset by the positive population balance in the following period.

The population dynamics of Central and Western Transdanubia is more favourable than in the whole country. The following figure presents the comparison of the estimated population number with the real population data for the years 2006 and 2011.

Figure 4

Difference between real and projected population as a percentage of the population by regions*



Source: Own calculation based on the data of HCSO.
* An indicator per real population number: (real-estimated)/real.

The divergence between the population projection and the real data in 2006 is less than 1% in each region. In 2011, the projected value slightly exceeds the real one in Northern Great Plain and Northern Hungary, while in Central Hungary, the population number is underestimated. This points to the fact that the estimation was made with lower in-migration and fertility for Central Hungary and with lower out-migration and higher fertility for Northern Hungary and Northern Great Plain than the real trends were.

In the divergence from the estimated migration, the economic crisis could play a role, which the population projection obviously could not reckon with.

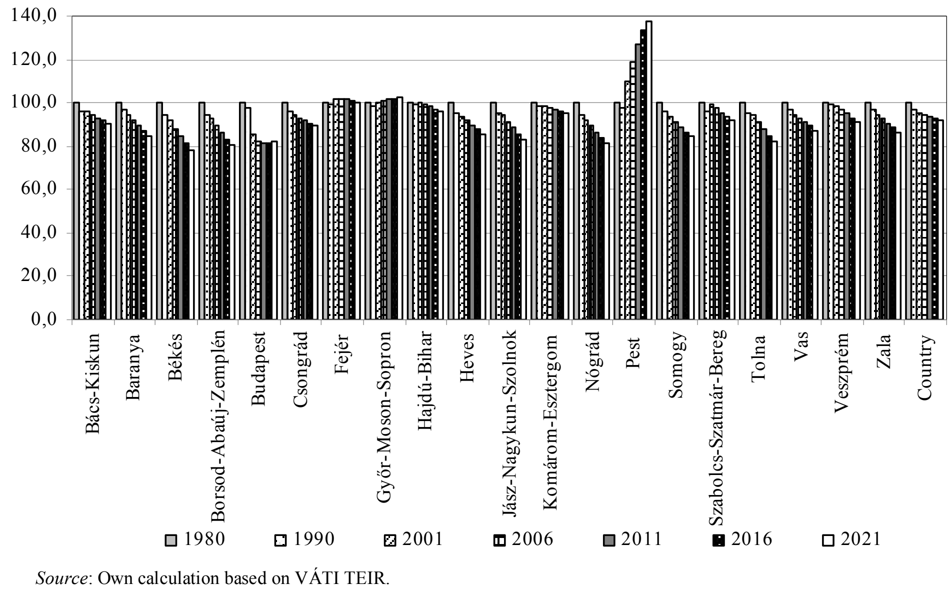
County level population projection

After regions, the population projections for smaller territorial units are compared to the real trends. It is again well demonstrated in Figure 5 that the population increase in Central Hungary described in the foregoing is due to the growth in Pest county.

A significant population growth can be expected only in Pest county, while a slight one in Fejér and Győr-Moson-Sopron counties. The most drastic fall will occur in Békés county, where the population number will decrease by 22% compared to 1980, while a nearly 20% population loss will occur in Tolna, Nógrád, Borsod-Abaúj-Zemplén counties and in Budapest according to the projections.

Figure 5

Probable change in population size between 1980 and 2021 by counties (1980=100%)

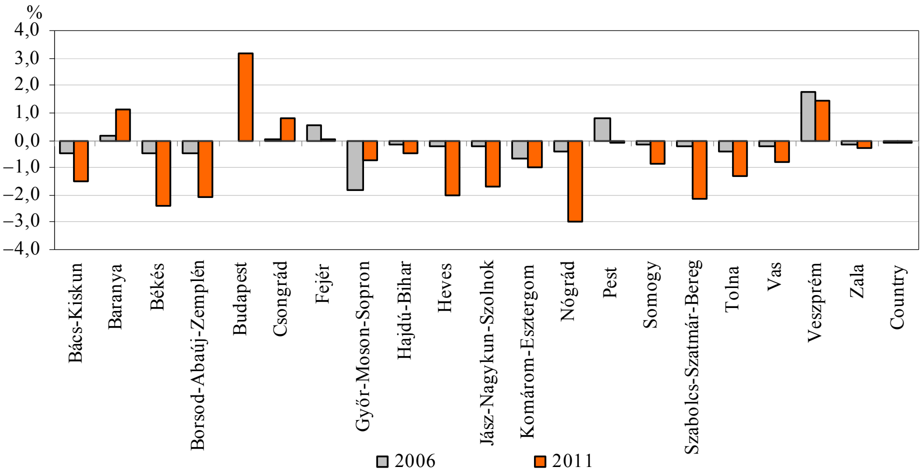


Source: Own calculation based on VÁTI TEIR.

The population projection for 2006 differs significantly from the real data only in two counties (the difference as a percentage of the population is more than 1%): it was overestimated in Győr-Moson-Sopron county and underestimated in Veszprém county.

Figure 6

Difference between real and projected population as a percentage of the population by counties



Source: Own calculation based on VÁTI TEIR.

In 2011, the difference exceeded 2% in five counties, only the population size of Budapest was underestimated by the projection model, while in case of Békés, Borsod, Nógrád and Szabolcs-Szatmár-Bereg counties, it was overestimated.

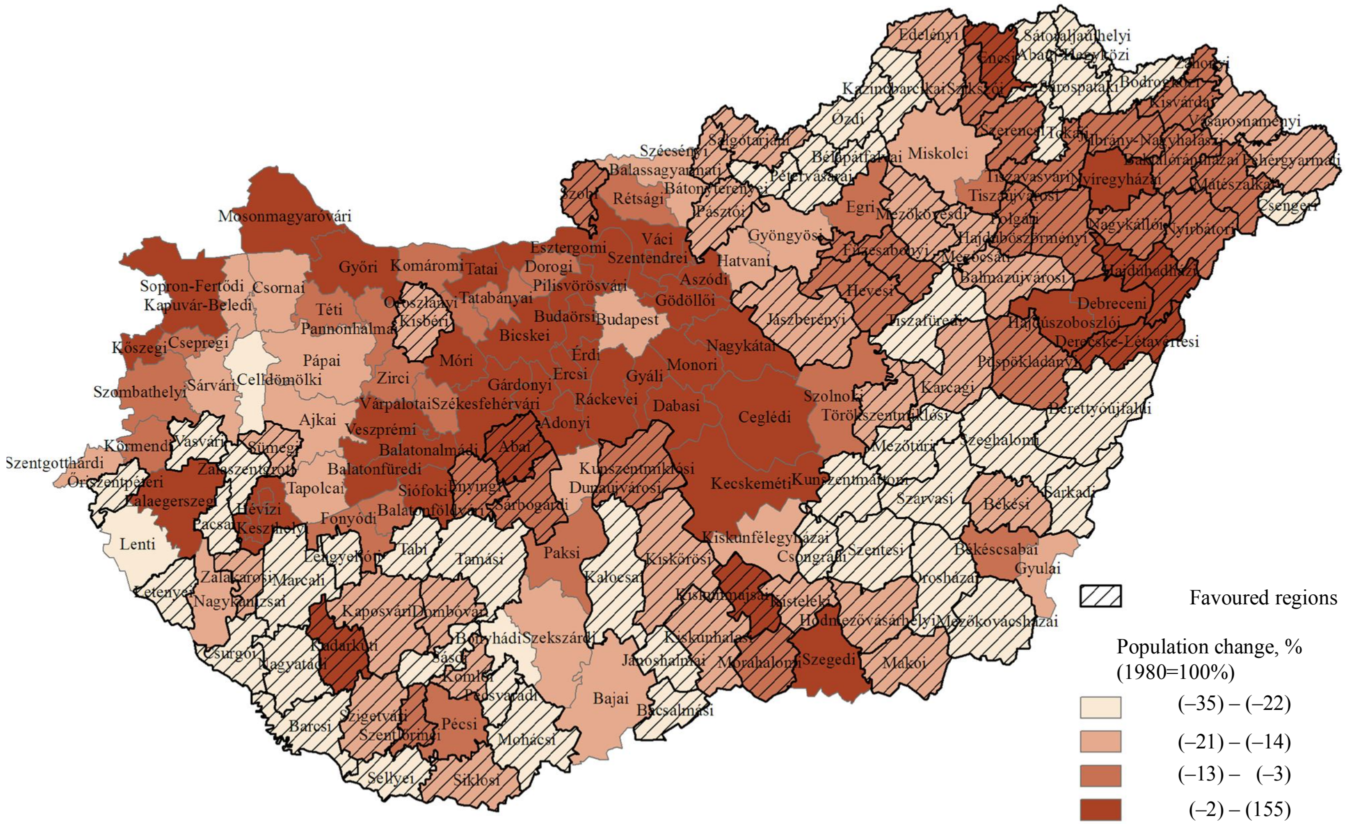
The economic crisis deepened inequalities, which brought large uncertainties into the system, and this also contributed to the fact that larger differences could be observed between projections and real population data of 2011. Projections cannot take into account the effects of such changes of course, especially for the reason that, even if these prevail for a relatively short period, demographic trends follow already a different line after the 'restoration'.

Micro-regional population projection

By examining population dynamics on micro-regional level, we can come to two conclusions. The population of most micro-regions is continuously decreasing, which is not surprising, as the tendency is similar for the population of the whole country. There are however 37 micro-regions, where the balance of population is positive, and in another 26 micro-regions, the population loss is lower than the national average. There are micro-regions in very good position, which belong to the catchment area of large cities and county seats. Eight micro-regions are in the direct catchment area of Budapest (Veresegyháza, Gödöllő, Érd, Budaörs, Szentendre, Pilisvörösvár, Ráckeve, Dunakeszi), and another eight micro-regions belong to the larger catchment area of the capital. From these micro-regions, Budapest is easily accessible within one hour. Although the Gárdony micro-region is not located in Central Hungary, the transport conditions are very good there as well, and according to projections, its population will increase by some one third by 2021 compared to 1980.

Figure 7

Projected population change between1980 and 2021 and favoured regions*



Sources: own calculation based on VÁTI TEIR.
* *According to Tipold–Faluvégi (2007).*

The other, much more sensitive issue of the study is the considerable population growth of micro-regions densely populated by Roma people, e.g. the micro-region of Hajdúhadház (25.3%).

While in case of county seats and the agglomeration of Budapest, all demographic trends affect positively the population increase, i.e. the fertility rate, the balance of migration, the balance of natural increase/decrease and the age-structure are all favourable, in the micro-region Hajdúhadház, the population growth results only from the outstandingly high fertility rate, which offsets the effect of the negative trend in migration (Hablicsek 2007b).

25.5% of the favoured micro-regions are in Northern Great Plain, 22.3% in Northern Hungary, 20.2% in Southern Great Plain, but this rate is nearly 20% in Southern Transdanubia as well with 18 favoured micro-regions. Central Hungary, Central Transdanubia and Western Transdanubia together account for only 12.8% of all favoured micro-regions. The serious situation in Northern Great Plain is also shown by the fact that 86% of all micro-regions of the region are favoured ones, i.e. out of the 28 micro-regions there, 24 belong to this category. The proportion of disadvantaged micro-regions is high is Southern Transdanubia, Northern Hungary and Southern Great Plain as well (72%, 75% and 76%, respectively).

Micro-regional population projections and real data

In 2006, the population projection on country level differed from the real population number only by 0.06%. This difference was 0.1% in 2011. Both values are negative, i.e. a smaller population decrease was assumed than the real trend was.

Table 4

Difference between real and projected data as a percentage of the population, 2006

Micro-region	Difference, %	Micro-region	Difference, %
Gödöllői	1.08	Adonyi	–10.12
Siófoki	1.16	Balatonföldvári	3.05
Dunakeszi	1.35	Salgótarjáni	1.79
Pilisvörösvári	1.37	Bodrogközi	1.07
Budaörsi	1.45	Mezőtúri	1.06
Ráckevei	1.49	Őriszentpéteri	1.01
Veresegyházi	2.11		
Bátonyterenyei	2.35		
Dunaújvárosi	2.58		

Source: Own calculation.

Based on this indicator, the well separable groups of micro regions are as follows:

– In 2006, micro-regions, which surpassed expectations, i.e. where the population growth exceeded the projection by more than 1%, were, apart from a few exceptions, dynamically developing regions belonging to the Budapest agglomeration.

– Overestimated population number with a more than 1% difference was observed in 6 micro-regions; among them, the difference is striking, more than 10% in the Adony micro-region..

Table 5

Difference between real and projected data as a percentage of the population, 2011

Micro-region	Difference, %	Micro-region	Difference, %
Ráckevei	2.08	Kiskunmajsai	–15.52
Nyíregyházai	2.11	Adonyi	–14.75
Sopron-Fertődi	2.54	Bodrogközi	–8.62
Székesfehérvári	2.72	Füzesabonyi	–6.19
Pécsi	2.88	Polgári	–5.64
Dunaújvárosi	2.95	Bélapátfalvai	–5.60
Veszprémi	3.03	Sarkadi	–5.52
Váci	3.09	Mezőkovácsházai	–5.48
Budapest	3.16	Mezőcsáti	–5.45
Szegedi	3.22	Balatonföldvári	–5.44
Debreceni	3.27	Szerencsi	–5.17
Bátonyterenyei	3.56	Hevesi	–5.16
Lengyeltóti	4.14	Salgótarjáni	–5.03
Gárdonyi	4.48		
Dunakeszi	6.86		

Source: Own calculation.

In 2011, the projected population number in the catchment area of some big cities was less than the real data, while, first of all, in some micro-regions of Northern Hungary and Northern Great Plain overestimated population numbers can be observed.[5]

Thus, the above results indicate that the extent of migration towards larger towns is higher than the predicted trends. Due to lack of jobs, young people and those with higher educational attainment move probably to a larger extent to towns than it was expected by the projection. These are of course only assumptions, exploring the reasons and checking the hypotheses need further research..

Regional factors influencing population dynamics: educational attainment and the effect of the Roma population

In the opinion of Hablicsek (2006), regional differences will diminish in the following period due to the „educational boom”. If it is true, the white micro-regions in Figure 8 will shift to the grey ones. In my opinion, however, the out-migration of people with higher educational attainment from disadvantaged regions will continue. Detailed migration data for analyzing the reason for migration, the educational attainment and age

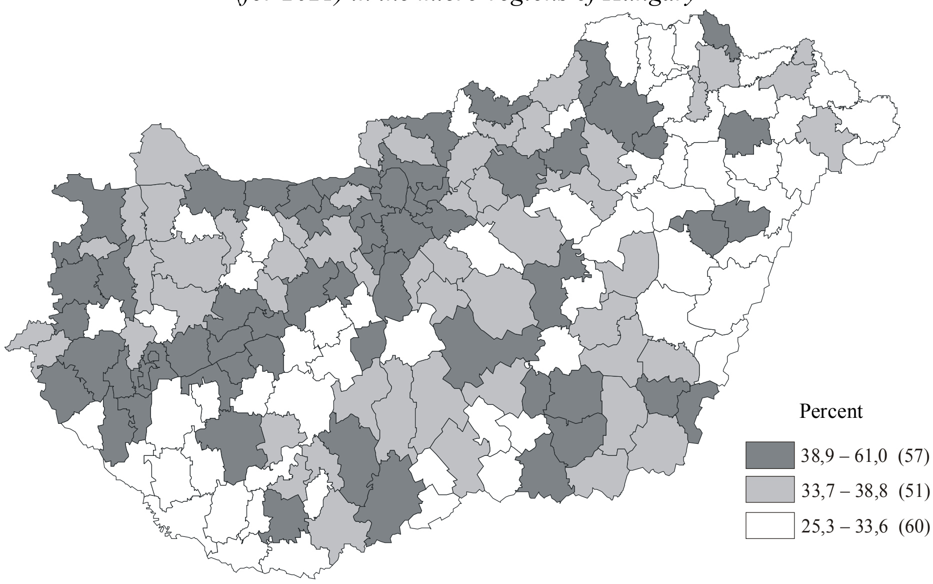
5 The interpretation of micro-regional data and drawing the conclusion need however care. The composition of micro-regions is continuously changing (the number of micro-regions was 138 in 1994, 150 from 1997, 168 from 2004, 174 from 2007 and 175 from 1 January 2011), settlements were transferred from one micro-region to another or even from one county to another, which distors the results. Although differences deriving from the changes are tried to be homogenized in the VÁTI TEIR system, it does not work perfectly, and the long-range population projection does not naturally ccalculate with the separation of settlements (http://hu.wikipedia.org/ wiki/Kistérség).

of migrants are unfortunately not available. My hypothesis is that those with higher educational attainment do not find a proper job in the disadvantaged regions, therefore, they settle down in towns or in their catchment areas. Thus, the "educational boom" induces not the decrease but just the increase of regional inequalities, i.e. regions suffering considerable population loss will be faced with further depopulation. The scientific literature has already indicated this trend, since, in contrast with all previous expectations, regional inequalities rather increase than diminish (Spéder 2002, Kulcsár 2009, Bódi 2010).

The correlation between educational attainment and belonging to the Roma population is of course very close. It is well known that the educational attainment of the Roma population is by far below that of the non-Roma population (Polónyi–Tímár 2006). In the regions afflicted by the crisis, the proportion of the Roma population is outstandingly high. On the one hand, this was explained by the fact that out-migrants with higher educational attainment have better chances to find a job in case of out-migration as well. The educational attainment of the Roma population is generally lower, so they would find a job with more difficulties even if they moved, therefore, they rather stay. On the other hand, in regions with declining economy, their chances for out-migration are worse due to the lower property prices, while these regions may be therefore attractive in respect of settling. All these result in the increase in the proportion and the concentration of the Roma population in these economically underdeveloped regions (Kertesi–Ábrahám 1996).

Figure 8

Projection of proportion of people with at least secondary educational attainment (for 2021) in the micro-regions of Hungary



Sources: Hablicsek (2007a).

The projections of László Hablicsek covered the estimation of the proportion of the Roma population as well. This is shown in Table 6.

Table 6

Number and proportion of Roma population between 1991and 2021

Region	Estimated number of Roma population, thousand persons				Proportion of Roma population, %			
	1991	2001	2011	2021	1991	2001	2011	2021
Budapest	40.4	59.6	80.5	103.2	2.0	3.4	4.8	6.1
Central Hungary without Budapest	65.8	97.0	132.7	173.0	2.2	3.4	4.6	5.7
Central Transdanubia	22.8	31.0	39.8	49.5	2.1	2.8	3.6	4.6
Western Transdanubia	22.2	26.4	30.2	33.7	2.2	2.6	3.0	3.5
Southern Transdanubia	63.5	72.4	79.3	84.9	6.3	7.3	8.4	9.6
Northern Hungary	123.8	151.5	181.9	215.5	9.4	11.6	14.9	18.8
Northern Great Plain	114.5	129.6	145.0	161.3	7.4	8.3	9.6	11.2
Southern Great Plain	35.5	41.8	48.7	56.3	2.5	3.0	3.7	4.4
Hungary, total	448.1	549.7	657.6	774.2	4.3	5.4	6.6	7.9
Change on country level, %, previous data=100		+22.7	+19.6	+17.7				

Source: Hablicsek (2007b), own calculation.

According to the projection of László Hablicsek, the population number will be around 9 million by 2050, of which the Roma population will be one million, i.e. 11% of the population will belong to the Roma minority. He estimated the number of immigrants at one million as well. In the baseline variant, the increase in the Roma population shows a declining trend, and the decrease in the number of the country's population will probably decelerate as well.

Along with this, the proportion of the Roma population in the total population will continue to grow[6], and by 2050 it may reach or exceed 12% (Polónyi–Tímár 2006, Hablicsek 2007b).

In three regions, i.e. in Northern Hungary, Northern Great Plain and Southern Transdanubia, the high proportion of disadvantaged micro-regions and the high proportion of Roma population prevail in parallel. The exception is Southern Great Plain, where the proportion of disadvantaged micro-regions is high (more than 20%), but the proportion of Roma population is quite low (4.44%).

Socio-economic development and population dynamics

Demographic trends are partly the reasons and partly the consequences of the socio-economic situation of a region. Researchers often assume significant correlation between the state of development and the age-structure of a region. The generally accepted opinion is that in underdeveloped settlements and regions with limited resources and services the old-age dependency ratio is increasing.

6 According to the projections of Hablicsek (2000), the number of Roma population increases yearly by 9–10 thousand.

However, it can be demonstrated that the favourable age-structure is not by all means accompanied with advanced socio-economic development. In some micro-regions of Hungary, the disadvantaged situation and underdevelopment is coupled with favourable age-structure. In other ones, the good geographical position (it is near a town, there are tourist attractions, etc.) partly offsets the unfavourable age-structure and the demographic disadvantage.

Migration is one of the most important indicators of regional socio-economic inequalities. Migration from the less developed regions to the more developed ones is induced among others by the better job opportunities (Teaford 2008, Brown–Glasgow 2008). However, this situation has to be examined with a more tinged approach. Some disadvantaged regions are characterized by the ageing of the population and out-migration, while in others the situation is just the opposite. Thus, it is not true in all instances that in underdeveloped regions the population is ageing and they are left by young people, while the developed ones are always characterized by increasing population and young age-structure.

Hereinafter, I will examine the correlation of the features of the age structure with the population projection and the socio-economic development. I present a typology, which pictures the relation between the dynamics of the age-structure and the economic development which is not simple at all.

As a first step of the multivariate statistical analysis, I applied a variance reduction method, the principal component analysis (PCA)[7] in order to define the main socio-economic characteristics. In the second step, by involving the resulted principal components I grouped the micro-regions by a cluster analysis.[8]

Table 7

Rotated component matrix

Indicators	1st principal component: economic force	2nd principal component: young population
In- and out-migration, %[a)]	0.907	0.119
Number of enterprises per 1,000 inhabitants[a)]	0.903	0.065
Unemployment rate[b)], %[a)]	–0.860	0.281
Logarithm of income/capita[a)]	0.855	–0.106
Population change, %[c)]	0.836	0.418
Change in dwelling stock %[c)]	0.835	0.304
Proportion of new dwellings, %[a)]	0.833	0.307
Number of passenger cars per 100 persons[a)]	0.790	–0.414
Population density, persons/km^2[a)]	0.697	0.304
Proportion of population under 14 years of age, %[a)]	–0.234	0.931
Birth rate per 1,000 persons[a)]	0.001	0.888
Old-age dependency ratio, %[a)]	–0.243	–0.817
Number of persons per households[a)]	0.289	0.790

Source: Own calculation.

a) In 2009. b) Proportion of jobseekers in the population aged 18–65 years. c) 2009/2001.

7 For more details about the method see Bartók (1983), Francia (1976) és Czirfusz (2010).

8 I worked with the data of the VÁTI TEIR system in the analysis.

The KMO value (how the variables fit in the model) is 0.839, which shows that the variables involved in the analysis fit well in the PCA model. In the applied principal component analysis, the total variance explained is 78.3%, if only eigenvalue components higher than one are taken into account. Two principal components meet this requirement. After rotation (applying varimax rotation), the variance explained by the first principal component is 50%, while the explanatory value of the second factor is 28.3%.

The first two strongest variables are the in-migration/out-migration ratio and the number of enterprises per 1,000 persons. The correlation between the principal component and the variables is very strong. In the first principal component, the third strongest variable is the unemployment rate, which is connected with negative sign to the principal component. There are some further demographic and economic indicators in the first principal component, therefore, I named this factor as 'economic development' factor.

The second factor correlates the best with the proportion of young people. The old-age dependency ratio is connected with a negative, while the birth rate is connected with a positive sign to this factor. Therefore, this factor could be named as 'young population' factor. Based on the results, we can say that the factors can be unambiguously identified and are suitable for further analysis.

In Table 8, micro-regions having extreme factor values are highlighted.

Table 8

Micro-regions with extreme factor values

Micro-region	Economic development	Young population
Őriszentpéteri	0.18	–2.69
Hévízi	1.27	–2.03
Gödöllői	2.27	1.35
Monori	2.06	1.32
Ráckevei	2.51	1.73
Budaörsi	3.63	1.87
Dunakeszi	3.84	2.09
Pilisvörösvári	2.28	1.44
Szentendrei	2.23	0.95
Veresegyházi	3.06	2.61
Érdi	2.69	1.47
Hajdúhadházi	–0.62	2.49
Edelényi	–1.46	2.13
Encsi	–1.22	2.49
Szikszói	–1.25	2.04
Baktalórántházai	–1.17	2.08

Source: Own calculation.

Among micro-regions having extreme values, there are economically developed, dynamically developing ones with young population near Budapest (Gödölli, Monor, Ráckeve, Budaörs, Dunakeszi, Pilisvörösvár, Szentendre, Veresegyháza, Érd) on the one hand and economically underdeveloped micro-regions having however young population (Hajdúhadház, Edelény, Encs, Szikszói, Baktalórántháza micro-regions, where the proportion of Roma population is high) on the other. The third group is made up by the

Hévíz micro-region with high proportion of old population but with quite good economic potential and by the Őriszentpéter micro-region with moderate economic conditions.

In order to create a micro-regional typology based on the correlation of the two examined factors, I made a cluster analysis on the factors. As a result of the cluster analysis based on the factor values, micro-regions are divided in four groups. The cluster centre values for the two factors are included in Table 9, while Figure 9 shows the location of the clusters.

Table 9

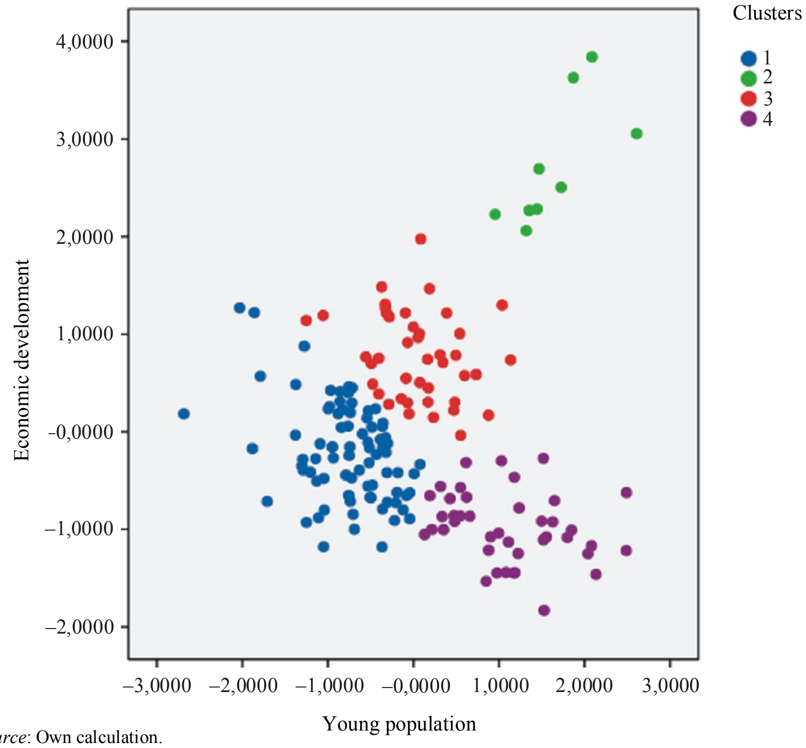
Cluster centres

Factors	Cluster			
	1	2	3	4
Economic development	–0,212	2,730	0,777	–0,978
Young population	–0,770	1,647	0,050	1,110
Number of micro-regions	81	9	42	41

Source: Own calculation.

Figure 9

Location of micro-regional clusters in the factor space



Source: Own calculation.

The aim of the cluster analysis[9] is to group the micro-regions by their economic force and age-structure. As a first step, I performed a hierarchical cluster analysis, then, I

9 For more details about cluster analysis see Obádovics–Popovics (2011).

applied K-mean method. Finally, after comparing the results, I accepted the cluster result of the K-mean method as final.

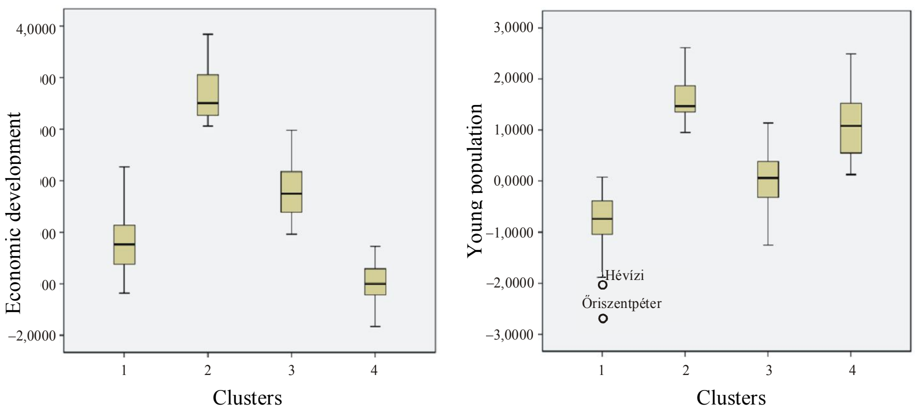
Based on the cluster centres, the groups can be characterized. It is well shown in the table that the fourth cluster is characterized by a favourable age-structure and young population, but it is comprised of micro-regions with the most underdeveloped economy. At the same time, in the second cluster, the favourable age-structure is accompanied by the strongest economic development. The poor, ageing micro-regions belong to the first cluster, while the ones slightly more developed than the average and having medium-level age-structure are in the third cluster.

The differences in the features of the clusters are shown by their spatial locality along the cluster axis (Figure 10).

As it shown in Figure 10, there are large differences among the clusters in respect of both age-structure and economic development. The situation is the best in the second cluster; these micro-regions are located in the agglomeration around Budapest. They are dynamically developing micro-regions with high population density, low unemployment rate, strong economy and favourable demographic characteristics.

Figure 10

Location of the clusters along the factors



Source: Own calculation.

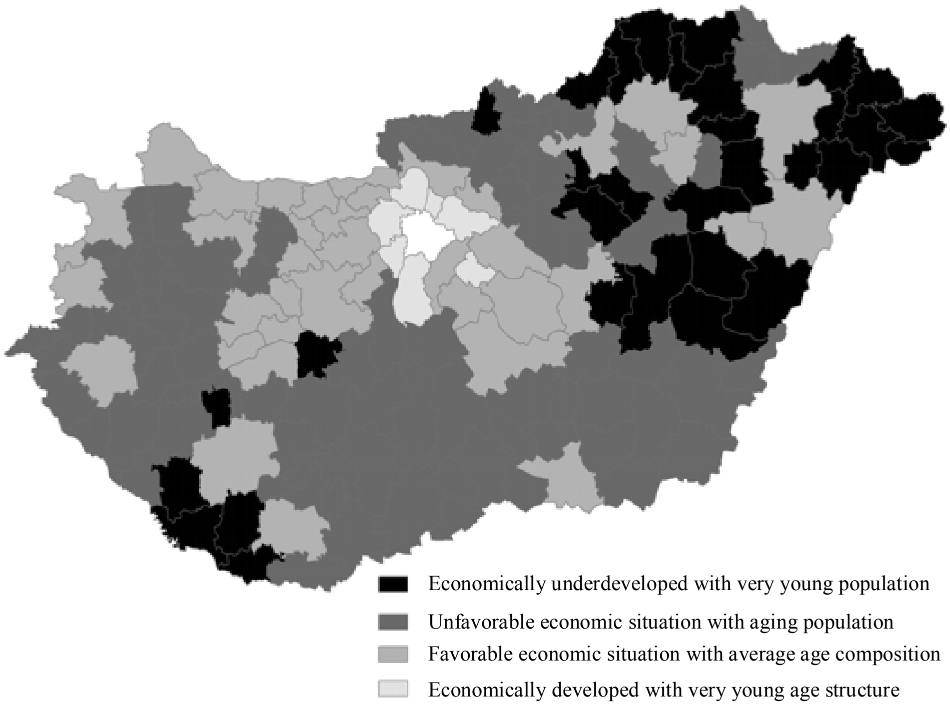
The fourth cluster is the other special group of micro-regions with highly underdeveloped economy and very young age-structure. This group comprises the most problematic micro-regions: there are no job opportunities, the level of educational attainment is low, while the unemployment rate and the proportion of long-term unemployed are high. The birth rate is especially high and the fertility rate is much higher than in the other regions. However, while the fertility and birth rates are outstandingly high, life expectancy at birth is lower than in the other parts of the country (Hablicsek 2009).

The geographical location of the clusters is shown in Figure 11.

Micro-regions with the youngest population are basically located in four regions of the country. The largest contiguous areas are in the north-eastern part of the country, mainly micro-regions in Borsod-Abaúj-Zemplén and Szabolcs-Szatmár-Bereg counties as well as in the agglomeration around the capital city are here. There are two another areas which are characterized by young population: one in Baranya county, along the Dráva and the other in the middle part of Northern Great Plain. However, the favourable age-structure is not always accompanied with economic development. In some micro-regions, the relation between the age-structure and the economic performance is positive, while in others it is negative. Economic development is the strongest in the central part of the country and in the northern part of Transdanubia.

Figure 11

Geographical location of clusters



Source: Own calculation.

Micro-regions in unfavourable economic situation and with young population are located in Southern Transdanubia and Northern and Eastern Hungary.

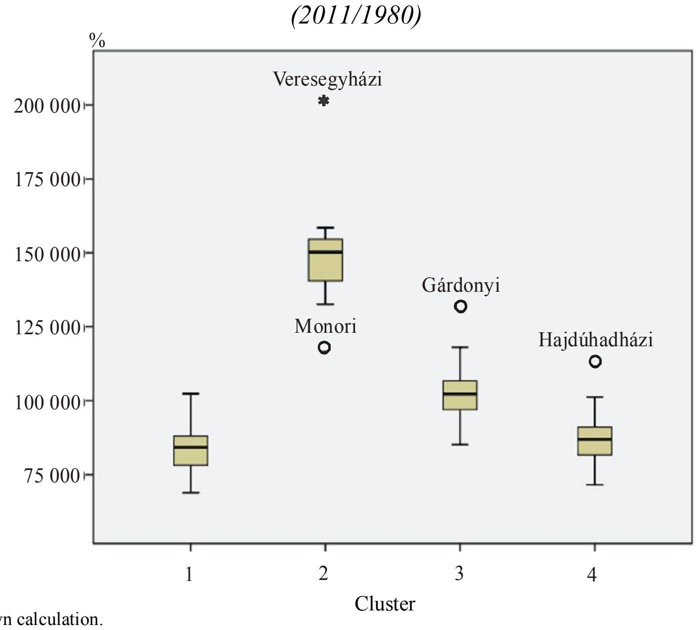
Population dynamics and projection of clusters

The population dynamics of the four established homogeneous groups show considerable differences as well. When examining the period between 1980 and 2011, we can see that, among the four groups, population dynamics is by far the strongest in the economically developed micro-regions around Budapest having favourable socio-economic characteristics and young age-structure. On the whole, considerable population loss is characteristic of the first and the fourth group, with the exception of some micro-regions, such as the Hajdúhadház micro-region in the economically underdeveloped group having however a young age-structure.

The socio-economic situation of the third cluster can be considered good, since its population loss equals the national average or lower, and some micro-regions (e.g. Gárdony) have a considerably positive population balance. It can be seen in Figure 12 that, in respect of population dynamics, the situation is the same in the first and the fourth cluster, while it is significantly better in the third and the second one. The population is decreasing in the first and the fourth cluster, it is stagnant in the third one – having however better value than the country average – while the population of the second cluster comprising only nine micro-regions is dynamically increasing.

Figure 12

Population dynamics in the micro-regions of the clusters (2011/1980)



Source: Own calculation.

Figure 13 shows how the real population dynamics data differ from the projections in each cluster.

The first cluster is the economically underdeveloped group with ageing population. The projection predicted a higher population number than the real number was. This cluster is characterized by a strong population decrease exceeding the projected one.

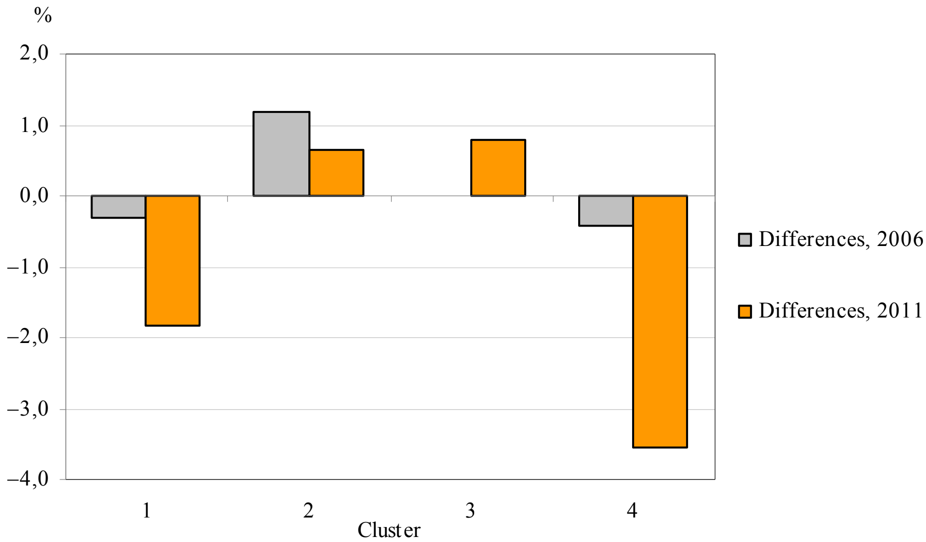
The second cluster includes the most developed micro-regions around Budapest with very young age-structure. In this group, the real population number is higher than the projected value, i.e. the population concentration is stronger than it could be deduced from the projection. The larger in-migration from the country and immigration from abroad, the outstanding fertility rate surpassed the expectations. In 2011, this cluster could record significant population surplus compared to 1980.

The third cluster comprises the moderately developed micro-regions with average age-structure. In 2006, the projection was completely accurate, but by 2011, the population number was higher than projected. The population of the micro-regions comprising the cluster is basically stagnant, so it is characterized by positive demographic trends compared to the national trend.

The population of the fourth cluster is decreasing except for the Hajdúhadház micro-region. The population projection predicted a considerably larger population for this group with very underdeveloped economy but with outstandingly high fertility rate. This also shows indirectly that the population increase moderates among groups with high fertility as well.

Figure 13

Differences between real population number and projected population data in the clusters



Source: Own calculation.

REFERENCES

Bartók János (1983): Faktoranalízis a gyakorlatban. TK Műhely, 2. (Factor analysis in practice)

Bódi Ferenc (2010): Népességfogyás, a szegénység ördögi köre. A Falu, 3. (Population decrease, the vicious circle of poverty)

Brown, David L. – Nina Glasgow (2008): Rural Retirement Migration. Springer

Czirfusz Márton (2010): Faktoranalízis, a látszatmegoldás. Tér és Társadalom, 1. (Factor analysis, a pseudo-solution)

Francia László (1976): A faktoranalízis és alkalmazása területi vizsgálatokban. In: Kulcsár Viktor (szerk.): A regionális elemzések módszerei. Akadémia Kiadó, Budapest (Factor analysis and its application in regional research)

H. Richter Mária (2002): A népesség becslése évközi időpontokra. A kanadai népesség-továbbszámítási modell adaptációja. Demográfia, 2–3. (Population projection for reference dates. An adaptation of Canadian population estimation model)

Hablicsek László (2000): Kísérlet a roma népesség előreszámítására 2050-ig. In: Horváth Ágota et al. (szerk.): Cigánynak születni. Tanulmányok, dokumentumok. Aktív Társadalom Alapítvány & Új Mandátum Kiadó, Budapest (Experimental calculations for the projection of the Roma population till 2050)

Hablicsek László (2001): A népességreprodukció alakulása a 20–21. században. KSH NKI, Budapest (The reproduction of Hungary's population in the 20^{th}–21^{st} centuries)

Hablicsek László (2006): A népesség várható alakulása a következő évtizedekben, különös tekintettel a területi különbségekre. In: Tausz Katalin (szerk.): A társadalmi kohézió erősítése. ÚMK Budapest (Population prospects in the next decades with special respect to the spatial differences)

Hablicsek László (2007a): Népességünk következő évtizedei – különös tekintettel a területi különbségekre. Demográfia, 4. (Next decades of our population with special respect to the spatial differences)

Hablicsek László (2007b): Kísérleti számítások a roma lakosság területi jellemzőinek alakulására és 2021-ig történő előrebecslésére. Demográfia, 1. (Experimental calculations for the spatial characteristics of the Roma population and its projection till 2021)

Hablicsek László (2009): A népesség szerkezete és jövője. Demográfiai portré. Jelentés a magyar népesség helyzetéről. KSH NKI, Budapest (Structure and future of the Hungarian population)

http://fogalomtar.eski.hu/index.php/Népesség-előreszámítás

Kertesi Gábor – Ábrahám Árpád (1996): A munkanélküliség regionális egyenlőtlenségei Magyarországon – 1990 és 1995 között. Közgazdasági Szemle, július–augusztus (Regional inequalities of unemployment in Hungary between 1990 and 1995)

Kulcsár László (2009): Társadalmi zárványok, vagy zárványtársadalmak? Területi Statisztika, 1. (Social enclaves or enclave societies?)

www.nepszamlalas2001.hu

Obádovics Csilla – Popovics Attila (2011): Klaszteranalízis. Melyiket válasszam? „Hitel, Világ, Stádium". Nemzetközi konferencia. 2010. november 3. Sopron, Konferencia kiadvány (CD kiadvány 2011) (Cluster analysis: How to choose?)

Polónyi István – Tímár János (2006): Oktatáspolitika és demográfia. Felsőoktatási Kutatóintézet, Budapest (Educational policy and demographics)

Spéder Zsolt (2002): A szegénység változó arcai. Tények és értelmezések. Századvég (The changing faces of poverty. Facts and meanings)

Teaford, Jon C. (2008): The American Suburb. The Basics. Routledge, New York

Tipold Ferenc – Faluvégi Albert (2007): Tájékoztató a területfejlesztés kedvezményezett térségei besorolásánál alkalmazott mutatókról és a számítás módszeréről. A 67/2007. (VI. 28.) OGY határozat és a 2007. évi CVII. törvénnyel módosított 2004. évi CVII. törvény alapján. Budapest (Information on the applied indicators for classification of favoured areas of regional development and the calculation method).

REGIONAL STATISTICS, 2013, VOL. 3: 79–97 DOI: 10.15196/RS03105

GÉZA TÓTH[a] – LÓRÁNT DÁVID[b] – LÁSZLÓ VASA[c]

Transportation and accessibility at European level

Abstract

This study aims at investigating the multiple and complex relationship between transport and tourism by various methods. Features of the relationship between the transport distance and tourism intensity will be detected and the connections between the accessibility of European regions and their tourism will be analysed.

Keywords: tourism flows, accessibility, modelling

Introduction

The relationship between tourism and transport has been in the focal point of studies for a long period of time (Hall 2010). However, relevant studies often discuss the possibilities of the disabled in tourism exclusively (Carda–Colea–Humphreya 2006, Apec 2003, European Communities 2004).

The matter covers questions whether such relationship between the two activities can be measured at all along with additional ones regarding the closeness and strength of relationship if it exists. It is also relevant to study if such relationship is observable in general or with detectable significant spatial disparities which is the particular subject this paper intends to investigate.

According to our initial hypothesis, although tourism performance is greatly impacted by the level of services provided by transport, related spatial disparities also play a significant role in forming it.

Objectives

Prior to its launch, the research intended to focus on performing a general investigation on the relationship between transport distance and tourism on the example of the European regions as well as to study the role of distance and accessibility as relevant to the topic in order to point out the significance and spatial aspects of this topic.

First of all, it is important to claim that we intend to study the relationship between tourism and transport in general along with an approach by which the spatial movements of individuals including tourists become more apprehensible. Agreeing with Hall's

a Hungarian Central Statistical Office, 1024 Budapest, Keleti Károly út 5-7., Hungary, E-mail: geza.toth@ksh.hu

b Károly Róbert College, 3200 Gyöngyös, Mátrai u. 36., Hungary, E-mail: davidlo@karolyrobert.hu

c Szent István University, 2103 Gödöllő Páter Károly út 1, Hungary, E-mail: Vasa.Laszlo@gtk.szie.hu

opinion we think that "there is a lack of vertical theoretical integration within broader perspectives of mobility" (Hall 2008, p. 15).

Obviously, we also acknowledge the demand for studies of this kind. However, we raise the question whether it is possible to develop a spatial model for tourism movements. Regarding this, Hall (2005) argues that if the physics analogue is maintained, the spatial form of human mobility can be measured by applying the classic Newtonian physics and within that, the number of travels can be described and forecasted assuredly on the macro-level whereas on the micro-level, individual human behaviour can be likened to that analysed by quantum physics in which a higher uncertainty is observed for individuals' travels. Nevertheless, a relationship between physical paradigms and community tourism attitudes exists (Hall–Page 2009). Before launching the study, the interrelationship between transport and tourism is required to be clarified to which a review on the most relevant approaches is given below.

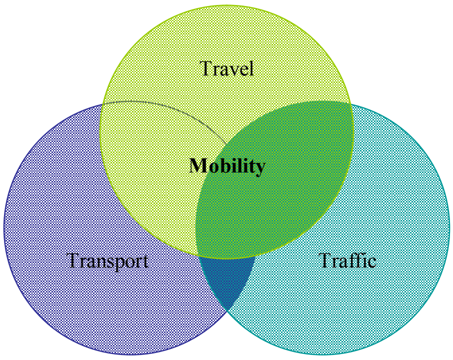
Literature review

The role of transport is manifested in connecting tourism demand and supply and in the internal features of supply, i.e. the destination to be accessed. Transport is one of the primary preconditions to the existence of tourism. It is a key element that links tourists to destinations to be accessed. Though the connection between tourism and transport has been widely examined previously (Page 2005, Prideaux 2000) there are still significant gaps in this research topic (Chew 1987, Gunn 1994, Hall 1991, Inskeep 1991, Page 1994 and 1999; Robbins–Thompson, 2007). As pointed out by Knowles (1993), in many cases researchers took transport into account as a passive element in tourism not as an integral part of tourism activities. Though the tourism product to be consumed by tourists, i.e. the set of services (accommodation, catering, entertainment and other services) based on attractions accordant with the motivation of tourists also includes transport.

It is important to emphasize that in the case of tourism, the demand for motion or mobility is among mankind's fundamental needs (Figure 1). Demand for travel is the extent of travel requisites defined or influenced by travel fees (Jászberényi-Pálfalvi 2009). In the tourism sector, three important definitions are used collaterally to describe related activities. Traffic is a change of locations usually accomplished by suitable technical appliances. For transport, the owners of vehicles expedite people or subjects by their own appliances. Travel is an activity when an individual goes somewhere. In this sense, it is a synonym for transport but it has various meanings i.e. the movement of people or subjects (Jászberényi–Pálfalvi 2009).

Figure 1

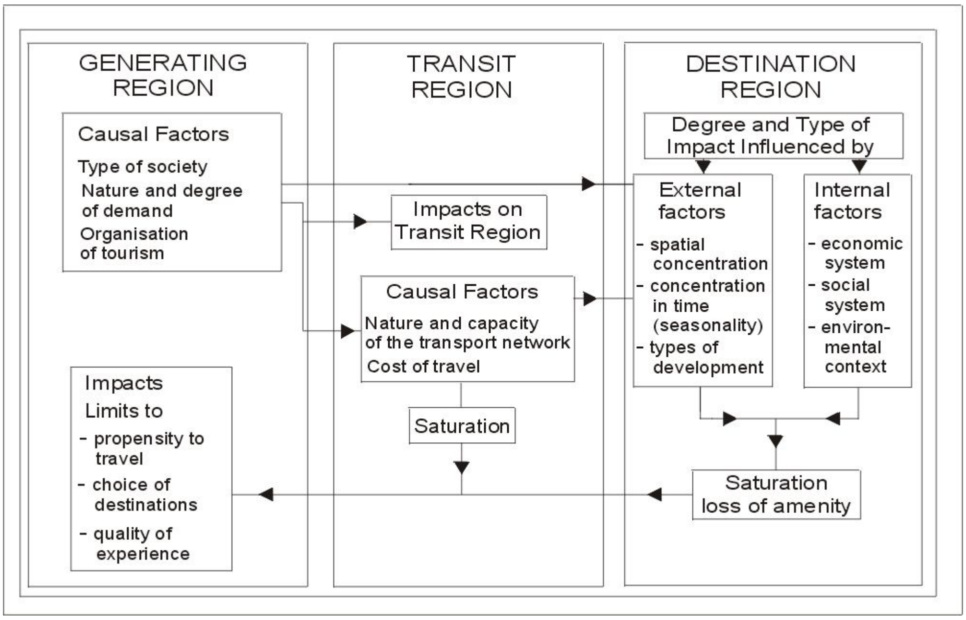
Components of mobility



During travels, travellers get from generating regions through transit regions to destination regions. Generating, transit and destination regions were distinguished by Pearce (1989) after Thurot (1980) while studying the impacts of tourism (Figure 2). For transit regions, the character and capacity of transport networks were studied with their limitations pointed out. When determining the carrying capacity (sporadic saturation) of regions (either generating, transit or destination ones), a specific causal significance was attributed to the features of infrastructure regarding the type and capacity of transport networks as well as to travel costs.

Figure 2

The relationship of causal factors and impacts with the generating, transit and destination regions



Source: After Pearce, D. 1989 and Thurot, J. M. 1980.

At the interface of transport and tourism, Hall and Page (1999) identified four fields to be studied: the link between the source market and host destination, mobility provision and access within the destination, mobility provision and access within an area with a relevant tourist attraction and the advancement of journeys along a recreation route itself also representing a tourism experience. Our study, due to its aspect, can be classified into the first field.

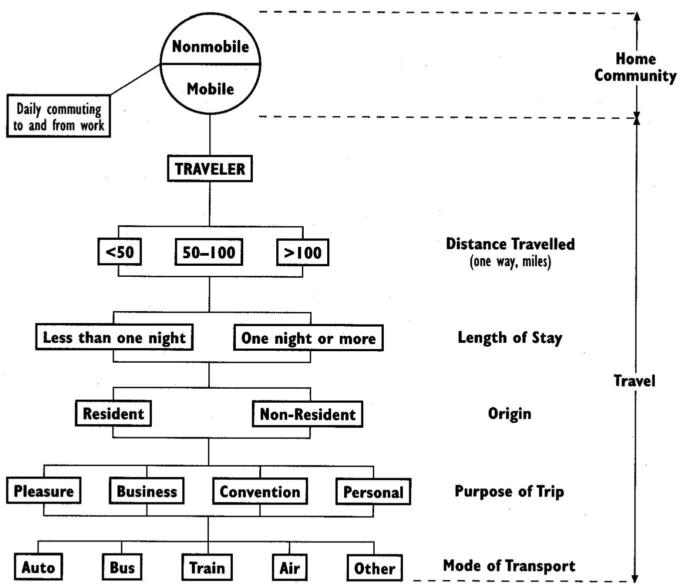
As presented by Hall (1999) relevant to this specific field, linking the generation and destination regions is not primarily accomplished for tourism purposes. As also pointed out, a rather limited number of research was conducted on how the opinions of travellers on a given destination to be accessed are influenced by the time, cost and type of travel required.

One of the methodologically most complicated issues of studying the connection between transport and tourism is how to separate tourism movements from transport capacities. There are several branches of transport that are used both by residents and tourists by choice, thus distinguishing such functions seems to be rather complicated.

The fundamental relationship between leisure and transport was defined by Halsall (1992). According to the author, transport is an essential part of tourist (recreational) behaviour and, additionally, it advances the achievement of recreation objectives while representing a recreational activity itself. The continuous reduction in relative travel costs and distances dramatically increases demands for recreational travel. The increasing use of cars for tourism has especially increased real travel distances and their share within travels for tourism purposes.

Figure 3

The dimensions of travel



Source: Gee, C. Y. – Makens, J. C. – Choy, D. J. L. 1997.

Travel is certainly a rather complex process influenced by a number of other factors. From this approach, the various dimensions of the travel process are demonstrated by the figure below (Figure 3). This present research neither goes into details regarding the above dimensions, nor does it intend to give a detailed analysis on the figure below.

Tourism and accessibility

From our point of view, a study into the relationship between transport and tourism can only be sufficient if adequately general. Not only should the level of services provided by transport infrastructure be taken into account but also features of those potentially participating in the travels, the criteria of travel etc. are also to be studied. To this, accessibility should be considered in our aspect. However, due to the multiplicity of literature it is also important to define accessibility in this respect.

Definitions for accessibility were often developed as an establishment of a spatial model or calculation. Thus a more detailed analysis into the topic is thought to be more expedient pointing out the wide range of compounds the relationship between accessibility and tourism is dependent on. Such approach is reflected by the definition according to which accessibility can be regarded as the sustainability potential of the built environment and the dimension of mankind's quality of life thus it is basically an approach of how the relative importance of certain spatial points are judged (Makri 2001).

This concept of accessibility includes several aspects as listed below:

1. Physical accessibility – whether a given point can be accessed despite existing physical obstacles letting tourists reach the required destination;
2. Mental accessibility – whether the given individual can apprehend and use the possibilities provided by a destination;
3. Social accessibility – whether, due to their social status, individuals have information and demand to travel;
4. Organisational accessibility – whether individuals have access to possibilities, information and services;
5. Financial accessibility – whether the use of individual or community transportation methods is affordable to reach the destination required.

The above definition of accessibility was found appropriate to be applied in our study as its elements are linked either directly or indirectly to the causal factors and impacts indicated by Figure 2 influencing tourist activities as well as the extent and character of flows between the generating and destination regions.

The content behind the definition of accessibility can certainly be modelled in various ways. However, in our study several examples were applied for quantification, regarding the basic definition, the same frame of contexture was studied.

Methods

The research included 3 methods applied. Being different in type, the three methodologies still have a rather similar motive of application. First, the spatial interaction model of concurrent destinations was used in order to model the amplitude of

European tourism flows. Calculated results of the model (with various distance approaches) were compared to the actual data, particularly guest nights, thus a general view on the existing relationship between transport distance and tourism was obtained.

By applying the second method namely the shift-share analysis in addition to the general results obtained by the above study, the relationship between the accessibility of European groups of countries and the number of guest nights was analysed. Two approaches of this method were applied with accessibility compared to changes in the number of guests in the first whereas to the spatial disparities of guest nights per bed in the second.

Finally, in order to quantify relationship between the number of international and domestic guest nights and accessibility, the loglinear analysis was applied. A more detailed description on the above methodologies will be provided below.

A research into the European tourism flows

Transport distance as one of the substantial indicators of travels is only one among the selection criteria of destinations. Regarding distance, a different overall view is drawn for movements for leisure purposes when compared to all dislocations. According to Bull (1994), the travel intensities with shortening distances will increase to a certain position followed by a decline and finally, a zero travel intensity observed at zero distance. At the bottom of this is the fact that too nearby thus too quickly accessible destinations are not attractive for visitors as are considered to be part of their everyday milieu.

It can also be concluded that several tourist destinations indicate a rather intensive development despite their locations relatively distant to their competitors. In many cases, poor accessibility can be practically balanced by other factors of attraction as well as a destination is conceivable where attraction is represented by unfavourable accessibility and the resultant wild destination.

As revealed by certain studies, accessibility has a role primarily in selecting tourist destinations (Thompson–Schofield 2007). Tourism in easily accessible towns indicates intensive development as opposed to those hard-to-access stagnates. According to a hypothesis, tourists during their travel decisions select the destinations to be reached first based on the local possibilities and attractions (Crompton 1992). In this decision-making, destinations sufficing the purposes of visitors and with similar type of endowments are taken into account (Celata 2007). Only after this primary selection is made will destinations be compared by accessibility. Thus accessibility primarily has or can theoretically have a role in substituting potentially visitable destinations. On the contrary, destinations capable of providing comparative advantages for tourists can attract a significant number of visitors even if with relatively unfavourable accessibility. Consequently, the matter of accessibility is relevant for destinations with similar endowments (seaside) whereas is less remarkable for those with individual attractions (historical towns, spas). Favourable accessibility itself does not necessarily represent an origin of competitiveness.

When studying tourism flows, the question of how individuals at a given point of origin will select among the accessible travel destinations can be raised. According to Haynes and Fortheringham (1990), spatial choice process has three characteristics. First of all, this is a discrete process. That is a given travel destination is either selected or not and there is a finite number of travel alternatives. Second, the number of alternatives is often extremely large. Third, the alternative destinations all have fixed spatial locations which limit the degree to which alternatives are substitutes for one another (Fotheringham et al. 2000). As a result of the above criteria, Fortheringham (1983, 1984, 1991) developed a spatial interaction model of competing destinations that is basically a single limited accessibility model by which first we intended to analyse the relationship between accessibility and tourism. Accordingly:

$$I_{ij}=(O_i; S_j; D_{ij}; A_j)$$

where I_{ij} is the interaction between the i^{th} origin and the j^{th} destination, O_i is the i^{th} place's ability as an origin to contribute to the interaction, S_j is the attractiveness of j as a destination, D_{ij} is the intervening distance between the origin and destination and A_j is the competing destinations variable being the accessibility of j^{th} destination relative to all others that may interact with the i^{th} origin, i.e.:

$$A_j = \sum_{\substack{k=1 \\ k\neq j}}^{m} S_i D_{ij}$$

Hereafter, the research attempts to focus on the type of relationship explored between theoretical accessibility calculated for tourism and statistical data on the number of visitors. Can the number of guest nights at hotel type units be estimated by applying the spatial interaction model? Within this, a further question concerns the role of transport, i.e. the distance in the estimation and whether Bull's (1994) theory can be validated by data.

The starting point of our study, for practical reasons, was the European Union's NUTS system as it is ensured that data are available at comparable regional levels. Moreover, the NUTS system was introduced not only in EU-27 member states but also in the EFTA countries, consequently data for the regions of Switzerland, Norway and Iceland were also taken into account in the analysis. For the calculations, NUTS2 data were used. The number of regions considered was 280 with transcontinental areas of France, Spain and Portugal excluded.

In the model, the universal accessibility definition was applied, i.e. given regions were not analysed by their main generating regions and how the appropriate destinations are accessed. In other words, theoretically, travels can be made from any region to any other one with tourist motivation. (Obviously, in practice, this is not the case; however, due to the features of modelling, it was put aside.)

The participatory capacity of departure regions in the interaction was attempted to be quantified by their population data.[1] The total number of population was included in the model despite being unlikely in real life that everyone would participate in tourism flows (see the definition of accessibility). Significant variations can be observed in implementing tourism activity based on age, sex, financial status, marital status and other considerations. By including the total population as potential tourists in the model, a

1 Data source was in this case as for other statistical data alike was the Eurostat database (http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/).

theoretical maximum was defined to which the amplitude and spatial distribution of real flows become well-measurable.

The tourist attraction of destinations was represented by the number of beds in hotel type units in the given region. We claim that by this the bulk of attraction is indicated directly as, not regardless to the general level of economic development and processes, the greater the attraction, the more beds there are at quarters.

Distance between the origin and destination regions was specified by the distance between the regional centres on road measured in minutes. In this study, when preparing the road network database, accessibility time in minutes for all road segments (for crossings-to-crossing sections) was defined with speeds appropriate to each road category. Within the networks, by programming GIS software, time demand of the optimal routes requiring the minimal access time was also established for the study area. This procedure corresponds to defining the optimal access path between two points of a graph where edges are the road segments and resistance data for the edges are the time data required for passing. Applying road distance data is apparently the first approach only as being otherwise obvious, other transport sub-sectors also play a relevant role in tourism-induced travels in the study regions. In our aspect, the time-space being significantly different from the geographical space and modelled by applying road distances (Dusek–Szalkai 2008) could be a good starting point in this present analysis.

The number of guests in a given region, i.e. in this particular case, the number of guest nights can be calculated as the sum of incoming tourism flows.

The spatial interaction model is based on a gravitational analogy as field intensity here is also studied in the relation of masses and distances. Our research intends to focus on the second one i.e. the topic of accessibility. To estimate the role of distance sensitivity in tourism flows, an analysis on the value γ constant can bind in the gravity model was relevant. Therefore calculations were performed by constants within the range between 0 and 2 followed by studying the strength of correlation between the calculated and actual values.

As concluded by Dusek (2003) in his work on the gravity model: "With the exponent increasing, the intensity of interregional connections becomes more distance sensible and collaterally the relevance of masses will gradually decline."

Table 1

Weighted means of correlation coefficients for various γ constants of the gravity model

γ	0,0	0,5	1,0	1,5	2,0

Source: Own calculation.

As seen above (Table 1), by applying the spatial interaction model, data on the number of guests can be sufficiently estimated, as coefficients of determination are relatively high, however, they can not be regarded as distance dependent. This is due to several reasons. On the one hand, the most important tourist destinations are located at the continent's periphery at positions relatively disadvantageous from the point of view of accessibility. On the other hand, tourism product as tourism experience is indefinable, one-shot and perishing i.e. can not be stored. E.g. in case somebody intends to spend the summer holidays at the seaside, such demand will not be replaced by spending it at a nearby although well accessible mountainous area but will undertake travelling to remote peripheries.

A study into accessibility and the number of guests by applying shift-share analysis

The following analysis intends to study the extent of the number of guests in European regions explained by accessibility and other local reasons. To this, the shift-share analysis was applied. Description on the method has been given in several spatial statistical publications and volumes (Houston 1967, Curtis 1972, Berzeg 1978, Stevens–Craig 1980) and an example for its application regarding accessibility in Hungary was provided by Tóth (2002).

Here, thus, a different approach was attempted. As already indicated earlier, accessibility has or can have a role primarily in substituting potentially visitable destinations (Celata 2007). Obviously, the question can be raised whether this is the case for all destination groups.

Therefore destinations with similar features were intended to be studied from the aspect of accessibility. European regions were classified into five groups based on the location of the countries involved. Our hypothesis in this respect was that for the contiguous groups of countries, several differences can be observed regarding the type and strength of relationship between accessibility and the number of guests.

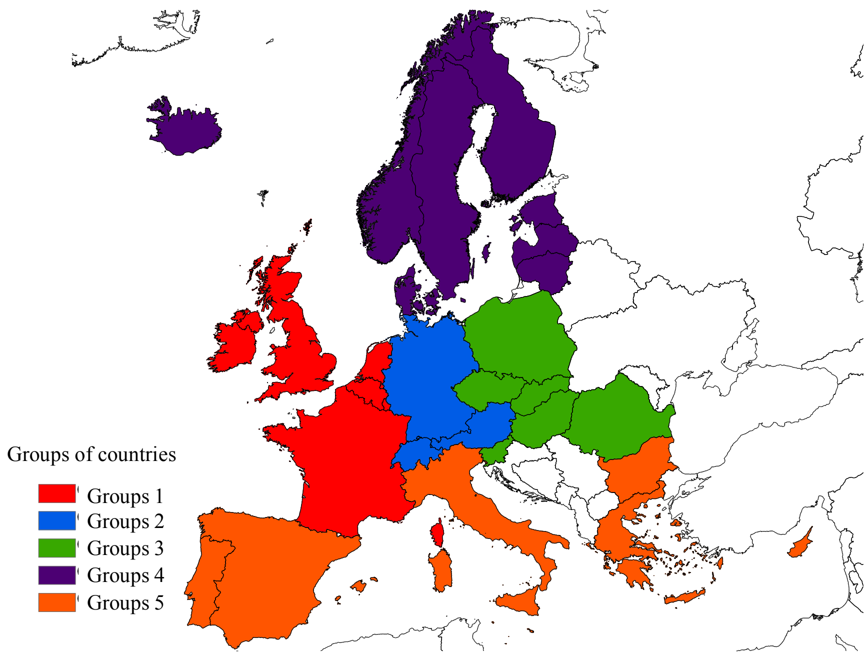
The groups and the countries included are as follows:

1. Western Europe: Belgium, the Netherlands, Luxemburg, France, United Kingdom, Ireland;
2. West Central Europe: Germany, Switzerland, Austria, Liechtenstein;
3. East Central Europe: Poland, the Czech Republic, Slovakia, Hungary, Romania, Slovenia;
4. Northern Europe: Denmark, Finland, Sweden, Norway, Iceland, Estonia, Latvia, Lithuania;
5. Southern Europe: Portugal, Spain, Italy, Greece, Bulgaria, Malta, Cyprus.
6. The groups and the countries included can be seen on Figure 1.

The groups and the countries included can be seen on Figure 1.

Figure 1

The groups of European countries



Accessibility in this respect was studied not only based on road transport data but also multimodal accessibility i.e. based on the use of various modes of conveyance and taking them into account collectively was attempted to be applied.

To this, data available on the Espon website2 were used. Downloadable data, among others, included multimodal accessibility of NUTS3 regions in the study area. As our research was intended to be carried out on the NUTS2 level such data was inappropriate thus population-weighted mean was applied.

First, changes in the number of guest nights were analysed between 2003 and 2009. In the second study, the specific method of shift-share analysis with the spatial disparities of guest nights per bed in 2009 factorised was applied. We intended to explore the amplitude accessibility and other local factors which are responsible for spatial disparities. (It is not possible to define the influencing specific local factors by the analysis, only the extent changes in the number of guests deviating from the European average is influenced by accessibility (in other words, the extent positive or negative deviation or in short surplus or deficiency in the number of guests compared to the average in the number of guest nights is entailed), and other factors characteristic for the given region (including: the level of urbanisation, seaside or mountain location etc.)

2 http://www.espon.eu/main/Menu_Publications/Menu_TerritorialObservations/trendsinaccessibility.html

Table 2

Surplus/deficiency in the number of guests and its components (2003/2009)

(%)

Regions	Total	Spatial	Accessibility
Western Europe	100	–69	169
West Central Europe	100	–6 703	6 803
East Central Europe	–100	136	–236
Northern Europe	–100	–56	–44
Southern Europe	100	226	–126

Source: Own calculation.

As indicated by the data in Tables 2 and 3, accessibility plays a more important role in the changes in the number of guests than spatial dimensions i.e. other local conditions for 3 of the 5 groups of countries as having higher absolute values. It is due to their accessibility position that Western and West Central Europe have more advantageous trends whereas countries in East Central Europe show slower dynamics compared to the European average also primarily due to their accessibility. Disadvantageous accessibility further spoils disadvantageous local conditions in Northern Europe while regarding the countries in Southern Europe accessibility can slightly worsen favourable local endowments. Accessibility of Southern Europe can not be disadvantageous within the continent to impede the increase in the number of guests exceeding the European average. Table 3 indicates the components of changes in the number of guests between 2003 and 2009.

Regarding the factorisation of data of guest nights per bed in 2009, a somewhat different overall view is seen (Tables 4 and 5). In this respect, one can observe a more important role of accessibility only for the West Central European countries compared to local conditions for the number of guest nights per bed. For the other groups of countries, however, it can be seen that conditions basically determined by local endowments can only be modified either in a positive or negative way by accessibility.

Summarised, while accessibility plays a significant role in the changes in the number of guests in yet more groups of countries, its role in effectiveness (capacity exploitage) is not relevant.

Table 3

The share of regions in surplus/deficiency in the number of guests and its components (2003/2009)

(%)

Regions	Surplus in the number of guests	Deficiency in the number of guests	The impact of territoriality is positive	The impact of territoriality is negative	The impact of accessibility is positive	The impact of accessibility is negative
Western Europe	27.5	–	–	11.3	32.0	–
West Central Europe	1.5	–	–	57.6	68.0	–
East Central Europe	–	6.6	5.3	–	–	10.7
Northern Europe	–	93.4	–	31.1	–	28.0
Southern Europe	71.0	–	94.7	–	–	61.3
Total	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*

Source: Own calculation.

Table 4

Surplus/deficiency in the number of guests and its components, 2009

(%)

Regions	Total	Spatial	Accessibility
Western Europe	100	83	17
West Central Europe	100	–8 804	8 904
East Central Europe	–100	–71	–29
Northern Europe	100	216	–116
Southern Europe	–100	–74	–26

Source: Own calculation.

Table 5

The share of regions in surplus/deficiency in the number of guests and its components, 2009

(%)

Regions	Surplus in the number of guests	Deficiency in the number of guests	The impact of territoriality is positive	The impact of territoriality is negative	The impact of accessibility is positive	The impact of accessibility is negative
Western Europe	86.2	–	71.1	–	33.6	–
West Central Europe	0.3	–	–	28.0	66.4	–
East Central Europe	–	40.6	–	28.6	–	27.3
Northern Europe	13.5	–	28.9	–	–	36.4
Southern Europe	–	59.4	–	43.4	–	36.2
Total	100.0	100.0	100.0	100.0	100.0	100.0

Source: Own calculation.

The loglinear theoretical model

Where there is no destination due to the lack of tourist attractions none will be established even after having potential infrastructural development implemented. In turn, among destinations with similarly advantageous endowments for tourism, theoretically a higher flow tends towards regions with better accessibility conditions.

This final research primarily intends to examine the degree guest nights of a destination depend upon accessibility and local conditions in cases with existing tourist attraction found.

The hypothesis, in this present case, also is that a relationship between accessibility and the number of guest nights is assumed. In our study, the extent and components of this will be analysed by applying the loglinear model.

The loglinear model examines when and in which context our variables are interdependent on each other. The method interprets the connection between variables as follows: in case any of the variables fall into any of the defined categories then this makes the likelihood of such variables falling into another category according to other features high. Such attribution is usually called interaction between the variables (Füstös 1985).

An average contingency table is taken as a starting point (Tables 6, 7 and 8). The regions studied were classified, by using the accessibility data of Espon described above, into four accessibility groups. Thus, data on the number of guests for all groups of countries could be classified based on accessibility into these four groups. As, at present, no region in West Central Europe can be classified into the worst, accessibility Group 4, for methodological reasons the unavailable data had to be replaced. This was achieved by placing the minimum values of the contingency table into the empty cells.

X and Y represent two (random) variables with range carriers with I and J elements, in which:

$p_{ij}=P(X=x_i, Y=y_j)$.

$m_{ij}=n*p_{ij}$–t, in other words, m_{ij} is a general element of the contingency table, and

$\xi_{ij}=\log(m_{ij})$.

Having the logarithms of all elements of the contingency table taken, a **ξ**–table or in another approach a matrix is obtained.

According to the notation generally applied, the calculation of any optional rows or columns of the table and the mean of the whole table can be described by the formulae below:

$$\xi_{i*} = \frac{\sum_j \xi_{ij}}{J}, \quad \xi_{*j} = \frac{\sum_i \xi_{ij}}{I}, \xi_{**} = \frac{\sum_i \sum_j \xi_{ij}}{I \cdot J}$$

Consequently, any element of this new table containing the logarithms of the original table can be noted in the following form:

$$\xi_{ij} = \xi_{**} + (\xi_{i*} - \xi_{**}) + (\xi_{*j} - \xi_{**}) + \left\lfloor \xi_{ij} - (\xi_{i*} - \xi_{**}) - (\xi_{*j} - \xi_{**}) - \xi_{**} \right\rfloor$$

This can be interpreted as (any elements can be generated) a sum of the total average, the average analogous to the given row, to the given column and the appropriate row–column interaction.

Where mij is the actual number of cases in cells i–j, I is the group of countries (i=5), J is the accessibility category (j=4), ξi* is the logarithm of the number of total, domestic and international guest nights of the ith group of countries for the various accessibility groups, ξ*j is the logarithm of the total, domestic and international guest nights for jth accessibility group for the groups of countries, and ξij is the likelihood that the observed total, domestic or international guest nights falls into the cells i–j of the table, compared to the probabilities defined by all secondary parameters above.

Having ξij=log(mij) replaced into the relationship above, an additive formula is obtained as below:

$$\log m_{ij} = \mu + \lambda_i^x + \lambda_j^y + \lambda_{ij}^{xy}$$

$$m_{ij} = e^{\mu} \cdot e^{\lambda_i^x} \cdot e^{\lambda_j^y} \cdot e^{\lambda_{ij}^{xy}}$$

where μ marks the total average, i–indexed λ is the effect of row, j–indexed λ is the effect of column whereas the ij index marks the interaction.

The advantages of this resolution are demonstrated by the following formulae; i.e. a resolution was applied where the effects of row and column and the interactions sum zero or, in other words, their impact on the whole table is zero:

$$\sum_i \lambda_i^x = \sum_j \lambda_j^y = 0 \quad (\prod_i e^{\lambda_i^x} = 1)$$

$$\sum_i \lambda_{ij}^{xy} = \sum_j \lambda_{ij}^{xy} = 0$$

Table 6

Distribution of guest nights at hotel type units in 2009

(%)

Accessibility groups	Group of countries					Total
	Western Europe	West Central Europe	East Central Europe	Northern Europe	Southern Europe	
1	8.6	9.5	0.9	0.5	5.3	24.9
2	4.5	9.8	0.8	0.6	18.0	33.7
3	4.2	1.1	3.2	2.4	16.4	27.3
4	1.1	–	1.7	2.5	8.8	14.1
Total	*18.4*	*20.5*	*6.6*	*6.0*	*48.5*	*100.0*

Source: Own calculation.

Table 7

Domestic guest nights in hotel type units in 2009

(%)

Accessibility groups	Group of countries					Total
	Western Europe	West Central Europe	East Central Europe	Northern Europe	Southern Europe	
1	7.4	11.1	0.3	0.6	4.3	23.6
2	5.8	10.4	0.6	0.8	16.3	33.9
3	5.2	1.8	4.0	2.8	15.3	29.1
4	1.3	–	2.4	2.7	7.0	13.4
Total	*19.6*	*23.3*	*7.3*	*6.8*	*42.9*	*100.0*

Source: Own calculation.

Table 8

International guest nights in hotel type units in 2009

(%)

Accessibility groups	Group of countries					Total
	Western Europe	West Central Europe	East Central Europe	Northern Europe	Southern Europe	
1	10.5	7.1	1.9	0.5	7.0	27.1
2	2.6	8.8	1.0	0.3	20.6	33.4
3	2.6	0.1	2.0	1.7	18.2	24.5
4	0.7	–	0.7	2.1	11.6	15.1
Total	*16.4*	*16.0*	*5.6*	*4.7*	*57.4*	*100.0*

Source: Own calculation.

By applying the loglinear model, three cases (i.e. the spatial distribution of total, domestic and international guest nights) were analysed. Our null–hypothesis was that our data were independent, in other words, there was no interaction between the two variables in the three cases. According to this hypothesis, the saturated model (containing all interactions, i.e. in this case the accessibility–region interaction) and the model without interaction will fit each other.

Results of the loglinear model

The threshold $\chi 2$ value (to the level of 95%) is 5.99, however, our data indicated values much higher thus values derived by neglecting interactions do not fit the original table of convergence and the null–hypothesis is rejected. In other words, the groups of countries and the accessibility groups, when compared with the total, domestic and international guest nights, are not independent of each other, and the actual data cannot be explained by the exception of interaction between the two variables.

Hereafter, we aimed to explore the provable as well as the quantifiable impacts of country groups and accessibility on the total, domestic, and international guest nights at hotel type units in the regions in the groups of countries.

Our study was conducted for the year 2009. The value of e^{μ} in the table according to total guest nights is 62,626,762, whereas that of domestic ones is 37,964,527 and for international ones it is 17,376025. The following table (Table 9) contains the power of e of the appropriate interactions. Consequently, by the following e^{μ} values and the table calculated, basic data for the previous tables can be generated (Tables 6,7 and 8).

For example, the total number of guest nights of regions for accessibility Group 1 of Western Europe can be obtained when the above value of 62,626,762 is multiplied by, taken from the following table, the impact of Western Europe (1.29), the impact of the first group (1.04), and the interaction between these two (2.25). Here, the result is 190,127,082 (being the basic data of the first row of the adequate table). All other cell values were obtained in a similar way.

The results themselves can provide information on how variables or interactions between variables can influence guest nights. Values exceeding 1 increase incomes whereas those less than 1 reduce them.

Based on the above, it can be concluded that no absolute interaction exists between the increase of accessibility and that of guest nights. Although areas with the most favourable accessibility (Accessibility Group 1) also indicate the highest value of interaction for international guest nights, the most favourable value regarding domestic guest nights can be observed in Group 3. Taken as a whole, no absolute connection can be found between the regions' accessibilities and the number of guest nights.

There is a significant difference between domestic and international guest nights in respect to how high the multiplier is in locations with the best accessibility. As with regions with the best accessibility, the multiplier for international incomes is somewhat higher than that for domestic ones, so it can be concluded that international guest nights are much more influenced by locations with favourable accessibility than domestic ones.

Interactions on the level of country groups basically reflect the spatial differences between incomes. The significant differences of interaction between domestic and international guest nights present in the given groups of countries are, however, also worth mentioning. They are derived primarily from the rather high spatial concentration of international guest nights compared to the somewhat more even distribution of domestic ones.

These variables have their impact on the incomes not only independently, but in interaction with each other as well. Now, it becomes apparent from the tables that the multiplier for the total guest nights among the regions impacted by Group 1 is the highest mainly in the West Central Europe – the regions with the most significant incomes are indicated in brackets – (Oberbayern, Berlin, Darmstadt), whereas for accessibility Group 2, the multiplier is the most significant also in this group of countries (Tirol, Schleswig–Holstein, Salzburg). The positive impact of Group 3 is represented mainly in East Central Europe (Zachodniopomorskie, Małopolskie, Severovýchod), while that of Group 4 is observable in Northern Europe (Norra Mellansverige, Vestlandet, Sør–Østlandet).

When the components i.e. the domestic and international guest nights are examined a difference is seen although for all groups the multiplier is the highest in the group of countries for both the domestic and international ones where it was seen in the 'total' category, however with a various extent. In three of the four accessibility groups, the multiplier for domestic ones is somewhat lower compared to that seen for international ones. In Group 2, the multiplier for international guest nights is more than double of the domestic ones. This refers to the rather great spatial and accessibility concentration of international guest nights in Europe.

Conclusions

Based on the research carried out it was concluded that the spatial interaction model is adequately suitable to estimate data on the number of guests, they cannot be seen as distance dependent. With the results recognised we argued the estimates on the relationship between the intensity of travels (that was modelled by the number of guest nights) and distance acknowledged (Bull 1994).

The results of the shift–share analysis carried out indicated accessibility playing a more important role than spatial dimension for 3 of the 5 groups of countries. Regarding the data of guest nights per bed in 2009, a more important role of accessibility is observed exclusively for West Central European countries compared to local conditions. In other words, conditions determined by basically local endowments can only be modified by accessibility. Therefore, while accessibility plays a significant role in the changes regarding the number of guests still in more groups of countries, its role in effectiveness (capacity exploitage) is not relevant. Our relevant hypothesis, according to which regarding the relationship between accessibility and tourism, a significant difference exist was proved true.

Table 9

Results of the loglinear analyis

Parameter		Total	Domestic	International
		No. Of guest nights		
Accessibility	Group 1	1.04	0.82	1.73
	Group 2	1.15	1.21	1.37
	Group 3	1.27	1.57	0.84
	Group 4	0.66	0.65	0.50
Group of countries	Western Europe	1.29	1.48	1.32
	West Central Europe	0.98	1.03	0.42
	East Central Europe	0.50	0.41	0.62
	Northern Europe	0.42	0.48	0.44
	Southern Europe	3.83	3.33	6.52
Region–Accessibility	Group 1 – a Western Europe	2.25	2.18	2.26
	Group 1 – West Central Europe	3.31	4.72	4.75
	Group 1 – East Central Europe	0.64	0.34	0.87
	Group 1 – Northern Europe	0.44	0.50	0.35
	Group 1 – Southern Europe	0.47	0.56	0.31
	Group 2 – Western Europe	1.08	1.15	0.72
	Group 2 – West Central Europe	3.09	3.01	7.50
	Group 2 – East Central Europe	0.47	0.42	0.60
	Group 2 – Northern Europe	0.44	0.48	0.27
	Group 2 – Southern Europe	1.44	1.46	1.14
	Group 3 – Western Europe	0.90	0.80	1.13
	Group 3 – West Central Europe	0.33	0.40	0.13
	Group 3 – East Central Europe	1.80	2.21	1.86
	Group 3 – Northern Europe	1.59	1.34	2.24
	Group 3 – Southern Europe	1.19	1.05	1.64
	Group 4 – Western Europe	0.46	0.50	0.54
	Group 4 – West Central Europe	0.30	0.17	0.22
	Group 4 – East Central Europe	1.85	3.17	1.03
	Group 4 – Northern Europe	3.20	3.12	4.72
	Group 4 – Southern Europe	1.23	1.17	1.74
	Group 5 – Western Europe	2.25	2.18	2.26
	Group 5 – West Central Europe	3.31	4.72	4.75
	Group 5 – East Central Europe	0.64	0.34	0.87
	Group 5 – Northern Europe	0.44	0.50	0.35
	Group 5 – Southern Europe	0.47	0.56	0.31

Source: Own calculation.

As indicated by the loglinear model, no definite connection is observed between improving accessibility and the increase of guest nights. The number of international guest nights is more sensible to more advantageous accessibility compared to domestic ones. Significant interrelationship differences are observed between international and domestic guest nights which can be primarily derived from the spatial concentration of international guest nights compared to which domestic ones indicate a more even distribution. This refers to the rather great spatial and accessibility concentration of international guest nights in Europe.

The estimate conceptualised in the null–hypotheses, according to which transport impacts tourism productivity, was proved true. Although the type and strength of relationship between tourism productivity and the level of services provided by transport can vary when different approaches are applied, we claim that the matter is worth paying attention to in the field of tourism planning.

REFERENCES

Apec tourism working group (2003): *Best practices in tourism accessibility for travellers with restricted physical ability* Final report, Singapore

http://www.tc.gov.yk.ca/pdf/accessibility_best_practices_in_tourism.pdf p. 88.

Berzeg, K. (1978): The Empirical Content of Shift–Share Analysis *Journal of Regional Science* 18 (3): 463–468.

Bull, A. (1994): *The Economics of Travel and Tourism* Pitman Publishing Melbourne

Carda, J. A.– Colea, S. T.–Humphreya, A. H. (2006): A comparison of the Accessibility and Attitudinal Barriers Model: Travel providers and travellers with physical disabilities *Asia Pacific Journal of Tourism Research* 11 (2): 161–175.

Celata F. (2007): Geographic marginality, transport accessibility and tourism development In: Celant A. (ed.) *Global Tourism and Regional Competitiveness* pp. 37–46. Patron, Bologna.

Chew, J. (1987): Transport and tourism in the year 2000 *Tourism management* 8 (2): 83–85.

Crompton, J. L. (1992): Structure of vacation destination choice sets *Annals of Tourism Research* 19 (3): 420–434.

Curtis, W.C. (1972): Shift–share analysis as a technique in rural development research *American Journal of Agricultural Economics* 54 (2): 267–270.

Dusek, T. – Szalkai, G. (2008): *Differences between geographical space and time spaces: theoretical consequences and Hungarian examples* 48th Congress of the European Regional Science Association, Liverpool.

Dusek, T. (2003) A gravitációs modell és a gravitációs törvény összehasonlítása (Comparing the Gravity Model and the Law of Gravitation) *Tér és Társadalom* 17 (1): 41–58.

European Communities (2004): Improving information on accessible tourism for disabled people, Luxembourg http://ec.europa.eu/enterprise/sectors/tourism/files/studies/improving_information_on_accessibility/improving_accessibility_en.pdf p. 33.

Fotheringham, A.S (1984): Spatial flows and spatial patterns *Environment and Planning A* 16 (4): 529–543.

Fotheringham, A.S (1991): Migration and spatial structure: The development of the competing the destinations model In: Stillwell, J. –Congdon, P. (eds) *Migration Models: Macro and Micro Approaches* Belhaven, London.

Fotheringham, A.S. – Brunsdon, C. – Charlton, M. (2000) *Quantitative Geography: Perspectives on Spatial Data Analysis* Sage, London.

Fotheringham, A.S. (1983): A new set of spatial–interaction models: The theory of competing destinations *Environment and Planning A* 15 (1): 15–36.

Füstös, L. (1985) Loglineáris modell kontingencia táblák elemzésére (Loglinear model to analyse contingency tables). MTA Szociológiai Kutató Int. , Budapest.

Gee, C.Y. – Makens, J.C. – Choy, D.J.L. (1997): *The Travel Industry* John Wiley and Sons, New York.

Gunn, C. A. (1994): *Tourism Planning Basic Concepts Cases*. Taylor & Francis, Washington.

Hale, C.M.–Page, S. J. (1999) *Geography of Recreation and Tourism: Place and Space* Routledge, London.

Hall, C. M (2010): Spatial analysis: A critical tool for tourism geographies. In: Wilson, J. *Space Place and Tourism New Perspectives in Tourism Geographies* pp: 1–16. Routledge, London.

Hall, C. M. (1991): *Introduction to Tourism In Australia: Impacts, Planning and Development* Longman Cheshire, Melbourne.

Hall, C.M. – Page, J. (2009): Progress in Tourism Management: From the geography of tourism to geographies of tourism – A review *Tourism Management* 30 (1): 3–16.

Hall, C.M. (2005). Time, space, tourism and social physics *Tourism Recreation Research* 30(1): 93–98.

Hall, C.M. (2008). Of time and space and other things: Laws of tourism and geographies of contemporary mobilities In: Burns, P. –Novelli, M. (Eds.) *Tourism and mobilities* pp. 15–32 CABI, Wallingford.

Hall, D. R. (1999): Conceptualising tourism transport: inequality and externality issues. *Journal of Transport Geography* 7 (3): 181–188.

Halsall, D. (ed.) (1982): *Transport for Recreation* IBG Transport Geography Study Group, Lancaster.

Haynes, K.E. – Fotheringham, A.S. (1990): The impact of space on the application of discrete choice models *Review of Regional Studies* 20 (2): 39–49.

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

Inskeep, E. (1991): *Tourism Planning. An Integrated and Sustainable Approach*. Van Nostrand Reinhold, New York.

Jászberényi, M.–Pálfalvi, J. 2009. Nemzetközi közlekedés és turizmus (International transport and tourism), Bologna – Tankönyvsorozat, Aula Kiadó, Budapest.

Knowles, R. (1993) Research agendas for transport geography in 1990s. *Journal of Transport Geography* 1 (1): 3–11.

Makrí, M.B., (2001): *Accessibility Indices. A Tool for Comprehensive Land–Use Planning* Proceedings of the TLEnet 5th workshop, The Nordic Research Network on Modelling Transport, Land–Use and the Environment, Sweden.

Page, S. (1994): *Transport for Tourism* Routledge, London.

Page, S. (1999): *Transport and Tourism* Longman, New York.

Page, S. (2005): *Transport and Tourism: Global Perspectives* Pearson Education, Harlow.

Pearce, D. (1989). *Tourist development* Longman, New York.

Prideaux, B. (1993): Possible Effects of New Transport Technologies in the Tourist Industry in the 21st Century. *The Australian Transport Research Forum* 18:245–258.

Robbins, D.–Thompson, K. (2007): Guest editorial: Special issue on transport at tourist destinations. *Journal of Transport Geography* 15 (2): 80–82.

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

Thurot, J.M. (1980). *Capacité de charge et production touristique* CHET, Aix–en–Provance.

Thompson, K. – Schofield, P. (2007): An Investigation of the Relationship between Public Transport Performance and Destination Satisfaction *Journal of Transport Geography* 15 (2): 136–144.

Tóth, G. (2002): Kísérlet autópályáink területfejlesztő hatásának bemutatására (An attempt to represent the impact of motorways in regional development in Hungary). *Területi Statisztika* 42 (6) : 493–505.

REGIONAL STATISTICS, 2013, VOL. 3: 98–116 DOI: 10.15196/RS03106

BALÁZS LENGYEL[a,b] – IZABELLA SZAKÁLNÉ KANÓ[c1]

Related variety and regional growth in Hungary: towards a transition economy approach

Abstract

The aim of this paper is to adapt related variety calculations to the special case of Hungarian regional development in the late post-socialist transition period. First, we test regional employment growth in rising and declining regions separately, in order to distinguish those areas that could cope with economic transition and those that could not. We find that related variety speeded up growth in the dynamic regions but at the same time pushed lagging regions to a downhill path that might be due to their unflexible industry structure. Second, regional variety measures are decomposed into domestic and foreign subsets and a new variable, ownership variety is introduced. Findings suggest that regional employment growth is due to related variety in the domestic set in earlier phases whereas the economy has evolved into a stage in which relatedness among foreign firms enhances regional employment growth significantly.

Keywords: related variety, regional employment growth, foreign-owned firms, post-socialist transition

Introduction

Co-location of companies are central concepts in the literature on regional economic growth. However, a fundamental debate has undermined the common understanding of regional employment growth in a region until recently: does regional specialization or diversity favours growth? Nevertheless, co-located companies from similar industries in industrial districts (Marshall 1890) or regional clusters (Porter 2003) are claimed to increase their productivity because they migth obtain very special and detailed knowledge from each other. On the other hand, companies co-located in economically diversified city-regions share and combine knowledge across industries more easily and therefore are more likely to create new variety and incremental innovation (Jacobs 1969; Glaeser *et al.* 1992). Thus, regional specialization and diversity might drive regional economic dynamics simultaneously, but on different basis.

A recent stream of literature claims that neither regional specialization nor regional diversity can provide efficient conditions for regional growth. Learning among very similar firms in a specialized region might not lead to growth, because the probability of

1 a Centre for Economic and Regional Studies, Hungarian Academy of Sciences, Budaörsi út 45, 1112 Budapest, Hungary
b International Business School Budapest, Tárogató út 2-4, 1021 Budapest, Hungary, blengyel@gmail.com
c Faculty of Economics and Business Administration, University of Szeged, Kálvária sgt. 1, 6722 Szeged, Hungary, kano.izabella@eco.u-szeged.hu.

obtaining new knowledge is low; on the contrary, firms have to share some knowledge in order to understand each other. Thus, a golden mean of technological proximity between co-located firms are needed in order that inter-firm learning provide ground for regional growth (Boschma 2005). This phenomena was first captured quantitatively by Frenken et al. (2007) by the formulation of related variety in a region, which builds on the concept of technological proximity and demonstrates that regional employment growth does not depent on diversity per se but the extent, to which industries are related to each other, is crucial.

The contribution of our paper is twofolds. First, we address the role of related variety in regional employment decline and growth in a post-socialist economy. Related variety might capture hidden characteristics of regional dynamics over economic transition, because previously prosperous industrial areas faced economic downturn and technological relatedness might have a special role in that. Second, we argue that additional proximity dimensions have to be involved in related variety calculations in order to understand regional economic growth in these areas. This latter step is important, because lagging firms may be isolated from the dynamic ones in terms of institutional or social conditions; thus not technological division but the lack of institutional or social proximity is what hinders inter-firm learning and regional growth.

Our demonstration is based on two distinctions. First, we analyse the role of related variety in rising and declining regions separately. Second, we decompose variety measures into domestic and foreign subsets. This latter step, namely the introduction of ownership categories into regional related variety calculations, requires an additional level of entropy decomposition and a new variable: ownership variety.

The findings suggest that related variety had a dual role in Hungarian regional development over the 1998-2005 period. It speeded up employment growth in those regions that could face the challenges of transformation. However, related variety had a negative effect on employment growth in declining regions, which suggests that technological relatedness pushed regions with uncompetitive industrial structure on a downhill path. We also find that regional employment growth is affected positively by related variety among the domestic set of firms initially but relatedness in the foreign subset started to contribute significantly to growth at a later stage.

The remainder of this paper is as follows. We give a brief overview from the Hungarian regional development in the post-socialist period, which is followed by the introduction of our data. The adaptation of related variety calculation to our special case is described in the fourth section. The fifth section gives an overview of results and discussion of findings as well as future research can be found the sixth section.

Hungarian regional development

The determining role of FDI, the remaining presence of some state-controlled services, and stagnating domestic companies have been the main features of transition economies in their current development model (Szanyi 2003). In the first half of the transition period, from 1990 to 1995, a massive economic downturn occurred in Hungary. Big state-owned companies either went bankrupt or got privatized; the latter was followed by basic restructuring. Consequently, unemployment rate, and especially long-term unemployment increased dramatically. MNEs started to carry out large investment projects in the tradable and services sectors (eg. automotive and ICT) and untraded sectors with secure local markets (eg. energy and communication) of Hungary. Simple, cheap unskilled labour-based activities were developed by additional investments (Iwasaki 2007).

Economic catching up started from year 1995, and employment rate approached again the level of year 1992 at the end of the period of our investigation. New, higher value-added activities were launched, which utilized local skilled labour and engineering talent; some of the foreign companies started to locate their R&D functions to their Hungarian sites (Lengyel–Cadil 2009).

The transition had a major footprint on regional development. Previously specialized industrial regions fell back quickly because they lost their market and became unable to meet the challenges of global competition (Lux 2009). Both regional specialization and spatial concentration of industries were proved to have a negative effect on regional employment growth (Lengyel–Szakálné 2013). Regional development is thought to be investment-driven, in which foreign direct investments have central role.

However, the role of foreign-owned firms in regional dynamics is still unclear. On the one hand, foreign-owned firms imported new knowledge to the economy, and many argue that this created positive spillover effects (Halpern–Muraközy 2007, Szanyi et al. 2011). On the other hand, domestic suppliers had only marginal roles in supplier networks of MNEs because their local decisions were usually determined by the parent company headquarters abroad (Grosz 2006). In many cases suppliers and competitors of these MNEs were mainly de-novo foreign firms that had followed their main partners into Hungary (Békés 2005). Thus, a dual structure of economy has evolved in Hungary, which can be characterised by a sharp foreign-domestic gap (Farkas 2000).

The dual economic structure and the gap between foreign-owned and domestic firms have had effected regional development as well. Those regions became relatively more competitive, in which foreign-owned companies invested (Lengyel 2003). However, foreign-owned firms have had positive effects in the relatively developed regions only and some argue that they have even destructed lagging regions (Lengyel–Leydesdorff 2011, 2013).

One might conclude that regional decline and catching up, the transition period itself, and the gap between foreign-owned and domestic companies created a unique field for testing the role of related variety in regional employment growth over post-socialist era. The current paper addresses two central questions:

1. What is the role of related variety in regional employment decline and growth?
2. How did domestic and foreign related variety affect regional employment growth?

Data

The information used for the empirical analysis in this paper was collected from the annual census-type data of Hungarian firms, which were compiled from financial statements associated with tax reporting submitted to the National Tax Authority in Hungary by legal entities using double-entry bookkeeping. The observation period covers 1998 and 2005 on a yearly basis. The data includes all industries and contains basic information for each sample firm, including the LAU1 region of company seats, NACE 4-digit industrial classification codes, the annual average number of employees, the amount of equity capital held by type of owners, and major financial indices at the end of the term.

Foreign ownership is attributed to a firm when 10% or more shares of the stocks of a firm are in foreign hands (HSCO 2007). This standard definition of the Hungarian Statistical Office considers a significant foreign interest in all of these firms even if domestic ownership is higher than foreign ownership in the firm.

All industries are present in the data, although for practical reasons we have excluded agriculture from the analysis and focus only on manufacturing and service sectors. One can find the distribution of firms along industry classifications according to 2-digit NACE categories in Appendix 1.

A major limitation of our data is that due to a change in company codes in year 2002, firms cannot be traced over the whole period. There is even a huge shift in terms of firm numbers across the two periods (Table 1). Although the data represents the total economically active population in a rather similar way (42% in 1998 and 48% in 2002); the regional, sectoral, ownership distribution of firms might be very different across the two datasets. Therefore, we analyse regional employment growth in two distinct periods: 1998-2001 and 2002-2005.

Table 1

Employment and number of firms in our data

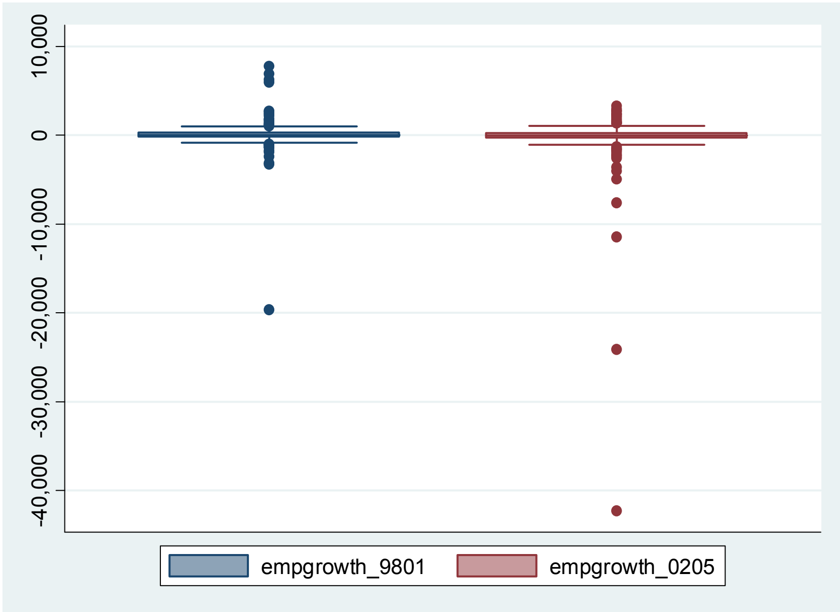
	1998	2002
Employment in the data	1,781,466	2,092,942
Share in economically active population, %	42	48
Domestic employment	1,196,222	1,563,175
Foreign employment	585,244	529,767
Number of firms	112,075	298,031
Number of domestic firms	93,736	272,111
Number of foreign firms	18,339	25,920

Note: Economically active population was 4.263 million employees in 1998; and 4.298 million employees in 2002 http://www.ksh.hu/docs/eng/xstadat/xstadat_long/h_qli001.html

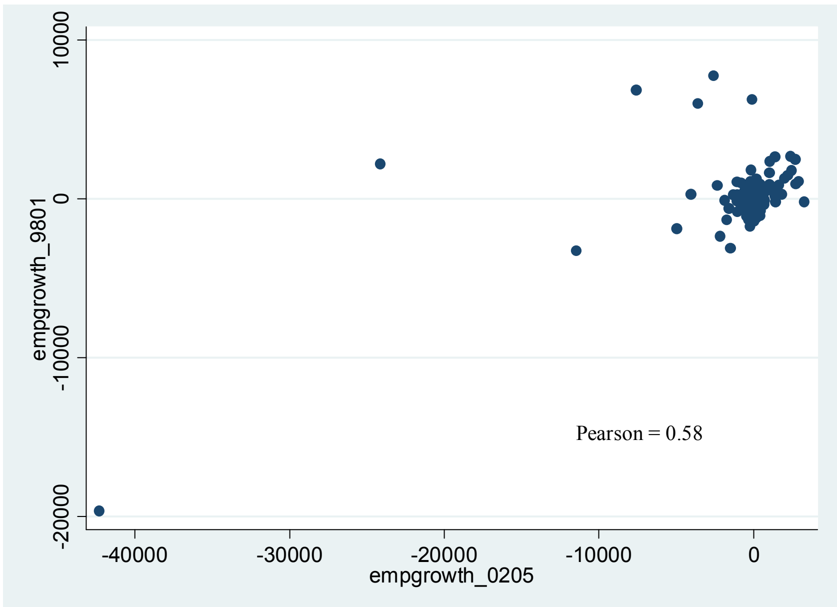
Despite company-level differences in the data, regional employment growth seems to be comparable across the two timeperiods under investigation. The average of employment growth across regions are around zero in both timesets; however, distributions differ. While the distribution of EMPGRO_9801 is close to be symmetric: Budapest (–19,418) is the only outlier, the EMPGRO_0205 distribution is more skewed towards negative values (Figure 1a). The twoway association between the two variables depicts a medium strong correlation; employment growth in most of the regions deviate around zero (Figure 1b).

Figure 1a, b

Employment growth in regions, 1998–2001 and 2002–2005



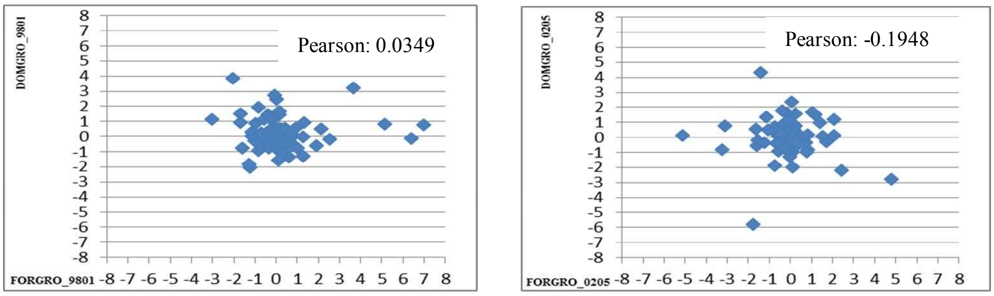
Regional employment growth across periods of investigation



Domestic and foreign employment growth scatter around zero in most of the regions and in both timesets. However, foreign growth varies on a somewhat larger scale than domestic growth: some regions stand out or lag behind more drastically in terms of foreign growth than in domestic growth (Figure 2). Maps in Appendix 2 suggest that regional employment growth stood out in the agglomeration of Budapest in both periods (except the city itself). Foreign companies contributed to growth mainly in those regions that are proximate to Budapest or lie between the capital and the Austrian border. The level of employment has decreased in the peripheral regions of the country.

Figure 2

Foreign versus domestic employment growth in subregions,
1998–2001 and 2002–2005



Note: Budapest is left out for reasons of visualization.

Methods

We follow the seminal work of Frenken et al. (2007) in variable creation. Their argument claims that two co-located firms are technologically unrelated when they don't share two-digit level NACE codes, and might not be able to learn from each other. Two co-located firms are technologically related when they share the same two-digit level NACE codes but don't share the four-digit level NACE code. Related firms might share enough knowledge but are not too proximate, therefore they can not only understand but might also learn new things from each other.

Related variety calculation is as follows. Let pi be the four-digit NACE share of employment and *Pg* the two-digit level NACE shares of employment that is derived by summing the four-digit shares. Then the variety of economic activity (V) in a region can be phrased as the sum of probabilistic entropy of four-digit level NACE shares (1). This variety can be decomposed to unrelated variety and related variety (2). Unrelated variety (UV) is given as the sum of probabilistic entropy of two-digit level NACE shares (3). Related variety (RV) is the sum of probabilistic entropy of four-digit level NACE shares within each two-digit level NACE shares (5) aggregated at the regional level (4).

$$V = \sum_{g=1}^{G} \sum_{i \in S_g} p_i \log_2\left(\frac{1}{p_i}\right) \qquad (1)$$

$$V = UV + RV \qquad (2)$$

$$UV = \sum_{g=1}^{G} P_g \log_2\left(\frac{1}{P_g}\right) \quad (3)$$

$$RV = \sum_{g=1}^{G} P_g H_g \quad (4)$$

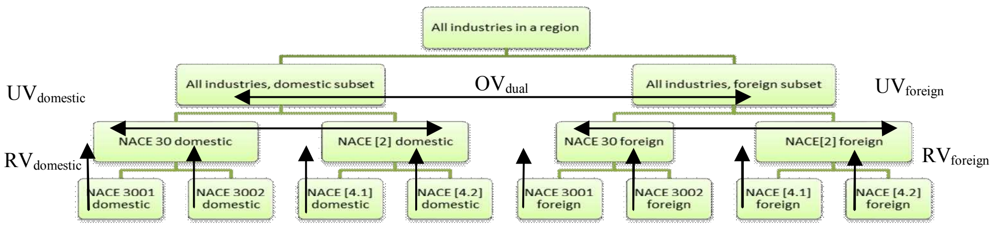
$$H_g = \sum_{i \in S_g} \frac{p_i}{P_g} \log_2\left(\frac{1}{p_i/P_g}\right) \quad (5)$$

As it follows from the above equations, related variety measures the extent to which the technological knowledge base of firms are related in a region. Relatedness is formulated on the base of technological proximity between firms. Other type of proximities into related variety calculation might be straightforward in less developed economies, because institutional gap between dynamic and lagging firms can set back inter-firm learning. The case of post-socialist transition is a good illustration: majority of domestic firms had only limited abilities to learn from the local sites of multinational companies even if they were technologically close to each other (Békés 2005, Grosz 2006).

We address this issue with introducing ownership categories into regional related variety calculations. This needs another level of entropy aggregation but enables us to decompose variety measures into domestic and foreign subsets. Unlike in previous papers, in which related variety was decomposed into subsets of manufacturing and service industries (Mameli et al. 2012) or high-tech manufacturing (Hartog et al. 2012), the introduction of ownership categories requires an additional level of entropy decomposition and a new variable: ownership variety.

Figure 3

Unrelated variety, related variety and firm ownership, a dual economy model



The formulation is visualized in Figure 3. Let poi be the share of employment in industries with four-digit NACE codes combined with ownership categories. Let poi sum up to Pog that is the share of employment in two-digit NACE codes combined with ownership categories. Also, let the sum of Pog be Po, the share of employment in all industries combined with ownership categories. Finally, let 'd' indicate domestic set of firms and 'f' indicate foreign set of firms.

Economic variety measured in the region will be equal to the entropy of the employment distribution of the finest bin structure that is the four-digit NACE code combined with ownership category (6). Then, variety in a region equals with the variety measured in the ownership distribution (OVdual), plus domestic and foreign unrelated varieties (UVD and UVF), plus domestic and foreign related varieties (RVD and RVF).

$$V = \sum_{o=f,d} \sum_{g=1}^{G} \sum_{i \in S_g} p_{oi} log_2 \left(\frac{1}{p_{oi}}\right) \quad (6)$$

$$V = OV_{dual} + UV_{dual} + RV_{dual} \quad (7)$$

$$OV_{dual} = \sum_{o=f,d} P_o log_2 \left(\frac{1}{P_o}\right) \quad (8)$$

$$UV_{F,D} = \sum_{o=f,d} P_o \sum_{g=1}^{G} \frac{P_{og}}{P_o} log_2 \left(\frac{1}{P_{og}/P_o}\right) \quad (9)$$

$$RV_{F,D} = \sum_{o=f,d} P_o \sum_{g=1}^{G} P_{og} \sum_{i \in S_g} \frac{p_{oi}}{P_{og}} log_2 \left(\frac{1}{p_{oi}/P_{og}}\right) \quad (10)$$

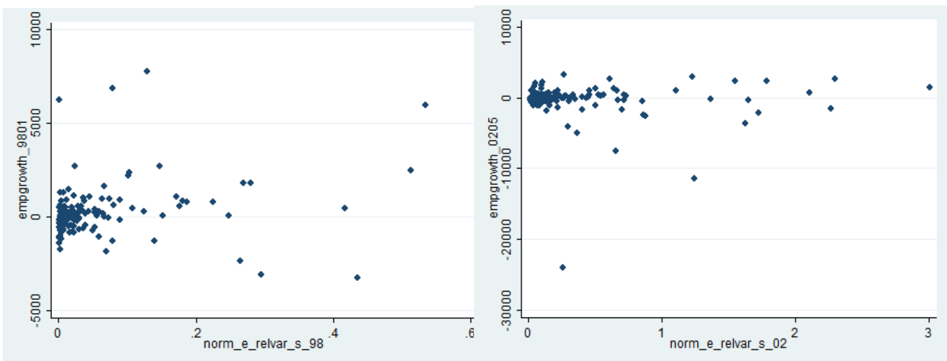
Because Hungarian subregions vary in terms of the size of their economy, we normalized variety measures by the number of employment in the appropriate categories. Accordingly, unrelated variety, related variety, and ownership variety have been divided by number of employment in the region; domestic and foreign unrelated and related variety measures have been divided by the number of employment in the respective subset in the region.

Results

According to a two-way scatter plot, Hungarian regional employment growth does not seem to be determined by related variety of economic activity in the region (Figure 4). However, one might observe a widening gap between growing and declining regions as related variety increases. Preliminary findings reflect on this issue, and the directions for further work are set as follows.

Figure 4

Related variety and employment growth



The variables created with the methods above are tested in linear regression with heteroskedasticity-robust standard errors in which employment growth is the dependent variable. The distinct time periods are handled separately and we illustrate a reverse effect of explanatory variables in declining and growing regions. As a next step we look at the effect of domestic and foreign related variety on employment growth. Finally, domestic and foreign employment growth will be analysed.

Pearson correlation values in Table 2 indicate high level of association between several variables in both timesets. For example, EMPGRO_F_9801 and EMPGRO_D_0205 correlate on a very high degree with the respected EMPGRO variables. Correlation values are also too high among few pairs of variety measures, namely UNRELVAR_98 and UV_D_98, just as between RELVAR_98, RV_D_98, and RV_F_98. Therefore, classic variety measures and variety measures decomposed to domestic and foreign subsets cannot be tested together, in order to avoid multicollinearity of explanatory variables.

A variety of control variables are introduced to the models (VIF values remains below 10 in all cases); for description of control variables see Appendix 3. Budapest is left out when looking at separate set of growing and declining regions because the capital is an extreme outlier in both of our time periods. County dummies are used in order to capture region-specific effects.

Table 2

Pearson correlation values, 1998 and 2002

1998		1	2	3	4	5	6	7	8	9
1	EMPGRO_98-01	1.0000								
2	EMPGRO_F_98-01	0.9335	1.0000							
3	EMPGRO_D_98-01	0.4045	0.0498	1.0000						
4	UNRELVAR_98	0.0444	–0.0115	0.1529	1.0000					
5	RELVAR_98	–0.7336	–0.8004	–0.0019	0.0582	1.0000				
6	OV_dual_98	0.0497	–0.0552	0.2792	0.4523	0.1212	1.0000			
7	UV_D_98	0.0315	–0.0061	0.1032	0.7755	0.0517	0.4290	1.0000		
8	RV_D_98	–0.7397	–0.8006	–0.0185	0.0784	0.9953	0.1165	0.1074	1.0000	
9	UV_F_98	0.0006	0.0040	–0.0085	0.2476	–0.0174	–0.2537	0.1914	–0.0099	1.0000
10	RV_F_98	–0.6565	–0.7093	–0.0196	0.0917	0.8766	0.0728	0.1827	0.8903	0.0101
2002		**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
11	EMPGRO_02-05	1.0000								
12	EMPGRO_F_02-05	0.3253	1.0000							
13	EMPGRO_D_02-05	0.9585	0.0423	1.0000						
14	UNRELVAR_02	0.1184	0.0271	0.1169	1.0000					
15	RELVAR_02	–0.8011	–0.2226	–0.7793	0.0641	1.0000				
16	OV_dual_02	0.0007	–0.1894	0.0578	0.1695	0.1048	1.0000			
17	UV_D_02	0.0424	–0.0643	0.0642	0.7596	0.0794	0.4103	1.0000		
18	RV_D_02	–0.7974	–0.2585	–0.7646	0.0771	0.9920	0.1240	0.1185	1.0000	
19	UV_F_02	0.0264	0.0294	0.0190	0.1507	–0.0254	–0.3262	0.1134	–0.0247	1.0000
20	RV_F_02	0.0021	–0.0092	0.0050	0.3923	0.0655	0.1445	0.4931	0.0570	0.0798

Table 3

Regional employment growth, heteroskedasticity-robust linear regression

	1998–2001								2002–2005							
	All subregions				Growth>0		Growth<0		All subregions				Growth>0		Growth<0	
	Model 1		Model2		Model 3		Model 4		Model 5		Model6		Model 7		Model 8	
CONSTANT	–228,16		40,27		947,8		–663,24	***	–1663,94	**	–23140,5		598,04	***	1212,22	
	(–1,02)		(0,16)		(1,39)		(–4,26)		(–2,49)		(–0,81)		(3.91)		(1,2)	
UNRELVAR	22776,62	**	–10658,2		–25241,5		33793,42	***	45787,56	***	47670,13	**	–11014,1	**	–31352,9	
	(2,26)		(–0,68)		(–1,27)		(3,43)		–3,29		(2,3)		(–2,64)		(–1,05)	
RELVAR	–2091,69	***	3437,39		2023,41		–4846,82	***	–956,2	***	–473,22		729,05	**	4853,92	*
	(–23.19)		(1,18)		(0,96)		(–3,64)		(–106,4)		(–0,73)		(2,48)		(1,69)	
HHI_INI	–1578,54	**	–1825,25	**												
	(–2,3)		(–2,11)													
HHI_END	3329,75	**	3955,12	**			720,13	**	2286,35	*	2171,588	**	2207,52	***		
	(2.21)		(2,6)				(2,37)		(1,85)		–2,14		(2,75)			
EMPL_INI					0,083	***	–0,012	**							–0,251	**
					(2,85)		(–2,38)								(–2,54)	
COUNTY DUMMIES	NO		YES		YES		YES		NO		YES		YES		YES	
Adjusted R-square	0,56		0,68		0,65		0,82		0,67		0,69		0,57		0,63	
N	168		168		83		84		168		168		77		90	

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively. Budapest was left out from Models 3,4,7,8..

Table 3 summarizes findings of the original related variety calculation (Frenken et al 2007); unrelated variety is calculated by (3) and related variety is formulated by (4) and (5). Interestingly, both Model 1 and Model 5 attribute a significant positive effect to unrelated variety and a significant negative effect to related variety in influencing regional employment growth. These happen to be controversial to what one expects because unrelated variety –being a measure of diversity– is thought to induce a portfolio effect. The higher the unrelated variety the more resistance against external shocks in the region; unrelated variety is expected to reduce the decline. On the other hand, related variety –being a measure of local learning capacities– is expected to enhance regional growth. Although the unexpected sign disappears when county dummies are introduced into the model in the 1998-2001 period (Model 2), unrelated variety still has a significant negative effect on growth in the 2002-2005 period (Model 6). In sum, we don't find a clear effect of related variety –that accords with previous findings in the literature– on regional growth in Hungary.

Our related idea, that merits and needs further efforts to prove in upcoming papers, is that growing and declining regions of transition economies might differ regarding the effect of related variety. In our case, inter-firm learning and technological relatedness is beneficial in those regions that could cope with post-socialist transformation. However, a large number of regions could not break out from the path they had been locked into previously and therefore technological relatedness might be another burden for them. In other words, related variety might even enhance the lag of these latter regions that did not meet the challenges of transition because local learning occurs among uncompetitive industries that lock the region into a declining path. Preliminary results in Table 3 support the idea; however, further work is needed for a detailed demonstration.

Related variety enhances growth in rising regions over the 2002-2005 period and also eases the drop of employment in lagging ones (Model 7 and Model 8). However, related variety had a negative effect on growth in declining regions over 1998-2001 indicating that it speeded up decline in backsliding regions. In our understanding, these results imply that technological relatedness and local learning have a dual effect in Hungary. Furthermore, these effects might change over time, which will have to be investigated over a longer timescale.

One might also argue that unrelated variety performs a sort of portfolio effect as proposed by Frenken et al (2007), because the higher unrelated variety the slower employment drop in sinking regions. This suggests that regions with a relatively diversified economy suffer less from path-dependent recession than regions with a more specialized economy. Accordingly, unrelated variety eases employment loss in regions that drop behind over the 1998-2001 period (Model 4), but also hinders employment growth in prospering regions in 2002-2005 (Model 7).

Table 4

Regional employment growth, dual economy model, heteroskedasticity-robust linear regression

	1998–2001								2002–2005			
	All regions		Growth>0		Growth<0		All regions		Growth>0		Growth<0	
	Model 1		Model2		Model 3		Model 4		Model 5		Model6	
Constant	270,96		–73,03		499,44		1351,48		–57,59		–1326,73	
	(0,88)		(–0,3)		(0,79)		(1,18)		(–0,16)		(–1,49)	
OV_DUAL	12103,58		–12921,9		–1098,88		12823,44		13447,41	**	35190,59	
	(1,18)		(1,25)		(–0,15)		(1,35)		(2,26)		(1,38)	
UNRELVAR_D	93775,77	**	–293318,8	*	203956,7	*	64486,65		–214526,0	***	–422333	
	(2,05)		(–1,85)		(1,78)		(0,35)		(–3,07)		(–1,06)	
RELVAR_D	–215887	***	422332,5	**	–727769	***	–245876,7	***	92814,9		248635	
	(–18,06)		(2,43)		(3,37)		(–81,56)		(1,6)		(1,15)	
UNRELVAR_F	–49,36		533,57	*	41,43		20,77		193,38		320,7347	
	(–0,22)		(1,81)		(0,15)		(0,09)		(1,03)		(0,63)	
RELVAR_F	–85,29		–436,91		2823,14		24,7		150,27	**	3111,361	***
	(–0,11)		(–0,61)		(1,56)		(0,29)		(2,09)		(3,55)	
HHI_INI			4472,28	***								
			(2,82)									
HHI_END									2093,77	**		
									(2,56)			
DOMEMPL											–0,213	***
											(–3,25)	
FOREMPL			0,22	***	–0,126	***						
			(4,52)		(–3,8)							
AVRFIRMSIZE							–243,77	*				
							(–1,52)					
AVRDOMFIRMSIZE	–20,21	*			–22,54	**						
	(–1,89)				(–2,26)							
COUNTY DUMMIES	NO		YES		YES		NO		YES		YES	
Adjusted R-square	0,57		0,81		0,73		0,67		0,56		0,63	
N	168		83		84		168		77		89	

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively. Budapest was left out from Models 3,4,7,8.

These first findings remain unchanged when unrelated and related varieties are decomposed into domestic and foreign subsets (Table 4). Note that the normalization process (described in Section 4) enables us to compare the effect of domestic and foreign populations even if foreign employment is minor compared to domestic employment in many subregions. Linear regressions with heteroskedasticity-robust standard errors in Table 4 suggest that regional employment growth is mainly affected by technological relatedness in the domestic subset of companies in the 1998–2001 period. Inter-firm learning opportunities among domestic companies might speed up regional growth (Model 2) but also has widened the lag of declining regions (Model 3). Technological relatedness among foreign firms has a significant positive effect on growth over 2002–2005, whereas no such effect was found in the 1998–2001 period.

Regional economies might have evolved over the transition period in a complex way in terms of how domestic and foreign-owned firms have been involved in local learning. Technological relatedness among domestic companies could have widened the gap between dynamic and declining regions due to the lock-in phenomenon that we described earlier. The economy has probably developed into a stage, in which relatedness of co-located domestic companies have influenced regional employment growth in all regions regardless of their previous industry structure. However, we have not found a significant effect of domestic related variety in the 2002–2005 period. On the other hand, local learning among co-located foreign-owned companies might have become crucial factor for regional employment growth. This issue merits a deeper analysis, in our opinion, in which the whole transition period shall be investigated by further papers.

The last insight addresses the cross-effect of related variety of the decomposed sets on domestic and foreign employment growth separately (Table 5). Interestingly, preliminary findings imply that relatedness among foreign firms remains unimportant for both domestic and foreign employment growth. On the other hand, relatedness among domestic companies favoured foreign employment growth in the 1998-2001 period, while domestic unrelated variety hindered foreign growth significantly. One might argue that foreign growth in Hungary does not depend on local relatedness, because foreign companies compete and co-operate not on the local level and their strategies are developed in distant headquarters. However, foreign firms might also benefit from local inter-firm learning and new relatedness measures –that capture relatedness between foreign and domestic subsets– might provide new insights to this issue.

Domestic employment growth is affected by domestic unrelated variety and domestic related variety in a similar manner as compared to the full set of firms. Unrelated variety hinders growth and related variety speeds up growth. Ownership variety also has a positive significant effect on regional employment growth in domestic firms. Since this latter variable is a simple probabilistic entropy measure of employment distribution in two categories, its' positive impact on growth implies that the closer domestic employment to foreign employment in absolute values the higher domestic growth. Thus, our result confirms previous findings that domestic employment growth is high in those regions where foreign employment is relatively also high.

Table 5

Domestic and foreign employment growth in the regions, OLS

	Domestic growth > 0				Foreign growth > 0			
	1998–2001		2002–2005		1998–2001		2002–2005	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
OV-dual	12307.77	7280.28	8189.539	9600.39	4688.786	5381.067	1211.4	8844.976
	(1.98) *	(1.14)	(2.26) **	(2.16) **	(0.33)	(0.37)	(0.21)	(1.61)
UNRELVAR_D	–312181.5		–158275.4			–619449.1		–235567.5
	(–3.85) ***		(–3.45) ***			(–3.29) ***		(–2.82) ***
RELVAR_D	305722.5		205093.3			756708.3		121910.9
	(3.01) ***		(7.42) ***			(3.16) ***		(1.66)
UNRELVAR_F		–186.53		–111.07	–838.5058		–2412.697	
		(–0.47)		(–0.25)	(–0.72)		(–1.05)	
RELVAR_F		–600.27		7.65	–809.9376		80.70932	
		(–0.55)		(0.09)	(–0.42)		(0.01)	
Constant	485.3697	381.43	168.6739	117.2	546.5496	759.548	482.689	482.3078
	(3.11) ***	(2.19) ***	(1.08)	(0.58)	(1.76) *	(2.66) ***	(1.94)	(2.41) **
Adjusted R-square	0.15	0.01	0.39	0.02	0.01	0.1	0.02 *	0.08
N	83	83	97	97	81	81	64	61

Note: ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Discussion

In this paper we have analysed the role of unrelated variety and related variety in Hungary over the 1998–2005 period. Our first special focus was a distinction between regional decline and regional growth. Second, we decomposed the classic variety measures into subsets of domestic and foreign firms.

Our first findings suggest that related variety in a region has a dual effect in regional development over post-socialist transition. Technological relatedness increases employment growth in those regions that could cope with new challenges of market economy because it creates learning opportunities among co-located firms that might lead to innovative outputs. However, related variety also speeds up employment decline in those regions that could not compete on the global economy. These regions might be locked into downhill paths, in which technological relatedness among uncompetitive industries is another obstacle. Future research on a longer timescale data shall pay extra attention to this phenomenon and upcoming papers shall focus on how the effect of related variety changes over time in a transition economy.

The second finding concerns technological relatedness in the domestic and foreign-owned subsets. While domestic related variety have predominantly influenced regional growth in the first half of the investigated period, technological relatedness among co-located foreign firms became important in the second half. A possible reason for this pattern is that the dual economic structure –which we usually talk about regarding Hungarian regional development– changes over time and foreign firms might become more integrated into the local texture of the economy at the later stages of the transition than initially. Foreign employment growth was positively affected by domestic related variety, which is a sign of cross-effects. Therefore, future papers will explore foreign-domestic relations in more details. For example, it is possible do re-organize the dual economy model of related variety calculations and introduce ownership categories on a lower level of aggregation (e.g. at two-digit or four-digit NACE code levels). These new models of decomposition might explain domestic-foreign relatedness better than the dual economy model presented in this paper. Also, other type of data (e.g. international trade at company level) might capture foreign-domestic proximities better than employment distribution.

REFERENCES

Békés, G. (2005): Location of manufacturing FDI in Hungary: how important are inter-company relationships? *MNB Working Paper Series* No. 2005/7.

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

Farkas, P. (2000): The Effects of Foreign Direct Investment on R&D and Innovation in Hungary. *IWE Working Papers* No. 108. Institute for World Economics, Hungarian Academy of Sciences.

Frenken, K.,–Van Oort, F.G.–Verburg, T. (2007) Related variety, unrelated variety and regional economic growth. *Regional Studies*,41(5): 685–697.

Glaeser, E.–Kallal, H. D.–Scheinkman, J. D.–Shleifer, A. (1992): Growth in Cities. *Journal of Political Economy* 100(6): 1126–1152.

Grosz, A. (2006): Clusterization Processes in the Automotive Industry, *Discussion Papers* No. 52, Centre for Regional Studies, Hungarian Academy of Sciences.

Halpern, L.– Muraközy, B. (2007) :Does distance matter in spillover? *Economics of Transition* 15, pp. 781–805.

Hartog, M.–Boschma, R. –Sotarauta, M. (2012):The impact of related variety on regional employment growth in Finland 1993-2006: High-tech versus medium/low-tech. *Industry and Innovation* 19(6): 459–476.

Hungarian Central Statistical Office (2007) National Accounts of Hungary, 2004-2005. Budapest. Downloaded from: http://portal.ksh.hu/pls/ksh/docs/hun/xftp/idoszaki/monsz/monsz0405.pdf

Iwasaki, I. (2007): Corporate Restructuring and the Role of Foreign Direct Investment in Hungary. In Dallago, B.–Iwasaki, I. (eds): *Corporate Restructuring and Governance in Transition Economies*. Basingstoke: Palgrave Macmillan, 178–210.

Jacobs, J. (1969): *The Economy of Cities*. New York: Random House.

Lengyel, B. Cadil, V. (2009): Innovation Policy Challenges in Transition Countries: Foreign Business R&D in the Czech Republic and Hungary. *Transition Studies Review*, 16(1): 174 188.

Lengyel, B. Leydesdorff, L. (2011): Regional Innovation Systems in Hungary: The Failing Synergy at the National Level. *Regional Studies*, 45(5): 677 693.

Lengyel, B., Leydesdorff, L. (2013) Diverse effects of FDI in regional innovation systems: synergy measurement based on complexity theory, and entropy statistics. In press, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2255521

Lengyel, B., Szakálné, K. I. (2013) Regional economic growth in Hungary 1998-2005: what does really matter in clusters? In Press. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2303748

Lengyel, I. (2003) The pyramid model: enhancing regional competitiveness in Hungary. *Acta Economica*, 53, 4: 323-342.

Lux, G. (2009) Divergent patterns of adaptation among Central European Old Industrial Regions. *European Spatial Research and Policy* 1, pp. 145-157.

Mameli, F., Iammarino, S., Boschma, R. (2012), Regional variety and employment growth in Italian labour market areas: services versus manufacturing industries, *Papers in Evolutionary Economic Geography* 12.03, Utrecht University

Marshall, A. (1890): *Principles of Economics*. London: Macmillan.

Porter, M. (2003): The Economic Performance of Regions. *Regional Studies*, 37(6-7): 549 578.

Szanyi, M. (2003): An FDI-Based Development Model for Hungary – New Challenges? *IWE Working Paper* 141, Institute for World Economics, Hungarian Academy of Sciences.

Szanyi, M. Iwasaki, I. Csizmadia, P. Illésy, M. Makó, Cs. (2011): Cluster Development in Hungary: Searching for a Critical Mass of Business via Cluster Mapping. In Dallago, B. Guglielmetti, C. (eds): *Local Economies and Global Competitiveness*, Basingstoke: Palgrave Macmillan, 113 133.

UNCTAD (2005): *World Investment Report 2005: Transnational Corporations and the Internationalization of R&D*. New York and Geneva: United Nations.

Acknowledgement

We acknowledge repeated discussions with Ron A. Boschma and Imre Lengyel. Miklós Szanyi helped us to get access to the data. The work of Balázs Lengyel was supported by the Hungarian Scientific Research Grant (PD 106290) and IBS Research Grant.

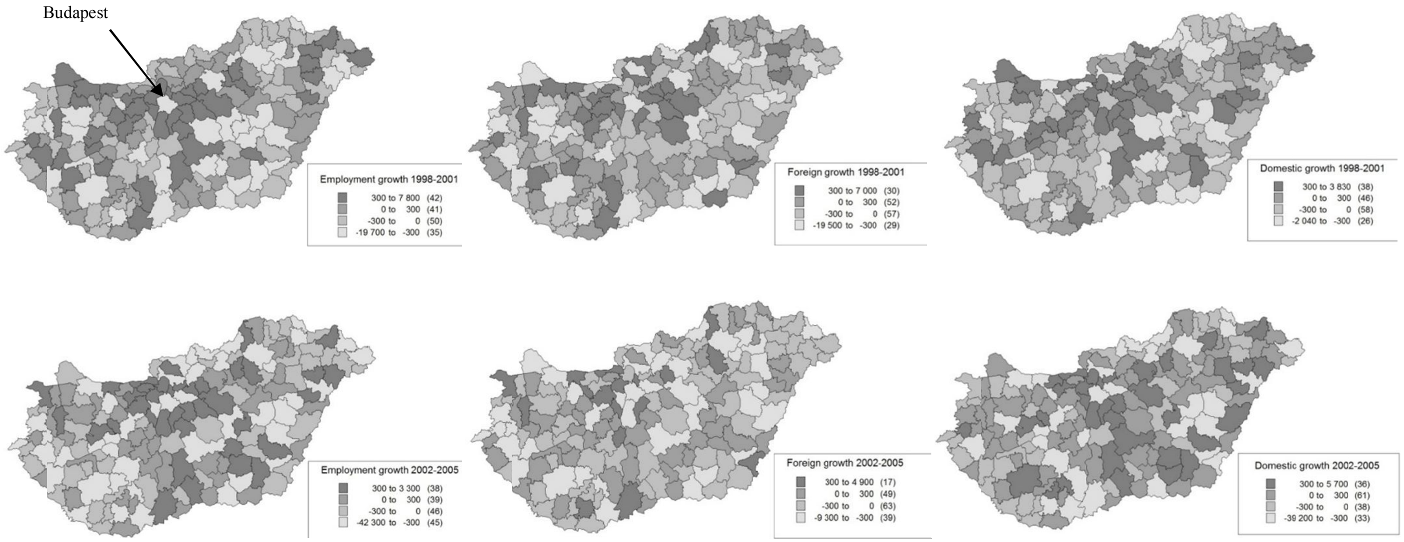
Appendix 1

Number of firms in industries, NACE 2 level

NACE2	1998	2002
15	2,557	4,226
16	7	6
17	727	1,394
18	1,071	2,463
19	348	554
20	1,129	2,631
21	229	495
22	2,363	6,14
23	7	13
24	501	711
25	960	1,754
26	676	1,48
27	227	333
28	2,478	5,532
29	1,95	3,454
30	131	307
31	669	1,192
32	486	936
33	753	1,648
34	221	294
35	123	263
36	965	2,776
37	83	211
40	201	287
41	214	307
45	10,025	27,805
50	4,801	11,767
51	20,118	32,088
52	12,988	40,815
55	4,05	13,708
60	2,58	7,057
61	47	104
62	46	76
63	1,754	3,634
64	315	953
65	575	797
66	24	94
67	532	2,801
70	6,334	16,879
71	765	1,622
72	3,143	11,118
73	627	1,687
74	17,213	50,024
80	833	5,203
85	1,523	12,03
90	471	790
91	439	2,519
92	3,051	12,012
93	745	3,052

Appendix 2

Employment growth, foreign employment growth, and domestic employment growth in Hungarian subregions; 1998–2001 and 2002–2002



Appendix 3

Description and statistics of control variables

Abbreviation	Description		Mean	St. Dev.	Min	Max
HHI_INI	Hirschman-Herfindhal index at the starting point of our period, calculated from company-level market shares.	'98	.1409016	.1509218	.0063626	.8974066
		'02	.1044115	.1210299	.0061086	.7810547
HHI_END	Hirschman-Herfindhal index at the ending point of our period, calculated from company-level market shares.	'98	.1298001	.1388309	.0093472	.9825625
		'02	.0941696	.1188845	.0070743	.767398
EMPL_INI	Absolute value of initial employment level in the region.	'98	10603.96	55852.08	172	720730
		'02	12457.99	61352.4	520	789646
DOMEMPL	Absolute value of initial domestic employment level in the region.	'98	7120.369	35993	157	463912
		'02	9304.613	44666.19	380	574023
FOREMPL	Absolute value of initial foreign employment level in the region.	'98	3483.595	19934.86	0	256818
		'02	3153.375	16834.9	0	215623
AVRFIRMSIZE	Average number of employees in firms in the region.	'98	511.6148	4502.682	4.598214	41518
		'02	7.423348	2.526928	2.915344	14.52036
AVRDOMFIRMSIZE	Average number of employees in domestic firms in the region.	'98	15.00056	6.982313	4.90625	57.17213
		'02	252.6752	3196.234	2.603261	41434

REGIONAL STATISTICS, 2013, VOL. 3: 117–140 DOI: 10.15196/RS03107

LENKA HUDRLÍKOVÁ[a] – JANA KRAMULOVÁ[b] – JAN ZEMAN[c]

Measuring Sustainable Development at the Lower Regional Level in the Czech Republic based on Composite Indicators

Measuring Sustainable Development in Czech LAU 1 Regions using Composite Indicators

Abstract

Measuring of sustainable development seems to be a great issue as there is neither unified set of indicators nor any preferred methodology how to do it. However, attempts to evaluate entities from the point of view of sustainable development regularly occur. The most problematic level according to sustainable development assessment seems to be "lower" regional level, such as LAU 1 (former NUTS 4) level. On one hand there are usually at this level already serious problems with data availability, on the other hand it is almost impossible to regularly perform detailed questionnaire surveys in all LAU 1 regions (77 districts in case of the Czech Republic), as it is done e.g. in cities. The aim of the paper is to decide, how to assess sustainability at this level. Finally we chose useful indicators, though different from indicators used at the national or NUTS 3 level, with data available for all LAU 1 regions. We succeeded in filling all three pillars of sustainable development (economic, social and environmental) with sufficient number of suitable indicators. In the first step we applied cluster analysis to find coherences among regions that are e.g. affected by similar problems. After that we constructed composite indicators in order to create ranking of all 77 districts. Ranking was derived from composite indicator approach. We constructed 10 composite indicators to test different methods of normalisation, weighting and aggregation. The results show the ranking of LAU 1 regions in the Czech Republic from the sustainability point of view, both including and excluding capital city of Prague as an outlying district. A good interconnection between cluster analysis and constructed composite indicators can be seen, which is supported by the discussion of the results.

Keywords: Sustainable development indicators, normalisation, weighting and aggregation methods, composite indicators, Czech LAU 1 regions, cluster analysis.

a Department of Economic Statistics, University of Economics in Prague, Nám. W. Churchilla 4, Praha 3, CZ 130 67, Czech Republic. E-mail: lenka.hudrlikova@vse.cz
b Department of Regional Studies, University of Economics in Prague, Nám. W. Churchilla 4, Praha 3, CZ 130 67, Czech Republic. Author is also working in the Czech Statistical Office. E-mail: jana.kramulova@vse.cz
c Department of Economic Statistics, University of Economics in Prague, Nám. W. Churchilla 4, Praha 3, CZ 130 67, Czech Republic. E-mail: janzeman06@gmail.com

Introduction

Measuring of sustainable development seems to be a great issue (see e.g. Parris et al., 2003) as there is neither unified set of indicators nor any preferred methodology how to do it. However, attempts to evaluate entities according to sustainable development regularly occur. Especially at the national level various indicators set are being created (e.g. Eurostat has one, see EUROSTAT 2013), as well as at the lowest level for cities (e.g. ECI by European Commission 2013, used in the Czech Republic by TIMUR 2012) or even enterprises. Similarly, at the "higher" regional levels (NUTS 2 and in case of the Czech Republic NUTS 3 level as well) there are also attempts to evaluate sustainability using a set of indicators (Progress Reports on the Czech Republic's Strategic Framework for Sustainable Development). The most problematic level according to sustainable development assessment seems to be "lower" regional level, such as LAU 1 (former NUTS 4) level. At this level there are already serious problems with data availability (i.e. methodological problems of regional GDP estimate or simply non-availability of reliable data). On the other hand it is almost impossible to regularly perform detailed questionnaire surveys in all regions at this level (77 in case of the Czech Republic), as it is done e.g. in case of ECIs. From these reasons it is not possible to use similar approach as in case of "higher" regional level or local level.

This paper is part of our project, which deals with the analysis of sustainable development in the Czech Republic at the regional level (NUTS 3, see e.g. Fischer et al., 2013). In this paper we decided to focus also on the lower level of administrative division – on the district level (LAU 1, formerly labelled NUTS 4) – as this level can be seen to be important as well (compare Lengyel et al., 2012). The main aim of the paper is to decide, how to assess sustainability at LAU 1 level. Unfortunately, this level has not been a subject to extensive research in the Czech Republic so far. An interesting approach is applied by Mederly et al. (2004) who analysed sustainability and quality of life in the Czech Republic at three different levels – regional, national and global. However, the regional level was limited exclusively to NUTS 3 level. They chose a huge number of 111 indicators that were initially analysed using correlation analysis and further deeper analysed. Another important approach was employed by the Czech Statistical Office (2010) again at the NUTS 3 level. Together with the Charles University in Prague Environment Centre, they analysed indicators in time series divided into three common pillars of sustainable development – economic, social and environmental. Some of these indicators were also used in a strategic document – Progress Reports on the Czech Republic's Strategic Framework for Sustainable Development in the year 2009 (Government Council for Sustainable Development et al. 2009). The new version of this document, published in the year 2012 (Government Council for Sustainable Development et al. 2012), already works with slightly different indicators in a different structure.

As there are no attempts at the LAU 1 level, this paper shows some opportunities, how to assess sustainability at this level. The paper is divided into several sections. The first one deals with the data availability and indicators finally included in the analysis (together with analysis of their correlations). The second section deals with potential coherences among Czech LAU 1 regions, e.g. due to affection by similar problems. Cluster analysis covering all selected indicators was applied to examine this. The third section introduces a brief overview of the methodology of composite indicators. Further the methods of normalisation, weighting and aggregation are used for composite indicators construction as well as for subsequent rankings creation. In the fourth section the results are presented and discussed. In the final section main conclusions are outlined.

Sustainable development and set of indicators

The main idea of sustainable development lies in searching for balance among economic development, social progress and equity and environmental responsibility. The Brundtland Commission definition (WCED, 1987, p. 8) is usually considered to be the main definition broadly accepted. It emphasizes especially the people's needs while expressing that "*Humanity has the ability to make development sustainable – to ensure that it meets the needs of the present without compromising the ability of future generations to meet their own needs*". However, many other definitions can be found, that bring the discussions on how to really understand the sustainable development and its targets and measurement (for several of them see e.g. Marsden et al. 2010, Byrch et al. 2009, Ciegis et al. 2009, Rassafi et al. 2006, Macháček 2004, p. 28–29 or Nováček et al. 1996, p. 16–19).

We focused on the Czech Republic as a case study. There are 77 districts (i.e. LAU 1 regions) in the Czech Republic including the capital city of Prague, which is very specific among other districts because it is not only a district (LAU 1), but at the same time also a region (NUTS 3) and even NUTS 2 unit. This is not so common in other countries, although capital cities often form a specific region with unusual characteristics. Usually the higher the level of classification, the broader area of the city is included (i.e. with suburban areas which have different characteristics from the core city). Cambridge Econometrics (2013) state that "In general, NUTS 3 regions are used to define the cities, in recognition of the fact that many cities are essentially spatially-concentrated cores of economic interaction which are smaller than NUTS 2 regions, with the important exceptions of the major conurbations such as Paris and London". For example London[d] as NUTS 1 is divided into Inner London and Outer London (NUTS 2 regions) and further into five NUTS 3 regions and then into 33 LAU 1 regions. On the other hand in case of Prague the area included in NUTS 2, NUTS 3 and LAU 1 classifications is exactly the same[e]. Therefore, we decided to perform two types of analysis – including and excluding Prague – and compare the results obtained.

From a statistical point of view and due to the need to meet certain conditions for the use of multivariate methods, the number of 77 units seems to be much more appropriate than the number of 14 regions forming the Czech Republic[f]. This was one of our reasons for focusing on this level. On the other hand the disadvantage of such small territorial units (e.g. districts) consists in the problem of data availability. The problem lies in the fact that each district often follows its own characteristics (indicators), which may differ across districts, or indicators available for some districts are no more available for other districts. According to these findings we were forced to find set of indicators different from the one used in the analysis at the NUTS 3 level (Fischer et al. 2013).

d For more details see Office for National Statistics (2013).
e See Methodology section in Statistical Yearbook of the Czech Republic 2012 (Czech Statistical Office 2012, p. 771)
f Appendix 1 shows map with 77 Czech LAU 1 districts.

Table 1

Correlation matrix of environmental pillar

Arable land (%)	Coefficient of ecological stability	Share of broadleaved species (%)	Specific emissions of solids (tonne per km2)	Specific emissions of sulphur dioxide (tonne per km2)	Specific emissions of nitrogen oxides (tonne per km2)	Specific emissions of carbon monoxide (tonne per km2)	Number of small-scale protected areas	Share of protected areas (NP+PLA+S-SPA) on region area (%)	Investment environmental protection expenditure by the investor registered office	Share of agricultural land (%)	Share of agricultural holdings having the agricultural land area 500+ ha (%)	Share of municipalities with established public water supply system covering all the municipality (%)	Share of municipalities with established sewerage system covering all the municipality (%)	Share of municipalities with established sewerage systém connected to a WWTP covering all the municipality (%)	Share of municipalities with established natural gas grid covering all the municipality (%)
1,000	**-0,904**	0,235	-0,098	-0,081	-0,056	-0,065	-0,154	-0,690	0,052	0,781	0,149	0,132	0,320	0,169	0,440
-0,904	1,000	-0,375	-0,050	-0,115	-0,139	-0,023	0,274	0,739	-0,090	-0,762	-0,128	-0,259	-0,375	-0,257	-0,431
0,235	-0,375	1,000	0,259	0,330	0,364	0,220	-0,158	-0,118	0,277	0,183	-0,228	0,464	0,346	0,397	0,433
-0,098	-0,050	0,259	1,000	0,706	0,810	**0,920**	-0,254	-0,046	0,216	-0,166	-0,173	0,024	-0,182	-0,082	-0,081
-0,081	-0,115	0,330	0,706	1,000	**0,942**	0,508	-0,341	-0,135	0,177	-0,276	0,050	0,173	-0,123	0,023	0,043
-0,056	-0,139	0,364	0,810	**0,942**	1,000	0,612	-0,309	-0,146	0,178	-0,207	-0,012	0,212	-0,097	0,048	0,015
-0,065	-0,023	0,220	**0,920**	0,508	0,612	1,000	-0,091	-0,006	0,246	-0,121	-0,185	-0,059	-0,193	-0,162	-0,050
-0,154	0,274	-0,158	-0,254	-0,341	-0,309	-0,091	1,000	0,361	0,091	-0,122	-0,177	0,086	0,185	0,061	0,072
-0,690	0,739	-0,118	-0,046	-0,135	-0,146	-0,006	0,361	1,000	-0,055	-0,485	-0,265	-0,109	-0,341	-0,247	-0,301
0,052	-0,090	0,277	0,216	0,177	0,178	0,246	0,091	-0,055	1,000	0,006	-0,133	0,126	0,158	0,190	0,270
0,781	-0,762	0,183	-0,166	-0,276	-0,207	-0,121	-0,122	-0,485	0,006	1,000	0,019	0,046	0,174	-0,016	0,375
0,149	-0,128	-0,228	-0,173	0,050	-0,012	-0,185	-0,177	-0,265	-0,133	0,019	1,000	-0,048	-0,073	-0,023	-0,030
0,132	-0,259	0,464	0,024	0,173	0,212	-0,059	0,086	-0,109	0,126	0,046	-0,048	1,000	0,579	0,575	0,507
0,320	-0,375	0,346	-0,182	-0,123	-0,097	-0,193	0,185	-0,341	0,158	0,174	-0,073	0,579	1,000	0,833	0,465
0,169	-0,257	0,397	-0,082	0,023	0,048	-0,162	0,061	-0,247	0,190	-0,016	-0,023	0,575	0,833	1,000	0,239
0,440	-0,431	0,433	-0,081	0,043	0,015	-0,050	0,072	-0,301	0,270	0,375	-0,030	0,507	0,465	0,239	1,000

Source: Own calculation.

Note: Gray marked indicators were discarded due to correlations. Bold underlined numbers in grey fields stand for values greater than ±0.9.

The set of indicators for districts was carefully chosen in order to be the best fitting to the sustainable development issue. It is widely recognized that sustainable development indicators are grouped into (most commonly) three pillars (economic, social and environmental). Part of the indicators chosen for the district level is the same as those used in the analysis of regions, some of them had to be adapted to the available data sources and the rest needed to be newly chosen after the discussions with experts. Our analysis focuses on the year 2010 being the most up-to-date year with data available for all selected indicators. We tried to find all relevant indicators that were available for the district level.

Before starting the main analysis we calculated the correlation matrices for each of the pillars to identify and eliminate redundant indicators (it is needed to pay a great attention to the context and explanation of each indicator in order not to discard the indicators that have a high degree of correlation with indicators which they are not directly related to with). Table 1 below shows correlation matrix of environmental pillar as an example. It contains 16 originally included indicators, 3 of them (two types of emissions and one indicator covering sewerage system) were eliminated according to the results of correlations. The highest (positive) correlation coefficients of 0.942 and 0.920 which lead to elimination of indicators were observed between the different types of emissions (serves as an example of indicators we can discard due to their relationship). On the contrary, high (negative) correlation coefficient of –0.904 was identified between indicators *Arable land* and *Coefficient of ecological stability*. As this can be considered as an example of indicators we cannot discard due to their incoherence, both indicators were left in the analysis.

After such adjustments in all three pillars we obtained following set of indicators we are using onwards. While Table 2 shows indicators selected into economic pillar, tables 3 and 4 show indicators in social and environmental pillar.

Table 2

Indicators in economic pillar (13 indicators)

EC 1	Density of motorways and 1st class roads (km per 100 km^2)
EC 2	Average value per building notification and/or permit (CZK thousands)
EC 3	Number of entrepreneurs (natural persons and legal persons) per 1000 inhabitants
EC 4	Foreign direct investment for 1000 inhabitants (CZK millions)
EC 5	Number of people receiving unemployment benefit per 100 job applicants
EC 6	Building permits per 1000 inhabitants
EC 7	Approximate value of construction projects permitted by planning and building control authorities (CZK millions)
EC 8	Domestic construction work “S” (CZK millions, current prices)
EC 9	Operated vehicles (per 1 inhabitant).
EC 10	Number of enterprises with more than 50 employees
EC 11	Share of total number of natural persons carrying out business (natural persons carrying out business in compliance with Trades Licensing Act, self-employed farmers and agricultural entrepreneurs, private entrepreneurs in business carrying out business activities governed by regulations other than the Trades Licensing Act of economically active inhabitants (%)
EC 12	Registered motor vehicles per 1 inhabitant (cars and vans)
EC 13	Share of the population living in towns (%)

Source: Own compilation based on expert discussion.

Table 3

Indicators in social pillar (17 indicators)

SO 1	Total general unemployment rate (%)
SO 2	Life expectancy at birth (years)
SO 3	Civil society – political participation (turnout in elections to municipal councils in %)
SO 4	Women and men in politics (share of women elected representatives in elections to municipal councils in %)
SO 5	Civil society – civic participation (number of mid-year population per 1 nongovernmental non-profit organization)
SO 6	Number of job vacancies per 100 applicants
SO 7	Age index (number of inhabitants aged +65 per 100 inhabitants aged 0-14)
SO 8	Share of municipalities with medical facilities (%)
SO 9	Share of municipalities with school (%)
SO 10	Average percentage of incapacity for work
SO 11	Average time of sick leave (days)
SO 12	Number of places in social services per 1000 inhabitants
SO 13	Number of doctors per 1000 people (outpatient care)
SO 14	Number of recipients of pensions per 100 inhabitants
SO 15	Average old-age pension (CZK)
SO 16	Number of disabled people-licensee holders per 100 inhabitants
SO 17	Total paid social benefits per 1 inhabitant (CZK)

Source: Own compilation based on expert discussion.

Table 4

Indicators in environmental pillar (13 indicators)

EN 1	Arable land (%)
EN 2	Coefficient of ecological stability
EN 3	Share of broadleaved species (%)
EN 4	Specific emissions of nitrogen oxides (tonne per km2)
EN 5	Specific emissions of carbon monoxide (tonne per km2)
EN 6	Number of small-scale protected areas
EN 7	Share of protected areas (NP + PLA + S – SPA) on region area (%)
EN 8	Investment environmental protection expenditure by the investor registered office
EN 9	Share of agricultural land (%)
EN 10	Share of agricultural holdings having the agricultural land area 500+ ha (%)
EN 11	Share of municipalities with established public water supply system covering whole municipality (%)
EN 12	Share of municipalities with established sewerage system connected to a WWTP covering whole municipality (%)
EN 13	Share of municipalities with established natural gas grid covering whole municipality (%)

Source: Own compilation based on expert discussion.

It is obvious, moving to LAU 1 level means that we can no more work with indicators such as e.g. GDP per capita or labour productivity, which are usually an essential part of the analysis at the higher regional level (NUTS 3 or NUTS 4). This set of indicators represents available but also reliable and relevant data at this level.

Analysis of similarities among Czech LAU 1 regions

After selecting the most proper indicators we decided to analyse the similarities in Czech LAU 1 regions according to all 43 indicators without paying attention in which pillar they are incorporated. For this issue we chose one of the multivariate statistical methods – cluster analysis. We tried to group homogenous LAU 1 regions and examine, if such regions with similar problems (e.g. unemployment, structurally affected regions, border regions or highly developed city regions) will belong to the same cluster.

Cluster analysis (Burns et al. 2009, Mooi et al. 2011 or Hebák et al. 2007) is a method of data classification, which performs the division of data into groups that contain units having something in common. The aim of cluster analysis is to divide n LAU 1 regions into k groups, called clusters, using p indicators. Like other types of statistical methods, cluster analysis has several variants which differ in the coalescing process as well; in our case hierarchical clustering is used. We used the Euclidean distance (see Equation (1)) as the distance between two points in the Euclidean space. "Euclidean distance is the most commonly used type when it comes to analyzing ratio or interval-scaled data." (Mooi et al. 2011, p. 245). Mimmack et al. (2001) state that when the cluster analysis is used for defining regions, which is our situation, Euclidean distance seems to be more proper than Mahalanobis distance. Furthermore, Ward's method as one of the methods of joining clusters, used as a linking clusters criterion in the sense of increase of the total intragroup sum of squared deviations of individual observations from the cluster average was applied. Increase is expressed as the sum of squares in the emerging cluster minus the sum of squares in both merging clusters as shown in Equation (2).

$$D_E(x_i, x_{i'}) = \sqrt{\sum_{j=1}^{p}(x_{ij} - x_{i'j})^2}, \qquad (1)$$

$$G = \sum_{h=1}^{k}\sum_{i=1}^{n_h}\sum_{j=1}^{p}(x_{hij} - \bar{x}_{hj})^2, \qquad (2)$$

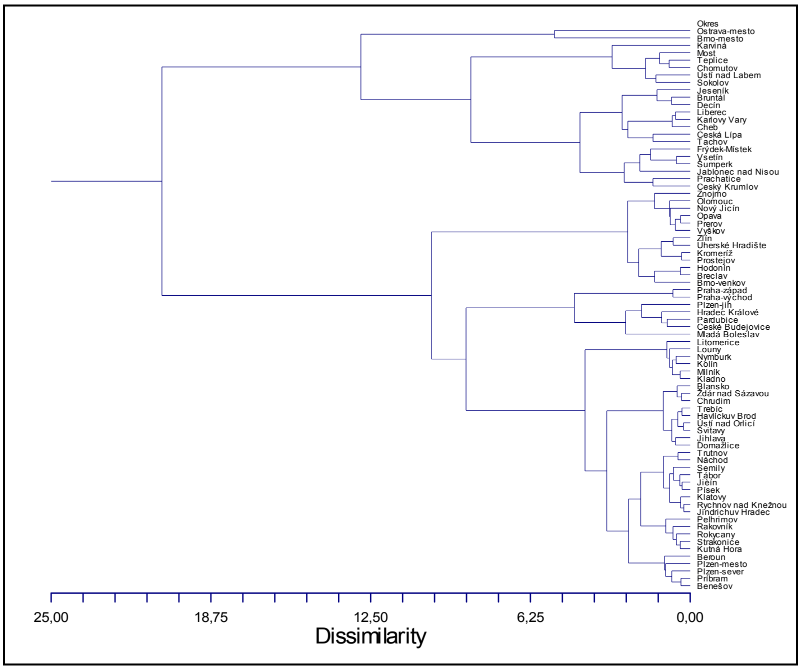
where G stands for Ward´s criterion, k stands for number of clusters, n_h stands for number of LAU 1 regions in h^{th} cluster, p stands for number of indicators.

Burns et al. (2009, p. 557) emphasize about Ward's method that "in general, this method is very efficient". Hebák et al. (2007, p. 135) sees another advantage of Ward's method. It forms similarly big cluster when eliminating the small ones.

The same approach (hierarchical clustering with Euclidean distance and Ward's method) was applied e.g. by Odehnal et al. (2012) when evaluating competitiveness of Ukrainian regions. We performed hierarchical clustering within the statistical software NCSS 2007 environment. Based on the results (Figure 1) we agreed on the final number of 6 clusters with the degree of dissimilarity of 6 as a reasonable number. The results of the cluster analysis were captured into individual choropleth maps shown below. It is important to mention that after selecting the indicators and initial data analysis we decided to remove from this analysis the capital city of Prague due to its specifics and incomparableness with other districts.

Figure 1

Dendrogram as a result of cluster analysis

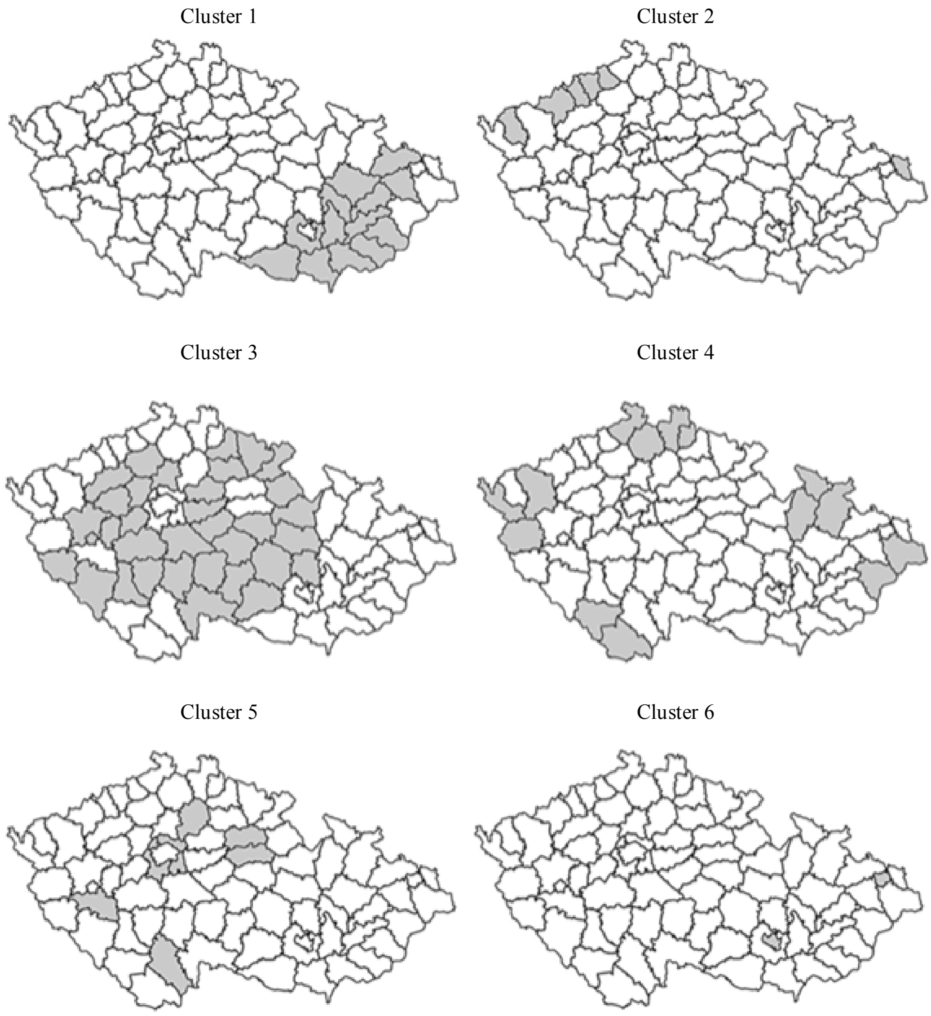


Source: NCSS 2007, own calculation.

Figure 2 shows the clusters in which many similarities can be found. Cluster 1 contains Moravian adjacent districts. Cluster 2 is formed by border districts, mostly situated northwest with one district situated northeast (districts with high unemployment rate). Cluster 3 is the largest cluster, made up of majority of the Czech districts and several Moravian districts. Cluster 4 is composed of 5 sub-clusters which do not have a common border, all of them are border (frontier) districts. Cluster 5 contains districts with medium-sized towns as centres and Prague surroundings. Cluster 6 is made up from the second and the third biggest cities in the Czech Republic (Brno and Ostrava). As mentioned above Prague as a “district-outlier” was discarded from this analysis. If left in analysis, Prague would form separate Cluster 7.

Figure 2

Six clusters identified in the Czech Republic



Source: Own analysis in Maps Generator environment.

To sum up, Figure 2 shows six clusters whose districts really have similar characteristics (not only in statistical sense, but also in the sense of sustainable development indicators), as Moravian districts, Czech districts, border districts, districts with big cities, or medium-sized towns. These conclusions encouraged us to continue in our analysis with the aim of evaluation of 77 Czech LAU 1 regions. In the next section we will already handle with indicators divided into pillars.

Methodology of composite indicators

Composite indicator (CI) is considered to be a useful tool for ranking. Overview of advantages and disadvantages of using composite indicators for sustainable development evaluation is done e.g. in Czesaný (2006). Examples of several CIs for assessment of sustainable development are listed in Parris et al. (2003). It allows expressing a multidimensional issue by one single number. We followed the generally accepted definition of sustainable development stating three areas – economic, environmental and social. Therefore aggregation needed to be performed in two steps. In the first step we aggregated indicators in each separate pillar. The second step consisted in merging all three pillars into one composite indicator. This chosen approach brought three main issues to be solved: method of transformation/normalization of the data, selection of suitable weighting scheme and finally selection of aggregation method.

As in most fields of economic reality also in sustainable development indicators are neither measured in the same units nor have the same direction. Higher values do not always reflect better performance. In other words higher value of an indicator may represent a worse performance (e.g. unemployment). Due to aforementioned reasons certain data transformation is required prior to the next analysis. The goal of the data transformation can be seen in adjustment for different ranges, different variances and outliers. There is a wide scale of normalization methods discussed e.g. in Nardo et al. (2009). Choosing the most appropriate method for normalization is crucial and depends not only on the type of data, but also on weighting and subsequent aggregation. Application of normalization can result in different outcomes for the CI. This paper deals with the two most common types: min-max method and z-scores.

The first considered transformation method is called min-max. Equation (3) is used for indicators for which the higher value the better, equation (4) for indicators for which the lower value the better.

$$I_{pn} = \frac{x_{pn} - min_n(x_p)}{max_n(x_{pn}) - min_n(x_p)}, \qquad (3)$$

$$I_{pn} = \frac{max_n(x_p) - x_{pn}}{max_n(x_{pn}) - min_n(x_p)}, \qquad (4)$$

where x_{pn} is the value of an indicator p for district n. Min-max method is based on minimum and maximum values. The advantage lies in the fact that the boundaries can be set and all indicators then get an identical range (0, 1). Each indicator reaches value between 0 and 1 even if it is the extreme value. The output is dimensionless and the relative distances remain. Nevertheless, the drawback reveals if outliers and/or extreme values are present. The computation of min-max method is based on extreme values (the minimum and the maximum) which can be outliers. These two values strongly influence the final output (see equations (3) and (4)). Still min-max approach is very popular and has been applied for construction of many composite indicators, e.g. the well-known Human Development Index (HDI), issued by the United Nations (Klugman 2011).

The second normalization method (z-scores) converts the data in order to have a common scale with a zero mean and standard deviation of one. For each indicator x_{pn} the average across districts $x_{pn} = \overline{n}$ and standard deviation across districts $\sigma'_{pn} = \overline{n}$ are calculated and used in formula (5).

$$I_{pn} = \frac{x_{pn} - x_{pn=\pi}}{\sigma'_{pn=\pi}}, \qquad (5)$$

This method provides no distortion from the mean; it adjusts different scales and different variance. The output is again dimensionless and because of applying of a linear transformation the relative differences can preserve. Z-scores method does not fully adjust for outliers. An indicator with extreme value has a greater effect on a composite indicator. This is desirable if exceptional behaviour should be rewarded, i.e. if an excellent performance on a few indicators is considered to be better than a lot of average performances. However, this effect can be reduced by applying proper aggregation method. Compared to min-max method, z-scores method is even more robust on outliers because it is based on variance instead of the range.

Weighting as the second step has a crucial effect on the outcome of the CI. There is not only one proper method always used. That is why this part of constructing a CI is the most discussed and criticised by opponents of CIs (Freudenberg (2003). Weighting methods can be divided into two main groups: statistical approaches and participatory approaches. The most common methods are listed in Nardo et al. (2009). No results from surveys, opinion poll, questionnaires etc. are available and useful for this analysis thus participatory methods cannot be used. This paper deals only with the first group of weighting methods (i.e. statistical methods). These methods are only data driven, that means no subjective value judgments are needed.

Using Equal weighting (EW) method, the equal weight is assigned for each indicator (according to equation (6))

$$w_p = 1/_{P}, \qquad (6)$$

where w_p is a weight for p^{th} indicator (p=1,...,P) for each district. It means all indicators are given the same weight for all LAU 1 regions. Equal weighting may be justified when there is no clear idea which method should be used. The main strength of EW method is definitely its simplicity. On one hand this approach is easy and clear, on the other hand, there is a risk that a pillar with more indicators will have a higher influence on the CI.

Weighting derived from principal component analysis (PCA) and factor analysis (FA) respectively deals with this issue. Both methods are very often used for data explanatory analysis. The PCA and FA explain the variance of the data through a few factors which are formed as a linear combinations of raw data. The original correlated set of indicators is changed into a new smaller set of uncorrelated variables. A detailed description of both methods can be found in e.g. Manly (2004), Morrison (2005), as well as in textbooks or handbooks about statistical software (e.g. StatSoft (2011)). In this analysis we carried out main components extraction and varimax rotation. Weights derived from the PCA are based on eigenvalues. After obtaining them it is necessary to decide about optimal number of components. Kaiser criterion suggests selecting all components which are associated to eigenvalues higher than one. Applying that criterion, in the economic and environmental pillars four factors sufficed, in social pillar five factors were covered into further computations. To obtain FA derived weights we followed approach proposed by Nicoletti et al. (1999). The weights had to be normalized by squared factor loadings, which are derived as portions of variance of the factor explained by particular variable. At the end they were scaled to unity sum. The main idea behind the usage of FA derived weights is correction of correlated indicators. Two highly correlated indicators are given lower weight because it is assumed that they can measure the same phenomenon. It is necessary to check the data for correlations before applying the weights on indicators, which was consistently done in section 1. But this method aims at further correcting of correlation.

The third step in the procedure of CI construction is aggregation. In practice linear aggregation (LIN) is the most widespread. The simplest method represents the weighted average as shown in equation (7), subject to conditions (8).

$$CI_n = \sum_{p=1}^{P} I_{pn} \cdot w_p, \qquad (7)$$

$$\sum_p w_p = 1 \text{ and } 0 \leq w_p \leq 1, \qquad (8)$$

where I_{pn} is a normalized indicator p (p=1,...,P) for district n (n=1,...,N) and w_p weight for indicator p (p=1,...,P). The fundamental topic of the aggregation is compensability among indicators. Linear aggregation implies full compensability, i.e. poor performance in one indicator can be compensated by sufficiently high values of others indicators. Compensability between indicators can be desirable only if various indicators are considered as substitutes. Even if full compensability can be weakened by the weighting scheme, another aggregation rules can suppress that.

Geometric aggregation (GEO) is only partially compensable (see formula 9).

$$CI_n = \prod_{p=1}^{P} (I_{pn})^{w_p}, \qquad (9)$$

where I_{pn} is a normalized indicator p (p=1,...,P) for district n (n=1,...,N) and w_p weight for indicator p (p=1,...,P). Geometric aggregation rewards districts with higher scores in stronger intensity because marginal utility of an increase in a low score is much higher than in a high score. Hence districts with low scores should prefer a linear rather than a geometric aggregation. The drawback lies in the requirement of strictly positive values of normalised indicators (i.e. I_{pn}>1) which means it is not applicable on normalized data by means of z-scores.

As it was already stated, aggregation was done in two steps – firstly within the pillar and after that aggregation of the three pillars. By applying above mentioned techniques we constructed ten different composite indicators. Table 5 shows all ten tested combinations.

Table 5

List of 10 tested combinations

	Normalization	Weighting	Aggregation in pillar	Aggregation of three pillars (without weights)
CI_1	Min-max	EW	Arithmetic mean	Arithmetic mean
CI_2	Min-max	EW	Arithmetic mean	Geometric mean
CI_3	Min-max	EW	Geometric mean	Arithmetic mean
CI_4	Min-max	EW	Geometric mean	Geometric mean
CI_5	Min-max	PCA/FA	Arithmetic mean	Arithmetic mean
CI_6	Min-max	PCA/FA	Arithmetic mean	Geometric mean
CI_7	Min-max	PCA/FA	Geometric mean	Arithmetic mean
CI_8	Min-max	PCA/FA	Geometric mean	Geometric mean
CI_9	z-scores	EW	Arithmetic mean	Arithmetic mean
CI_10	z-scores	PCA/FA	Arithmetic mean	Arithmetic mean

Source: Own construction.

Note: EW stands for equal weights within the pillars, PCA/FA for weights derived from principal component analysis and factor analysis.

More techniques were used in order to assess the robustness of ranking. We are aware of the fact that this is not the exhaustive list of techniques for normalization, weighting and aggregation. Our aim was to select only methods which are simple, easily understandable and only data driven. Used methods cover two main issues during constructing a composite indicator – correlation and compensability between various indicators (Paruolo et al., 2013). The differences in results as well as suitability of each CI are discussed in the next section.

Computations, results and discussion

This section introduces the main results. After normalisation of an indicator, the ranking of the districts remains the same regardless of chosen method of normalisation (min-max or z-scores). However, the values are different and further analyses are influenced by the chosen normalisation method.

Even more important is the weighting scheme. Table 6 shows weights derived from equal weighting and PCA/FA weighting within each pillar.

Table 6

Equal and PCA/FA weights within one pillar (in %)

Economic pillar	EW	PCA/FA	Social pillar	EW	PCA/FA	Environmental pillar	EW	PCA/FA
EC1	7.69	4.10	SO1	5.88	4.53	EN1	7.69	10.82
EC2	7.69	3.27	SO2	5.88	8.36	EN2	7.69	10.42
EC3	7.69	9.71	SO3	5.88	5.83	EN3	7.69	3.85
EC4	7.69	4.94	SO4	5.88	4.15	EN4	7.69	11.60
EC5	7.69	5.23	SO5	5.88	3.39	EN5	7.69	10.41
EC6	7.69	7.24	SO6	5.88	5.32	EN6	7.69	6.16
EC7	7.69	10.00	SO7	5.88	6.73	EN7	7.69	6.81
EC8	7.69	8.86	SO8	5.88	8.45	EN8	7.69	0.64
EC9	7.69	7.05	SO9	5.88	9.48	EN9	7.69	9.68
EC10	7.69	9.73	SO10	5.88	9.46	EN10	7.69	9.82
EC11	7.69	11.09	SO11	5.88	8.69	EN11	7.69	8.75
EC12	7.69	9.55	SO12	5.88	1.47	EN12	7.69	7.98
EC13	7.69	9.21	SO13	5.88	3.75	EN13	7.69	3.05
			SO14	5.88	5.58			
			SO15	5.88	6.56			
			SO16	5.88	1.91			
			SO17	5.88	6.35			

Source: Own computation.
Note: EW stands for equal weights, PCA/FA for weights derived from principal component analysis and factor analysis.

In the third step it was essential to decide about the most proper aggregation methods inside a pillar and of all three pillars. For aggregation within a pillar we used the weights computed in Table 6. We calculated all ten CIs presented in Table 5, i.e. all possible combinations of aggregation methods. However, we can conclude that especially geometric aggregation at pillar level produced unreliable results. Therefore, as a final ranking we decided to use combination recommended by Joint Research Centre[g]. Tarantola (2011) suggests using arithmetic average to combine indicators within a pillar and geometric average to merge pillars into one single composite indicator. The idea is simple: within one pillar there can be considered trade-off between indicators but the pillars should not be fully compensable. Final ranking in Table 7 is based on min-max normalization, weighted arithmetic average at pillar level and geometric average for combining three pillars. Two combinations meet these conditions, one with equal weights (CI_2) and one with PCA/FA weights (CI_6). Results for these two CIs are shown in Table 6. Unlike in cluster analysis in section 2, here is the capital city of Prague (Hlavní město Praha, hereinafter Prague) included in order to bring complete ranking of Czech districts. Number of cluster corresponds with results in section 2 (Prague was labelled as Cluster 7).

g Joint Research Centre provides scientific and technological support to European Union policies. It's Econometrics and Applied Statistics Unit aims besides other issues as well at composite indicators and ranking systems (see http://ipsc.jrc.ec.europa.eu/).

Table 7

Final rankings (including Prague)

	Cluster	EW	PCA/FA		Cluster	EW	PCA/FA
Hlavní město Praha	C7	1	1	Most	C2	40	55
Brno-město	C6	2	2	Ústí nad Orlicí	C3	41	37
Ústí nad Labem	C2	3	8	Sokolov	C2	42	40
Praha-západ	C5	4	3	Náchod	C3	43	29
Nový Jičín	C1	5	6	Tachov	C4	44	50
Zlín	C1	6	5	Blansko	C3	45	54
Karviná	C2	7	12	Znojmo	C1	46	43
Plzeň-město	C3	8	7	Trutnov	C3	47	36
Olomouc	C1	9	10	Prostějov	C1	48	56
Hradec Králové	C5	10	13	Jihlava	C3	49	41
Liberec	C4	11	4	Kolín	C3	50	58
Břeclav	C1	12	18	Žďár nad Sázavou	C3	51	52
Pardubice	C5	13	14	Rychnov nad Kněžnou	C3	52	47
Česká Lípa	C4	14	19	Bruntál	C4	53	60
Uherské Hradiště	C1	15	20	Jičín	C3	54	51
Praha-východ	C5	16	11	Tábor	C3	55	44
Brno-venkov	C1	17	21	Český Krumlov	C4	56	46
Cheb	C4	18	9	Svitavy	C3	57	62
Ostrava-město	C6	19	45	Semily	C3	58	49
Litoměřice	C3	20	27	Šumperk	C4	59	61
Opava	C1	21	15	Domažlice	C3	60	63
Mladá Boleslav	C5	22	33	Benešov	C3	61	48
Vyškov	C1	23	26	Jindřichův Hradec	C3	62	57
Frýdek-Místek	C4	24	24	Louny	C3	63	66
Kladno	C3	25	35	Havlíčkův Brod	C3	64	65
Přerov	C1	26	30	Rokycany	C3	65	70
Hodonín	C1	27	38	Příbram	C3	66	59
Vsetín	C4	28	31	Třebíč	C3	67	74
Mělník	C3	29	28	Chrudim	C3	68	71
Beroun	C3	30	32	Plzeň-sever	C3	69	69
Kroměříž	C1	31	39	Klatovy	C3	70	64
Děčín	C4	32	25	Prachatice	C4	71	73
Teplice	C2	33	42	Pelhřimov	C3	72	67
Jeseník	C4	34	16	Kutná Hora	C3	73	72
Karlovy Vary	C4	35	22	Písek	C3	74	68
Chomutov	C2	36	53	Plzeň-jih	C5	75	76
České Budějovice	C5	37	23	Rakovník	C3	76	75
Nymburk	C3	38	34	Strakonice	C3	77	77
Jablonec nad Nisou	C4	39	17				

Source: Own calculation.

Note: EW stands for equal weights within the pillars, PCA/FA for weights derived from principal component analysis and factor analysis.

We can see that Prague is at the first place in both cases. This might be a little bit surprising, because capital cities usually perform economically good, do not have so many social problems (low unemployment, high pensions, high life expectancy etc.), but the environmental pillar may not perform so good. Brno as the second biggest Czech city occupies second place. The main surprise for us was the third place of Ústí nad Labem district, because this district is in the long term connected with damaged environment and social problems with higher unemployment. On the other hand, quite a disproportion can be seen in case of Ostrava-město (ranked 19th and 45th), which was supposed to perform rather worse being a structurally affected LAU 1 region. In case of equal weights probably compensability of indicators causes relatively high ranking of this district.

The remaining rankings including Prague (for CI_1, CI_5, CI_9 and CI_10) are introduced in Appendix 2. All rankings (6 CIs) excluding Prague are listed in Appendix 3.

In order to summarize all results we decided to evaluate 77 Czech districts from the point of view of clusters established in section 2. Median seemed to be a suitable indicator for this target properly eliminating outlying values. Table 8 shows the outcomes including and excluding Prague (Cluster 7).

Table 8

Cluster medians

Cluster	Including Prague	Excluding Prague
C1	20.0	19.5
C2	40.0	39.0
C3	55.5	55.0
C4	34.5	34.5
C5	14.5	11.5
C6	10.5	6.0
C7	1.0	x

Source: Own calculation.

The resulting district rankings are not generally unexpected. Districts belonging to the same cluster very often reach similar ranking, i.e. in the overall order they are ranked close to each other. Considering cluster medians, Prague (C7) achieves the first place, followed by big cities (C6), which benefit from the fact that they usually perform well in two pillars (economic and social) having slightly worse results in the third (environmental) pillar. Districts classified to cluster C5 are surroundings of big cities, districts with smaller university cities or prospering industrial branches, so there is no astonishment this cluster is ranked the third. The fourth place take Moravian districts (C1) which have a slightly better performance in environmental pillar, there is lower share of bigger cities and they are more focused on agriculture. Clusters C2 and C4 (both border regions) have almost the same results; their common disadvantage can be seen in the distance from centres of economic performance. The worst result achieved cluster C3. The reason for this may lie in the fact that this cluster covers many diverse regions, i.e. when dividing the indicators into pillars this may play an important role.

Comparing the results when both including and excluding Prague brings almost no differences. Only small changes in values can be noticed in C6 and C5 when Prague is not part of the analysis. The differences are caused by definition of used methods which are endogenous.

Conclusion

As we aimed at the analysis of LAU 1 regions, we had to decide, how to assess sustainability there. Finally we chose useful indicators, though different from indicators used at the national or NUTS 3 level, with data available for all LAU 1 regions. We succeeded in filling all three pillars of sustainable development (economic, social and environmental) with sufficient number of suitable indicators. We found coherences among LAU 1 regions that were affected by similar problems. Cluster analysis was applied and six quite homogeneous clusters were identified (seven when including Prague respectively). After that, all 77 districts were ranked (both including and excluding Prague) according to sustainable development. Several normalisation and aggregation methods were used to compare selected indicators having diverse units of measurement. The results show the ranking of LAU 1 regions in the Czech Republic from the economic, social and environmental point of view (i.e. these three perspectives are aggregated into composite indicator). It was proved that the results obtained from cluster analysis performed in section 2 (all indicators together) correspond with the final rankings based on composite indicators computation (indicators separated into corresponding pillars).

Although we obtained exact rankings, our aim was to assess approximate rankings of the districts. The results should indicate the rough position of particular district among all other districts. Each method gives slightly different results, the suitability should be determined according to the aim of single analysis, i.e. if we want rather to assign the equal weights to all indicators or take into account correlations among indicators. In the same way the compensability when choosing the proper aggregation method should be considered. The proper methodological approach (i.e. proper composite indicator) should be selected according to specific requirements of the analysis.

Finally, it is necessary to add, that statistical approach to sustainable development (analysis of indicators) performed in this paper represents just one possible perspective. It may not (and usually does not) fully correspond with the feelings of people in the regions or with their subjective assessment of quality of their lives (different from sustainable development). Such analysis would exceed this paper, not only due to the necessity of vast time- and cost-demanding research and qualitative analysis of questionnaires or in-depth interviews but also due to uncertain data representativeness.

REFERENCES

Burns, R. B. – Burns, R. A. (2009): Additional advanced chapters, Chapter 23, Cluster Analysis: 552–567. In: *Business Research Methods and Statistics Using SPSS*. Sage Publications Ltd.

Byrch, C. – Kearins, K. – Milne, M. J. – Morgan, R. K. (2009): Sustainable Development: What does it really mean? *University of Auckland Business Review*, 11 (1): 1-7.

Cambridge Econometrics (2013): *Data Definitions*. [online]. Cambridge Econometrics, Cambridge. [cit. 2013-06-24].

<http://www.camecon.com/Europe/Regional_Local_Cities/KnowledgeBase/Appendices/EuRegM/Data_Definitions.aspx>.

Ciegis, R. – Ramanauskiene, J. – Martinkus, B. (2009): The Concept of Sustainable Development and its Use for Sustainability Scenarios. *Engineering Economics,* 62 (2): 28-37.

Czech Statistical Office (2012): *Statistical Yearbook of the Czech Republic 2012*. Český statistický úřad, Praha.

Czech Statistical Office (2010): *Vybrané oblasti udržitelného rozvoje v krajích České republiky 2010 [Selected Areas of Sustainable Development in the Regions of the Czech Republic 2010]*. Český statistický úřad, Praha.

Czesaný, S. (2006): Indikátory udržitelného rozvoje [Sustainable Development Indicators]. *Statistika*, 43 (5): 431-434.

European Commission (2013): *European Common Indicators*. [online]. European Commission, Bruxelles/ Luxembourg. [cit. 2013-06-20]. http://ec.europa.eu/environment/urban/common_indicators.htm.

EUROSTAT (2013): *Sustainable Development Indicators*. [online]. EUROSTAT, Luxembourg. [cit. 2013-06-20]. <http://epp.eurostat.ec.europa.eu/portal/page/portal/sdi/indicators>.

Fischer, J. – Helman, K. – Kramulová, J. – Petkovová, L. – Zeman, J. (2013): Sustainable development indicators at the regional level in the Czech Republic. *Statistika*, 93 (1): 5–18.

Freudenberg, M. (2003): Composite Indicators of Country Performance: A Critical Assessment. *OECD Science, Technology and Industry Working Papers,* 2003/16, OECD Publishing.

Government Council for Sustainable Development – Ministry of the Environment (2009): *Progress Report on the Czech Republic Sustainable Development Strategy. Summary*. Praha.

Government Council for Sustainable Development – Ministry of the Environment (2012): *Progress Report on the Czech Republic's Strategic Framework for Sustainable Development.* Praha.

Hebák, P. – Hustopecký, J. – Pecáková, I. – Plašil, M. – Průša, M. – Řezanková, H. – Vlach, P. – Svobodová, A. (2007): *Vícerozměrné statistické metody [3] [Multi-dimensional Statistical Methods [3]]*. Informatorium, Praha.

Klugman, J. (2011): *Human Development Report 2011. Sustainability and Equity: A Better Future for All.* United Nations Development Programme, New York.

Lengyel, I. – Szakálné Kanó, I. (2012): Competitiveness of Hungarian Urban Micro-regions: Localization Agglomeration Economies and Regional Competitiveness Function. *Regional Statistics*, Special Issue, 52 (2): 27–44.

Macháček, J. (2004): *Ekonomické souvislosti využívání kulturně historických lokalit [Economical Context of Taking Advantage of Cultural Historical Localities]*. Oeconomica, Praha.

Manly, B. F. J. (2004): Multivariate statistical methods: A primer. Chapman and Hall/CRC, London.

Marsden, G. – Kimble, M. – Nellthorp, J. – Kelly, C. (2010): Sustainability Assessment: The Definition Deficit. *International Journal of Sustainable Transportation*, 4 (4): 189-211.

Mederly, P. – Topercer, J. – Nováček, P. (2004): *Indikátory kvality života a udržitelného rozvoje – kvantitativní, vícerozměrný a variantní přístup [Indicators of Quality of Life and Sustainable Development – Quantitative, Multidimensional and Alternative Approach]*. UK FSV CESES, Praha.

Mimmack, G. M. – Mason, S. J. – Galpin, J. S. (2001): Choice of Distance Matrices in Cluster Analysis: Defining Regions. *Journal of Climate*. 14 (12): 2790–2797.

Mooi, E. – Sarstedt, A. (2011): Cluster Analysis: 237-284. In: *A Concise Guide to Market Research.* Springer-Verlag, Berlin Heidelberg.

Ministry of Environment – Czech Statistical Office – CENIA (2006): *Statistical Environmental Yearbook of the Czech Republic 2006*. Praha.

Morrison, D. F. (2005): *Multivariate Statistical Methods.* Thompson Brooks, California.

Nardo, M. – Saisana, M. – Saltelli, A. – Tarantola, S. – Hoffman, H. – Giovannini, E. (2009): *Handbook on constructing composite indicators: methodology and user guide*. OECD, Paris.

Nicoletti, G. – Scarpetta, S. – Boylaud, O. (1999): Summary indicators of product market regulation with an extension to employment protection legislation. *Economic Department Working Papers*. No. 226. OECD, Paris.

Nováček, P. – Mederly, P. (1996): *Strategie udržitelného rozvoje [Sustainable Development Strategy]*. G plus G, Praha.

Odehnal, J. – Sedlaik, M. – Michalek, J. (2012): A Competitiveness Evaluation of the Ukrainian Regions – Empirical Study. *Engineering Economics*. 23 (4): 406–413.

Office for National Statistics (2013): *NUTS: London Directory*. [online]. Office for National Statistics, UK. [cit. 2013-06-24]. <http://www.ons.gov.uk/ons/guide-method/geography/ beginner-s-guide/eurostat/london/index.html>.

Parris, T. M. – Kates, R. W. (2003): Characterizing and measuring sustainable development. *Annual Review of environment and resources*. 28 (1): 559-586.

Paruolo, P. – Saisana, M. – Saltelli, A. (2013): Ratings and rankings: voodoo or science? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 176 (3): 609–634.

Rassafi, A. A. – Poorzahedy, H. – Vaziri, M. (2006): An alternative definition of sustainable development using stability and chaos theories. *Sustainable Development*, 14 (1): 62-71.

StatSoft, Inc. (2011): *Electronic Statistics Textbook*. (Electronic Version). StatSoft. Available at: www.statsoft.com/textbook

Tarantola, S. (2011): *Aggregation rules*. Lecture at JRC Seminar on Composite Indicators and Rankings. Ispra, Italy.

TIMUR. (2012): *Týmová iniciativa pro místní udržitelný rozvoj [Team Initiative for Local Sustainable Development]*. [online]. Timur, Praha. [cit. 2012-09-30]. <http://www.timur.cz/>.

WCED. (1987): *Our Common Future*. Oxford University Press, Oxford.

Acknowledgements

This paper was prepared under the support of the University of Economics in Prague, project No. 11/2012 “Construction and verification of sustainable development indicators in the Czech Republic and its regions”.

Appendix

Appendix 1

77 Czech LAU 1 districts

Source: Ministry of Environment et al. (2006).

Appendix 2

Final results of analysis of 77 Czech LAU 1 regions (including Prague)

Cluster	LAU 1 Region	CI_9	CI_10	CI_1	CI_5
C1	Břeclav	6	3	4	8
C1	Brno-venkov	9	13	9	14
C1	Hodonín	21	29	19	24
C1	Znojmo	31	27	29	25
C1	Uherské Hradiště	18	20	14	17
C1	Olomouc	11	8	11	10
C1	Zlín	10	7	8	6
C1	Vyškov	24	24	20	20
C1	Nový Jičín	12	10	6	5
C1	Kroměříž	33	43	32	39
C1	Opava	25	18	23	16
C1	Přerov	32	38	30	34
C1	Prostějov	52	54	49	49
C2	Ústí nad Labem	5	12	5	11
C2	Karviná	15	21	13	18
C2	Sokolov	42	46	42	51
C2	Teplice	36	53	39	55
C2	Chomutov	35	56	41	60
C2	Most	44	68	43	67
C3	Litoměřice	19	25	16	22
C3	Beroun	23	19	27	21
C3	Mělník	26	22	28	19
C3	Kladno	22	30	26	32
C3	Náchod	41	31	35	28
C3	Blansko	46	55	44	56
C3	Kolín	45	47	48	48
C3	Jindřichův Hradec	55	51	55	50
C3	Ústí nad Orlicí	37	37	36	38
C3	Chrudim	64	64	59	63
C3	Třebíč	62	65	58	64
C3	Žďár nad Sázavou	53	48	50	44
C3	Rychnov nad Kněžnou	48	39	47	40
C3	Nymburk	39	34	37	36
C3	Jičín	54	52	51	47
C3	Plzeň-město	7	6	12	9
C3	Jihlava	50	40	53	42
C3	Rokycany	58	63	62	66
C3	Havlíčkův Brod	66	58	65	58
C3	Rakovník	75	74	74	74
C3	Svitavy	61	61	57	57

(Table continued the next page)

(Continued)

Cluster	LAU 1 Region	CI_9	CI_10	CI_1	CI_5
C3	Benešov	56	44	56	46
C3	Klatovy	71	59	71	59
C3	Domažlice	60	60	61	61
C3	Trutnov	51	41	52	41
C3	Semily	59	49	60	45
C3	Louny	70	75	67	72
C3	Plzeň-sever	63	62	68	69
C3	Kutná Hora	73	70	73	71
C3	Tábor	57	50	63	52
C3	Pelhřimov	74	66	75	68
C3	Příbram	67	57	70	65
C3	Strakonice	77	77	77	77
C3	Písek	76	73	76	75
C4	Česká Lípa	17	23	18	30
C4	Vsetín	30	36	31	33
C4	Frýdek-Místek	27	33	25	31
C4	Český Krumlov	49	42	54	43
C4	Prachatice	68	71	69	73
C4	Liberec	16	4	17	4
C4	Děčín	38	35	34	35
C4	Cheb	20	14	24	13
C4	Tachov	40	45	46	54
C4	Šumperk	69	67	66	62
C4	Jeseník	43	32	38	26
C4	Jablonec nad Nisou	47	28	45	27
C4	Bruntál	65	72	64	70
C4	Karlovy Vary	34	26	40	37
C5	Praha-západ	3	3	3	3
C5	Hradec Králové	8	11	7	12
C5	Praha-východ	4	5	10	7
C5	Pardubice	13	15	15	15
C5	České Budějovice	28	16	33	23
C5	Mladá Boleslav	14	17	21	29
C5	Plzeň-jih	72	76	72	76
C6	Brno-město	2	2	2	2
C6	Ostrava-město	29	69	22	53
C7	Hlavní město Praha	1	1	1	1

Source: Own computation.

Appendix 3

Final results of analysis of 77 Czech LAU 1 regions (excluding Prague)

Cluster	LAU 1 Region	CI_9	CI_10	CI_1	CI_5	CI_2	CI_6
C1	Zlín	5	6	3	3	5	4
C1	Olomouc	7	7	9	5	6	6
C1	Brno-venkov	8	11	6	11	11	15
C1	Břeclav	10	13	4	15	9	16
C1	Nový Jičín	12	12	12	7	8	7
C1	Uherské Hradiště	18	19	15	16	18	19
C1	Hodonín	23	27	20	25	22	36
C1	Opava	25	18	22	20	17	17
C1	Vyškov	28	28	25	26	28	31
C1	Přerov	31	37	29	27	34	33
C1	Znojmo	32	29	32	45	31	44
C1	Kroměříž	34	41	31	30	39	39
C1	Prostějov	51	56	48	48	53	54
C2	Ústí nad Labem	9	14	5	4	13	9
C2	Karviná	15	17	13	9	16	13
C2	Chomutov	33	53	38	35	56	48
C2	Teplice	35	49	35	32	47	41
C2	Sokolov	40	48	41	40	51	43
C2	Most	42	61	43	41	64	58
C3	Plzeň-město	3	3	8	6	4	5
C3	Litoměřice	20	26	17	19	24	26
C3	Kladno	22	24	23	23	29	25
C3	Beroun	26	22	27	31	25	32
C3	Mělník	27	25	28	29	23	29
C3	Ústí nad Orlicí	36	32	34	36	37	35
C3	Náchod	39	30	37	38	30	28
C3	Nymburk	41	39	39	39	38	38
C3	Kolín	43	50	46	50	48	55
C3	Blansko	44	55	42	44	55	52
C3	Jihlava	46	36	50	47	41	40
C3	Rychnov nad Kněžnou	48	42	47	52	42	47
C3	Trutnov	49	38	51	46	40	37
C3	Žďár nad Sázavou	50	45	49	49	43	49
C3	Jičín	53	54	52	53	50	53
C3	Jindřichův Hradec	54	52	54	59	54	57
C3	Benešov	55	46	56	57	46	46
C3	Tábor	56	47	55	51	45	42
C3	Semily	57	51	60	58	52	51
C3	Svitavy	58	60	57	55	57	59
C3	Rokycany	59	67	63	66	66	71

(Table continued the next page)

(Continued)

Cluster	LAU 1 Region	CI_9	CI_10	CI_1	CI_5	CI_2	CI_6
C3	Třebíč	60	65	58	65	61	68
C3	Domažlice	61	62	61	61	63	64
C3	Chrudim	62	64	59	67	62	70
C3	Havlíčkův Brod	63	58	62	62	58	61
C3	Příbram	64	57	67	63	60	56
C3	Plzeň-sever	65	63	69	69	69	67
C3	Klatovy	67	59	70	68	59	60
C3	Louny	70	73	66	64	72	66
C3	Pelhřimov	72	66	73	72	67	65
C3	Kutná Hora	73	70	72	71	70	72
C3	Rakovník	74	74	74	75	74	74
C3	Písek	75	72	75	73	73	69
C3	Strakonice	76	76	76	76	76	75
C4	Liberec	16	4	16	12	3	3
C4	Česká Lípa	17	23	19	17	26	21
C4	Frýdek-Místek	21	21	18	18	19	18
C4	Cheb	24	15	24	22	14	11
C4	Karlovy Vary	29	20	33	34	27	20
C4	Vsetín	30	35	30	28	35	34
C4	Děčín	37	34	36	33	36	27
C4	Tachov	38	44	44	42	49	45
C4	Jeseník	45	40	40	37	33	23
C4	Jablonec nad Nisou	47	31	45	43	32	22
C4	Český Krumlov	52	43	53	56	44	50
C4	Bruntál	66	71	64	54	68	62
C4	Šumperk	68	68	65	60	65	63
C4	Prachatice	69	69	68	70	71	73
C5	Praha-západ	2	2	2	2	2	2
C5	Praha-východ	4	5	10	14	7	8
C5	Hradec Králové	6	8	7	8	10	10
C5	Pardubice	11	10	14	10	12	12
C5	Mladá Boleslav	14	16	21	21	21	24
C5	České Budějovice	19	9	26	24	15	14
C5	Plzeň-jih	71	75	71	74	75	76
C6	Brno-město	1	1	1	1	1	1
C6	Ostrava-město	13	33	11	13	20	30

Source: Own computation.

REGIONAL STATISTICS, 2013, VOL. 3: 141–154 DOI: 10.15196/RS03108

ILDIKÓ KOVÁCS[a] – GÁBOR VALKÓ[b]

Sustainable consumption – consumers' reactions to CSR activities in Hungary

Abstract

The significance of socially responsible consumption as well as the question of the knowledge and information that consumers may have about producers of consumer product are increasingly appearing in the literature. In the case of companies, responsible corporate operation and to examine how information could be transferred to consumers from companies have become key issues especially in the last decade.

Socially responsible consumption, which is the incorporation of social and environmental concerns by individuals in their consumption choices, is growing. The aim of this research is to verify the existence of different profiles of socially conscious consumers and to study their social representation of consumption.

Keywords: sustainable consumption, conscious consumption, Corporate Social Responsibility, consumer segments, degrowth, sustainable development.

Introduction

Promoting corporate social responsibility (CSR) and sustainable consumption are parts of the European Sustainable Development Strategy. There are several programmes aiming at shaping the attitude of consumers for promoting sustainable consumption. Targets of these programmes can be facilitating conscious product choice and frugal consumption. Corporate social responsibility and conscious product choice can have a common beneficial effect towards sustainable consumption.

In our research, we have concentrated on two aspects: first the attitudes that Hungarian consumers have for the activities of socially conscious companies, and second we have examined whether there are separate consumer segments that are receptive to certain areas of CSR.

Sustainability and degrowth

Sustainability has become one of the leading political priorities in the EU recently. An important policy guiding principles of the EU is the involvement of business and social partners and to "enhance the social dialogue, corporate social responsibility and private-public partnerships to foster cooperation and common responsibilities to achieve

a Marketing Institute, Szent István University, 2103 Gödöllő, Páter Károly út 1., Hungary, E-mail: valko.kovacs.ildiko@gmail.com
b Hungarian Central Statistical Office, 1024 Budapest, Keleti Károly út 5-7., Hungary, E-mail: Gabor.Valko@ksh.hu

sustainable consumption and production" (EU 2006). Sustainable consumption and production is one of the seven key challenges in the EU SDS. An objective related to this challenge is to keep development within the carrying capacity of the earth that can be achieved by decoupling of growth from environmental pressures.

Economic growth itself has never been the most important source of human well-being. According to Ayres, technological development is the main driver of development of wealth and also of economic growth (1996). Nowadays, the development of technology mainly serves environment protection and efficient management of resources. The companies do not only protect the environment as a consequence of external regulators, but environmental care has also become a must for companies to remain competitive (Faucheux–Nicolaï 1998).

Limiting economic growth has become a leading aim of environmentalists recently. The representatives of degrowth paradigm desire to halt economic growth in the developed part of the world. The degrowth concept is created by Nicholas Georgescu-Roegen (Georgescu–Roegen 1995), and is further developed by Serge Latouche. Degrowth means the creation of a society where life is better, there is less work and consumption is limited (Latouche 2010). Latouche emphasizes the local scale where solutions should be found for most of the problems. The way to diminishing consumption leads through returning to local economy and the consumption of local products (relocalization). Another important point of the concept is the reduction of energy consumption. Degrowth should not affect the world equally; resources should be redistributed in line with the sustainability concept. Degrowth is defined by Kallis as a „socially sustainable and equitable reduction (and eventually stabilization) of society's throughput" (2011) where throughput refers to materials and energy taken from and waste returned to the environment. The reduction of consumption leads to the use of less energy and less materials as well as the production of less waste which eventually means the decrease of throughput. "Degrowth is impossible to achieve without a turn towards strong sustainable consumption", which is defined as an approach that focuses on appropriate levels and patterns of consumption compared to the weak sustainable consumption, which has the technological improvements in its focus (Lorek–Fuchs 2013).

Sustainability requires not only a change in the production of goods and services but the concept also contains the change in the consumption patterns. The consumption level of developed countries can not be compared to that of the third world countries. There are enormous differences between the consumption of the rich and the poor. About 20% of the world population is responsible for 80% of the total world consumption. Therefore, taking into account the intention of reaching equity in the sustainability concept, it is straightforward that the reduction in the consumption of the rich countries would be the way to achieve sustainability (Málovics et al. 2008).

Corporate social responsibility and consumption

The main idea of the corporate social responsibility (CSR) concept is that there are other roles of the companies in the society beyond manufacturing products, providing services and making profit. These roles include society and environmentally driven actions and commercial activities that increase the well-being of the community (Robins 2005).

However, the companies have to achieve these goals at the same time, one related to profit making and the other to social interests.

The proliferation of corporate social responsibility leads to a cohesive society and a sustainable economic system. Therefore, the European Commission has created a new definition of CSR as "the responsibility of enterprises for their impacts on society" (EU 2011).

The EU also recognized the importance of consumer decisions: „Consumer attention to CSR-related issues has grown in recent years, but significant barriers remain, such as insufficient awareness, the need sometimes to pay a price premium, and lack of easy access to the information necessary for making informed choices. Some enterprises play a pioneering role in helping consumers to make more sustainable choices. The revision of the Sustainable Consumption and Production Action Plan may provide an opportunity to identify new measures to facilitate more responsible consumption" (EU 2011).

In the last decade, due to regulations and market expectations – beside financial performance reports – statements on CSR have appeared in which the companies report on their social and environmental performance. Several researchers agree that CSR investments and attitudes will eventually help the company to perform better economic performance. (Metaxas–Metaxas 2010, Granek–Hassanali 2005, Hall 2000, Rondinelli–Berry 2000).

Several researches argue that the most important stakeholders of the European companies are the employees and so they are the main target group of the CSR activities. Therefore, the CSR activities towards the consumers are of secondary importance and those aiming at the consumers are regarded to be rather PR activities (Dawkins–Lewis 2003).

Doane (2005) argues that CSR is not efficient because the companies imitate the CSR activities of other companies instead of finding there own pattern of CSR. Voluntary reporting of the companies would lead to the recognition of socially conscious companies and it would change the consumption pattern of them. So, the consumers drive the change of businesses to perform in a more sustainable manner. Doene is sceptic in this sense because of the imitation of other companies that makes CSR inefficient.

Socially responsible consumer

Definitions in the literature are not consistent in the content of social responsibility. Some sources argue that only environmentally conscious purchase and social responsibility are related to the concept of social responsible consumption while others say that reducing the volume of consumption should also be part of the responsible consumer behaviour.

The definition of socially responsible consumer and the importance of research in this area came up first in the seventies when Anderson and Cunningham separated the consumers with high social consciousness according to demographic and social-psychological characteristics in 1972. They express that the socially conscious consumers are consumers who consider not only their own satisfaction but they also take into account the social welfare when making purchase decisions.

Roberts (1996) defined the socially responsible consumer as "one who purchases products and services perceived to have a positive (or less negative) influence on the environment or who patronizes businesses that attempt to effect related positive social change". This definition assumes two dimensions: environmental concern and a more general social concern.

Although consumption in general is in itself harmful to the environment, even those who are committed to sustainable consumption recognize that reduction of consumption or additional costs in order to lower the environmental pressure are not likely (Láng 2003).

Sustainable consumption is interpreted to mean consuming less and a kind of alternative or conscious consumption (Jackson 2004). The authors express that welfare does not depend on the volume of consumption. The expenditure of consumers has more than doubled in the UK in the last thirty years, but life-satisfaction does not show a significant change (Donovan et al. 2003). Various previous researches argue that more and more consumers consider "green" and socially conscious consumption important (Vágási 2000, Pakainé Kováts–Herczeg 1999, Borsi 1997).

Mohr et al. (2001) defined socially responsible consumer behaviour based on the concept of CSR. An approach to define CSR involves an attempt to list the major responsibilities of companies. According to Pepper et al., the pillars of sustainable consumption are as follows: pro environmental, pro social, and frugal (2009). Other researchers (McDonald et al. 2006) also argue the decrease of consumption and the „frugal lifestyle" (Lastoviczka et al. 1999). Webb et al. (2008) distinguish between three possible dimensions of socially responsible consumption: (1) purchases based on the corporate social responsibility activities of the companies, (2) recycling, (3) avoiding and reducing products harmful to the environment. Based on these dimensions, the Socially Responsible Purchase and Disposal (SRPD) scale has been developed. This scale measures four dimensions of responsible purchase: 1) influence of the companies' CSR performance on the purchases, 2) recycling activity of the consumers, 3) beside the traditional procurement criteria (price, availability, quality), other concerns related to responsibility emerge (e.g. environmental issues), 4) purchase criteria based on the environmental effects of the products.

Several researches argue that there is a gap between the attitude and behaviour and also between the values and actions (Young et al. 2010, Spaargaren–Koppen 2011, Öbereder et al. 2011). Young et al. claim that the 'attitude–behaviour gap' or 'values–action gap' is present at 30% of consumers who are concerned about environmental issues very much but they do not realize this in their purchases. Companies should have an active role in turning consumers socially conscious. For more sustainable consumption patterns, consumers need new ideas and information. The producers and retailers of products have a responsibility in providing the consumers with information and orientation on the possibilities of green consumption. (Hume 2010)

According to analyses of consumer attitude, there is positive motivation and willingness towards socially responsible companies but the actual consumption is lagging behind. Several researches, that include analyses of both attitude and consumption, have reached the same conclusion (Devinney et al. 2006, Eckhardt et al. 2010). CSR still has a minor affect on consumption decisions (Mohr et al. 2001).

Previous researches on the effect of CSR on purchasing decisions

There are not too many researches in the literature on the effect of CSR on consumer decisions. Several researches reveal that consumers attach more and more importance to the consumption of responsible products and monitoring of CSR activities of the firms. (Carrigan–Attalla 2001, Maignan 2001). Increased attention on CSR has a considerable effect on purchases (Brown–Dacin 1997, Sen–Bhattacharya 2001, Mohr–Webb 2005).

There is a considerable difference between the supply and demand sides of the market. On the supply side, firms are more and more engage themselves in CSR activities while on the demand side, consumers pay more attention to irresponsible corporate behaviour (Snider et al. 2003). Irresponsible corporate actions have a greater impact on consumers' purchases than responsible behaviour (Biehal–Sheinin 2007, Brown–Dacin 1997, Marin–Ruiz 2007, Bhattacharya–Sen 2004).

The survey methodology

The aim of the survey was to analyse the attitude of Hungarian consumers to CSR. The survey was carried out in Hungary on a sample of 510 respondents. The responses were weighted according to regions, types of settlements, age, sex and level of education and therefore are representative for these variables. 11 variables of the research model contained Likert scale questions on consumer opinions about the socially responsible activities of the companies. Based on the survey, latent variables could be created about the description of themes of responsible consumption. The awareness of social responsibility was surveyed by nominal scale while the importance of its areas by ordinal scale. The survey contained the following personal characteristics: sex, age, age group, level of education and residence.

The age of respondents was between 18 and 69 years. The distribution of respondents according to age groups was as follows: 18–29 years (26.1%), 30–39 years (20.4%), 40–49 years (21.0%), over 50 years (32.5%). Bearing in mind the topic of the survey, a core aspect of the selection of respondents was that they should take part in the decisions related to purchase of goods and services. 46.9% of the respondents were men and 53.1% of them are women. Primary school was the highest level of education for 10.2%, vocational training school for 24.7%, secondary school for 40.2% and higher education for 24.3% of the respondents. The place of residence is Budapest for 12.6%, county towns for 17.6%, other towns for 28.3% and villages for 41.4%.

Consumer segments created according to the variables of CSR and their characteristics

According to the responses for the questions related to social responsibility of companies, the respondents have a positive attitude towards the responsible activities of companies (Table 1).

Table 1

Characteristics of the variables

When possible, I by from companies…	Mean	Std. Deviation	Variance	Skewness	Std. Error	Kurtosis	Std. Error
that take care of local products	4.40	0.85	0.72	–1.48	0.11	1.93	0.22
that take care of environment	4.51	0.74	0.54	–1.59	0.11	2.45	0.22
that take care of working conditions and health protection	4.72	0.53	0.28	–2.00	0.11	5.06	0.22
that take care of local people	4.41	0.77	0.60	–1.44	0.11	2.57	0.22
that are fundraiser and supporting	4.28	1.01	1.02	–1.39	0.11	1.28	0.22
that take care of costumer complaints	4.47	0.77	0.59	–1.64	0.11	3.12	0.22
that recycle	4.28	0.99	0.98	–1.41	0.11	1.62	0.22
with responsible behaviour	4.65	0.59	0.35	–1.64	0.11	2.32	0.22
that take care of employees with disabilities	4.27	0.87	0.75	–1.04	0.11	0.63	0.22
that take care of satisfaction of employees	4.47	0.77	0.60	–1.53	0.11	2.41	0.22
that take care of working conditions	4.32	0.76	0.58	–1.03	0.11	1.18	0.22

Source: own elaboration.

The analysis of social responsibility of the companies was carried out by factors of variables. According to Cronbach's alfa and Kolmogorov–Smirnov tests (these tests show the reliability of the scale), the variables were suitable for the conditions of factor analysis. The KMO test showed that the data were suitable for factor analysis (KMO=0.755). According to the Bartlett test, the correlation matrix was significantly different from zero (Sig=0.000). The communality of variables contributes to the explanation of factors at a strong or medium level. The total variance explained by the factors is 74.59%, which is acceptable.

Table 2

Factor structure matrix

	Social	Environmental	Employees	Costumers
Variance explained, %	37.8	15.0	11.6	10.2
When possible, I buy from companies…				
that take care of employees with disabilities	**0.823**	0.166	0.339	0.021
that are fundraiser and supporting	**0.816**	0.317	0.262	0.307
that take care of local people	**0.672**	0.236	0.293	0.467
that take care of local products	0.185	**0.904**	0.312	0.257
that take care of environment	0.397	**0.860**	0.274	0.312
that take care of satisfaction of employees	0.210	0.367	**0.876**	0.172
that take care of working conditions	0.536	0.147	**0.785**	0.221
that take care of costumer complaints	0.242	0.291	0.197	**0.955**

Extraction Method: Principal Component Analysis. Rotation Method: Promax with Kaiser Normalization.
Source: Own elaboration.

Table 2 shows the factor structure. The Social factor has high coefficients in case of companies that take care of employees with disabilities and that are fundraiser and supporting. At the Environmental factor, both variables are important: the companies that take care of environment and of local products. The factor of Employees has high coefficients for the companies that take care of both employees' satisfaction and working conditions. The coefficient of the companies that take care of costumer complaints is important for the Costumer factor. Table 3 presents the correlation matrix between the factors.

Table 3

Component Correlation Matrix

Component	Social	Environmental	Employees	Costumers
Social	1.000	0.268	0.381	0.286
Environmental	0.268	1.000	0.297	0.306
Employees	0.381	0.297	1.000	0.204
Costumers	0.286	0.306	0.204	1.000

Extraction Method: Principal Component Analysis. Rotation Method: Promax with Kaiser Normalization.
Source: own elaboration.

Distinction between the CSR consumer groups by cluster analysis

In our research, we have tried to analyse whether the respondents can be grouped according to their characteristics. For this purpose, the data from factor analysis was used. The cluster analysis was carried out with K-means clustering. As a result, 4 clusters were separated, which are described below.

Cluster centres and the analysis of variance are presented in tables 4 and 5 and in Figure 1. Description of the segments by their demographic characteristics is summarised in tables 6–10.

Cluster 1 – Socially sensitive and urban

Ratio in the sample: 16.7%.

This group mainly relates the social responsibility of the companies with the importance of social aspects. They consider taking care of the working conditions very important. They also consider the two other characteristics, fundraising and supporting the local people very much likeable. The group evaluates environment protection neutral while the satisfaction of employees gets lower scores and the costumer relations higher scores than the average.

Most of the respondents in the group live in Budapest and in large cities; their age is typically over 40 and they have higher education.

Cluster 2 – Environmentalists

Ratio in the sample: 51.5%.

The group considers the manufacturing of environment friendly products (99.3%) and the use of local products (95.3%) essential. 87.1% of the respondents think that it is important to reuse materials. Social concerns are also important and the responsible behaviour with employees and costumers is regarded to be valuable compared to other groups.

The respondents in the group mainly live in Budapest and in other major cities; 59.2% of them are women and the majority has secondary or higher education.

Cluster 3 – Neutrals

Ratio in the sample: 12.1%.

Social responsibility of the companies is considered to be less important in this cluster. The only environmental characteristic that is regarded to be important is the reuse and recycling of materials. Handling of customer complaints is of less or neutral importance for 81% of the respondents in this group.

The respondents in this group are close to the average sample population in terms of age structure. Respondents with secondary education and those living in small towns are overrepresented while there is an equal number of men and women in the cluster.

Cluster 4 – Working conditions in rural areas

Ratio in the sample: 19.7%.

Social concerns are of less importance in this group. Within social concerns, supporting the local people is regarded to be less important. Fundraising and supporting is considered to be neutral or less important for 57.4% which is under the ratio of other clusters. Satisfaction of employees receives the main attention in this cluster.

The typical respondent in this cluster is a man under 40 years with primary or secondary education and lives in a small town.

Table 4

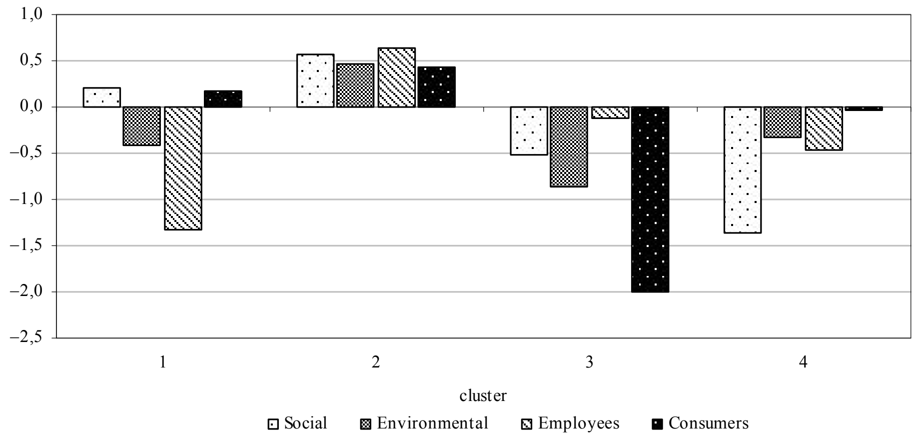
Final Cluster Centres

Component	Cluster 1 Socially sensitive and urban	Cluster 2 Environ-mentalists	Cluster 3 Neutrals	Cluster 4 Working conditions in rural areas
Social	0.214741	0.57107	–0.51098	–1.35947
Environmental	–0.41514	0.465954	–0.86624	–0.33501
Employees	–1.32807	0.636551	–0.12469	–0.46235
Consumers	0.164561	0.429667	–1.99985	–0.03648

Source: Own elaboration.

Figure 1

Cluster centres



Source: Own elaboration.

Table 5

Analysis of variance

Component	Cluster		Error		F	Sig.
	Mean Square	Df	Mean Square	df		
Social	98.034	3	0.430	510	228.239	0.000
Environmental	43.433	3	0.751	510	57.869	0.000
Employees	93.775	3	0.455	510	206.293	0.000
Consumers	99.934	3	0.418	510	238.874	0.000

Source: Own elaboration.

Table 6

Description of clusters by types of settlement

(%)

Types of settlement	Cluster 1 Socially sensitive and urban	Cluster 2 Environ-mentalists	Cluster 3 Neutrals	Cluster 4 Working conditions in rural areas	Total
Budapest	12.9	10.9	8.1	18.8	12.5
County towns	24.7	19.2	11.3	11.9	17.7
Other towns	23.5	21.5	46.8	39.6	28.5
Villages	38.8	48.3	33.9	29.7	41.3
Total	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*

Cramer's V=0.151, sig=0.000.

Source: Own elaboration.

Table 7

Description of clusters by regions

(%)

Region	Cluster 1 Socially sensitive and urban	Cluster 2 Environ-mentalists	Cluster 3 Neutrals	Cluster 4 Working conditions in rural areas	Total
Central Hungary	31.4	24.2	27.9	31.1	27.2
Central Transdanubia	3.5	8.0	13.1	15.5	9.3
Western Transdanubia	8.1	7.2	13.1	4.9	7.6
Southern Transdanubia	10.5	13.6	16.4	18.4	14.4
Northern Hungary	20.9	23.1	13.1	9.7	18.9
Northern Great Plain	5.8	8.7	9.8	8.7	8.4
Southern Great Plain	19.8	15.2	6.6	11.7	14.2
Total	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*

Cramer's V=0,142 p=0,029.
Source: Own elaboration.

Table 8

Description of clusters by sex

(%)

Sex	Cluster 1 Socially sensitive and urban	Cluster 2 Environ-mentalists	Cluster 3 Neutrals	Cluster 4 Working conditions in rural areas	Total
Men	43.0	40.8	50.0	63.7	46.8
Women	57.0	59.2	50.0	36.3	53.2
Total	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*

Cramer's V=0.178, sig=0.001.
Source: Own elaboration.

Table 9

Description of clusters by age

(%)

Age	Cluster 1 Socially sensitive and urban	Cluster 2 Environ-mentalists	Cluster 3 Neutrals	Cluster 4 Working conditions in rural areas	Total
18 – 29 years	19.8	29.2	30.6	47.5	31.4
30 – 39 years	9.3	12.5	17.7	15.8	13.3
40 – 49 years	25.6	17.8	12.9	17.8	18.5
Over 50 years	45.3	40.5	38.7	18.8	36.8
Total	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*

Cramer's V=0.140, sig=0.000
Source: Own elaboration.

Table 10

Description of clusters by education

(%)

Education	Cluster 1 Socially sensitive and urban	Cluster 2 Environ-mentalists	Cluster 3 Neutrals	Cluster 4 Working conditions in rural areas	Total
Primary school	50.0	44.7	27.5	40.0	42.5
Vocational training school	20.9	17.8	33.9	14.0	19.5
Secondary school	20.9	24.6	32.2	33.0	26.5
Higher education	8.1	12.9	6.4	13.0	11.3
Total	*100.0*	*100.0*	*100.0*	*100.0*	*100.0*

Cramer's V=0.133, sig=0.008

Source: Own elaboration.

Validation of the segments by discriminant analysis

In order to validate the segments created by cluster analysis, a canonical discriminant analysis has been carried out. The aim of the analysis was to control if the respondents fall in the same groups. The significance levels of the functions are presented in Table 11.

Table 11

Wilks' Lambda, Chi Square, degree of freedom and significance levels of the discriminant functions

Component	Wilks' Lambda	F	Df1	Df2	Sig.
Social	0.427	228.239	3	510	0.000
Environmental	0.746	57.869	3	510	0.000
Employees	0.452	206.293	3	510	0.000
Customers	0.416	238.874	3	510	0.000

Source: Own elaboration.

The results of the discriminant analysis are summarised in tables 12, 13 and 14.

Table 12

Pearson correlation coefficient matrix

Component		Social	Environmental	Employees	Customers
Correlation	Social	1.000	0.004	0.248	0.048
	Environmental	0.004	1.000	0.015	0.003
	Employees	0.248	0.015	1.000	0.152
	Customers	0.048	0.003	0.152	1.000

Source: own elaboration.

Table 13

Eigenvalues, variances and canonical correlation values of the three discriminant functions

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	2.217	54.376	54.376	0.830
2	1.151	28.244	82.621	0.732
3	0.708	17.379	100.000	0.644

a) First 3 canonical discriminant functions were used in the analysis.

Source: Own elaboration.

Table 14

Wilk's Lambda, Chi-square, degree of freedom and Significance values of the discriminant functions

Test of Function(s)	Wilks' Lambda	Chi-square	Df	Sig.
1 through 3	0.085	1257.920	12	0.000
2 through 3	0.272	662.918	6	0.000
3	0.585	272.762	2	0.000

Source: Own elaboration.

The results of the classification were validated by discriminant analysis, which showed that the regrouping only resulted minor differences compared to those of the cluster analysis. The two classifications resulted the same group for 95.4% of the respondents. The classification according to the cluster analysis was justified; because the two methods gave almost the same results (Table 15).

Table 15

Classification results according to cluster and discriminant analysis

		Cluster 1 Socially sensitive and urban	Cluster 2 Environ-mentalists	Cluster 3 Neutrals	Cluster 4 Working conditions in rural areas	Total
Original						
Count	1	84	0	0	2	86
	2	9	250	4	3	265
	3	2	1	59	1	62
	4	1	0	1	100	101
%	1	97	0	0	3	100
	2	3	94	1	1	100
	3	2	1	95	1	100
	4	1	0	1	98	100
Cross-validated						
Count	1	84	0	0	2	86
	2	9	249	4	4	265
	3	2	2	58	1	62
	4	1	0	1	100	101
%	1	97	0	0	3	100
	2	3	94	1	1	100
	3	2	2	94	1	100
	4	1	0	1	98	100

Source: Own elaboration.

Conclusions

In this research the attitudes related to the CSR activities of the firms was analysed on a representative sample of respondents in Hungary. The value structure of consumers is presented by factor analysis. The four factors are the social, environmental, employees and costumers factors. The consumers were segmented according to these factors and their demographic characteristics. The segmentation was carried out by cluster analysis and the success of the classification was validated by a discriminant analysis.

In our research it is proved that it is possible to separate and describe those consumers who are receptive to certain areas of the CSR activities of companies. Four segments are discriminated: socially sensitive, environmentalists, neutrals and those who find the working conditions the most important. There is generally a positive attitude of the consumers to the socially responsible companies.

Decision makers in the business sphere more and more take into account the attitudes of consumers related to corporate social responsibility of the firms. It is a competitive advantage if a firm can identify consumers likely to respond to socially responsible corporate behaviour.

REFERENCES

Anderson, W. T. JR. – Cunningham, W. H. (1972): The Socially Conscious Consumer, Journal of Marketing, 36, (3) p. 23-31.

Ayres, R. U. (1996): Limits to the Growth Paradigm, Ecological Economics, 19, (2), p. 117-134.

Bhattacharya, C. B. – Sen, S. (2004): Doing Better at Doing Good: When, Why and How Consumers Respond to Corporate Social Initiatives, California Management Review, 47 (Fall) p. 9-25.

Biehal, G. J. – Sheinin, D. A. (2007): The Influence of Corporate Messages on the Product Portfolio, Journal of Marketing, Vol. 71, No. 2, p. 12-25.

Borsi, K. (1997): Zöld stratégiák – Green strategies, Hungarian Journal of Marketing & Menedzsment, 31, p. 66-68.

Brown, T. J. – Dacin, P. A. (1997): The Company and the Product: Corporate Associations and Consumer Product Responses, Journal of Marketing, 61 (January), p. 68-84.

Carrigan, M. – Attalla, A. (2001): The myth of the ethical consumer – do ethics matter in purchase behaviour?, Journal of Consumer Marketing, 18 (7), p. 560-577.

Dawkins, J – Lewis, S. (2003): CSR in stakeholder expectations: And their implication for company strategy, Journal of Business Ethics; May 2003; 44, 2/3; ABI/INFORM Global, p. 185-193.

Devinney, T. M. – Auger, P. – Eckhardt, G. M. – Birtchnell, T. (2006): The Other CSR: Consumer Social Responsibility, Stanford Social Innovation Review, 70, (3), p. 299-326.

Doane, D. (2005): Beyond corporate social responsibility: minnows, mammoths and markets, Futures, 37, p. 215-229.

Donovan, N. – Halpern, D. – Sergeant, R. (2003): Life Satisfaction: The State of Knowledge and Implications for Government, Discussion Paper, Strategy Unit, Government UK psychology', Russell Sage Foundation, New York, p. 3-25.

Eckhardt, G. M. – Belk, R. – Devinney, T. M. (2010): Why don't consumers consume ethically?, Journal of Consumer Behaviour, 9, (6), p. 426-436.

EU (2006): Review of the EU Sustainable Development Strategy (EU SDS) − Renewed Strategy, Council of the European Union, 10917/06, Brussels, 26. 6. 2006.

EU (2011): Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – A Renewed EU Strategy 2011-14 for Corporate Social Responsibility, COM(2011) 681 final, Brussels, 25.10.2011

Faucheux, S. – Nicolaï, I. (1998): Environmental technological change and governance in sustainable development policy, Ecological Economics, 27, (3), p. 243-256.

Georgescu-Roegen, N. (1995): La décroissance, Entropie – Écologie – Économie, Paris, Éditions Sang de la terre, p. 254.
Granek, F. – Hassanali, M. (2005): The Toronto Region Sustainability Program: insights on the adoption of pollution prevention practices by small to medium-sized manufacturers in the Greater Toronto Area (GTA), Journal of Cleaner Production, 14, p. 572-579.
Hall, J. (2000): Environmental supply chain dynamics, Journal of Cleaner Production, 8, p. 455–471.
Hume, M. (2010): Compassion without action: Examining the young consumers consumption and attitude to sustainable consumption, Journal of World Business, 45, (4), October 2010, p. 385-394.
Jackson, T. (2005): Motivating Sustainable Consumption: A review of evidence on consumer behaviour and behavioural change, a report to the Sustainable Development Research Network, January 2005
Kallis, G. (2011): In defence of degrowth, Ecological Economics, 70, (5), p. 873-880.
Láng, I. (2003): A fenntartható fejlődés Johannesburg után – Sustainable development after Johannesburg, Agrinform Kiadóház, p. 147.
Latouche, S. (2010): Farewell to Growth, Polity Press, Cambridge, UK and Malden, MA
Lorek, S. – Fuchs, D (2013): Strong sustainable consumption governance – precondition for a degrowth path?, Journal of Cleaner Production, 38, p. 36-43.
Maignan, I. (2001): Consumers' Perceptions of Corporate Social Responsibilities: A Cross-Cultural Comparison, Journal of Business Ethics, 30 (1), p. 57-72.
Lastovicka, J. – Bettencourt, L. – Hughner, R. – Kuntze, R. (1999): "Lifestyle of the tight and frugal", The Journal of Consumer Research, 26, (1), p. 85-98.
Málovics, Gy. – Nagypál Csigéné, N. – Kraus, S. (2008): The role of corporate social responsibility in strong sustainability, The Journal of Socio-Economics, 37, p. 907-918.
Marin, L. – Ruiz, S. (2007): "I Need You Too!" Corporate Identity Attractiveness for Consumers and the Role of Social Responsibility, Journal of Business Ethics, 71 p. 245-260.
McDonald, S. – Oates, C. J. – Young, C. W. – Hwang, K. (2006): Toward Sustainable Consumption: Researching Voluntary Simplifiers, Psychology and Marketing, 23, (6), p. 515-534.
Metaxas, T. – Metaxas, M. (2010): Corporate Social Responsibility in Europe: Denmark, Hungary and Greece, Journal of Contemporary European Studies, 18 (1), p. 25-46.
Mohr, L. A. – Webb, D. J. (2005): The Effects of Corporate Social Responsibility and Price on Consumer Responses, Journal of Consumer Affairs, 39 (1), p.121-147.
Mohr, L. A. – Webb, D. J. – Harris, K. E. (2001): Do Consumers Expect Companies to be Socially Responsible? The Impact of Corporate Social Resposibility on Buying Behaviour, The Journal of Consumer Affairs, 35, (1), p. 45-72.
Öbereder, M. – Schlegelmilch, B. B. – Gruber, V. (2011): Why Don't Consumers Care About CSR?: A Qualitative Study Exploring the Role of CSR in Consumption Decisions, Empirical Paper, Journal of Business Ethics, 104 (4), p. 449-460.
Pakainé Kováts, J. – Herczeg, J. (1999): Ökotudatos üzleti magatartás – Ecoconscious business behaviour. In: Gidai, E.,Nováky, E., Tóth, A. (Eds.), Magyarország az ezredforduló után. MTA Jövőkutatási Bizottság, Budapest.
Pepper, M. – Jackson, T. – Uzzell, D. (2009): An examination of the values that motivate socially conscious and frugal consumer behaviours, International Journal of Consumer Studies, 33, p. 126-136.
Roberts, J. A. (1996): Will the Real Socially Consumer Please Step Forward?, Business Horizons, 39, (1), p. 79-83.
Robins, F. (2005): The Future of Corporate Social Responsibility, Asian Business & Management, 4, p. 95-115.
Sen, S. – Bhattacharya, C. B. (2001): Does Doing Good Always Lead to Doing Better? Consumer Reactions to Corporate Social Responsibility, Journal of Marketing Research, 38 (2), p. 225-243.
Snider, J. – Paul, R. H. – Martin, D. (2003): Corporate Social Responsibility in the 21st Century: A View from the World's Most Successful Firms, Journal of Business Ethics, 48 (2), p.175-187.
Spaargaren, G. – Kris van Koppen, C. S. A. (2011): Provider Strategies and the Greening of Consumption Practices: Exploring the Role of Companies in Sustainable Consumption, In: The new Middle Classes, Springer Netherlands, p.81-100., 306 p.
Vágási, M. (2000): A fenntartható fogyasztás és a környezettudatos fogyasztói magatartás – Sustainable consumption and environmentally conscious consumer behaviour, Hungarian Journal of Marketing and Management, 34, p. 39–44.
Webb, D. J. – Mohr, L. A. – Harris, K. E. (2008): A Re-examination of Socially Responsible Consumption and its Measurement, Journal of Business Research, 61, (2), p. 91-98.
Young, W. – Hwang, K. – McDonald, S. – Oates, C. J. (2010): Sustainable consumption: green consumer behaviour when purchasing products, Sustainable Development, 18, (1), p. 20-31.