GEOGRAPHY MATTERS: KOHONEN CLASSIFICATION OF DETERMINANTS OF FOREIGN DIRECT INVESTMENT IN TRANSITION ECONOMIES

Joel Deichmann

Abdolreza Eshghi

Dominique Haughton

Selin Sayek[1]

Nicholas Teebagy

Heikki Topi
Bentley College
Waltham, Massachusetts

Abstract

This exploratory study examines the extent to which non-spatial determinants of foreign direct investment (FDI) organize themselves in a manner that mimics the spatial proximity of twenty-five Eurasian transition economies. The Kohonen algorithm is used to create a self-organizing map (SOMs) of a data set that features vectors of twenty-one socioeconomic variables. In this analysis, clusters emerge among the Central European, Balkan, Baltic, and Caucasus/Central Asian regions, leaving Russia as a regional outlier. By introducing SOMs to the discussion of FDI and the factors governing its distribution, we demonstrate an untapped utility in the visualization and analysis of economic data.

Introduction

An interesting development in the transition of Central and Eastern European and Central Asian states (hereafter "transition economies") is the division of the region into groupings by academics and practitioners alike. These sub-regional groupings of countries are made on the basis of divergent initial conditions as well as the extent of political and economic reforms (DeMelo, Denizer, Gelb, and Tenev, 1997). Many of these groupings reflect the spatial proximity of countries (Michalak, 1995), and some of the groups have formally organized themselves as "blocs" in an effort to coordinate their transitions. One example of this is the Visegrád Group (Visegrád Group, 2002), which has initiated

collective cultural, economic, and political efforts highlighted by cooperation in European Union (EU) accession talks.

Academic researchers have also identified geographic clusters of countries that share certain macroeconomic characteristics. For example, Michalak (1995) subdivided the region into six distinct areas: the Visegrád Four (Czech Republic, Poland, Hungary, and Slovakia), the Balkans (former Yugoslavia plus Albania, Romania, Bulgaria, and Moldova), the Baltics (Latvia, Lithuania, and Estonia), Russia, the western ex-Soviet Republics (Belarus and Ukraine), and the southern ex-Republics (the remaining states of South Central Asia and the Caucasus). With the exception of the former Yugoslavia, where fortunes have polarized between Slovenia and Croatia in the north and Bosnia/Serbia in the south, these zones remain coherent in their progress toward market economies and democratic reforms, a pattern that is mimicked by the inflows of FDI. Most interestingly, each of these zones is fairly contiguous in its geography.

Perhaps the most consequential grouping of the transition economies is the one defined by the European Union (EU, 2002), as it provides the blueprint for EU expansion during the first decade of the new millennium. The EU reports that the following countries fulfill the political criteria, and are on target to complete the other requirements for accession in 2004: Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, and Slovenia[2]. The Commission identifies Bulgaria and Romania as the second tier of candidates, eligible to pursue the goal of 2007 membership.

In this paper we attempt to determine the extent to which the distribution of foreign direct investment (FDI) in the region mirrors the geographic clustering of the countries under investigation. We start by removing all geographic references from the analysis, focusing instead on twenty-one non-spatial macroeconomic determinants of FDI. In order to group the transition countries according to their abilities to attract FDI, we employ a Kohonen/Self-Organizing Map (SOM) clustering algorithm. The resulting map shows that the distribution of the determinants of FDI, as well as FDI itself, closely mirrors the geographic clustering of the countries in the region. This result suggests a strong link between geography, FDI inflows, and economic and social conditions of the recipient country, without depicting specific causal relationships.

The manner by which the countries are arranged on the self-organizing map in this study is of great interest, with implications for policy and our understanding of differential levels of progress among the transition economies. In addition, the results of this study provide decision makers at multinational enterprises with useful insights regarding the nature of the included countries as potential investment targets and their relative similarity based on a large number of socio-economic variables. Moreover, this study provides a novel approach to the analysis of foreign direct investment in Central and Eastern Europe and the former Soviet Union by clustering these countries on the basis of their suitability for capital inflows.

An Overview of FDI in Transition Economies

The growth of foreign direct investment (FDI) in the formerly communist states of Eurasia introduces many important research questions by virtue of the recency and magnitude of this growth. The cross-sectional determinants of location choice by foreign firms in transition economies have been widely investigated and modeled using regression and factor analysis (Deichmann, 2001; Tøndel, 2001; Garibaldi, Mora, Sahay, and Zettelmeyer 2001; Deichmann, Eshghi, Haughton, Sayek, and Teebagy, 2003). Countries featuring abundant human capital, superior infrastructure, vast resources, and aggressive reform programs in addition to longstanding economic linkages with the Triad states of North America, the EU, and the Asian Pacific are the early champions in attracting such investment, while those states lacking these characteristics stumble.

The European Bank for Reconstruction and Development (1999, p.79) reports that by 1998, Central and Eastern Europe had accumulated a stock of $57.5 billion in FDI (led by Poland with $15.1 billion), and the former Soviet Republics had attracted $23.7 billion (led by Russia with $8.9 billion). As a share of GDP throughout the region, annual inflows in 1998 represent as little as .4% (Russia) or as much as 2.47% (Azerbaijan), underscoring their varying economic importance to the host countries. The per capita cumulative value of FDI also varies wildly, from $1627 in Hungary to merely $22 in Tajikistan. The EBRD (1999) reports that investment inflows have grown consistently over the past decade, and represent as much as one quarter of GDP as was the case of Azerbaijan in 1998. Table 1 illustrates the importance of FDI in the economies of all transition countries under investigation, and highlights a clear gap between Central and Eastern Europe vis-à-vis the states of the former Soviet Union. In the former group, FDI in 2000 represented an average of 5.72% of GDP, while the corresponding figure for the latter was merely 1.48%.

The potential for FDI to yield positive economic impacts is widely recognized, and as such, the sustained growth of investment over the past decade has generally been viewed favorably by recipient countries. In other words, the perceived benefits of FDI overshadow the costs and lead toward convergence within Europe. Accelerating FDI inflows and evidence toward convergence with Western Europe feed the optimistic economic scenarios for the region's future.

As a corollary to the growth of FDI, the literature on FDI in the transition states continues to grow as well. An important question that prevails in this literature is what macroeconomic factors govern the spatial distribution of investment? Poland, Hungary, the Czech Republic, Slovenia, and the Baltics have dominated the regional investment scene, while most other former Soviet Republics (Russia being the lone exception) have been unsuccessful in attracting foreign capital. Several studies have attempted to explain the uneven distribution of FDI in the region. Deichmann et al. (2003) find that the major factors distinguishing between these countries include professional skills, infrastruc-

ture development, natural resources, favorable investment climate, trade policy, and market reforms. Garibaldi et al. (2001) find that the FDI patterns among the transition economies are best explained in terms of the standard set of economic variables, which include the level of economic reforms, trade liberalization, natural resource endowments, the privatization method, regulations regarding FDI flows and the local institutions. Lansbury et al. (1996) study the factors affecting investment in the Czech Republic, Slovakia, Hungary and Poland from the OECD countries, providing further evidence that the privatization programs, the unit labor costs and the local research intensity govern the FDI flows to the above-mentioned four transition economies.

Table 1
The Magnitude of FDI flows to Transition Economies in 2000, CEE compared to CIS

Central and Eastern Europe				**Former Soviet Republics**			
Country Name	GDP*	FDI	FDI Pct of GDP	Country Name	GDP	FDI	FDI Pct of GDP
Albania	3752	143	3.81%	Armenia	1913	140	7.31%
Bulgaria	11,994	1001	8.35%	Azerbaijan	5266	130	2.47%
Croatia	19,031	926	4.87%	Belarus	29,949	90	0.30%
Czech Rep	50,776	4583	9.03%	Georgia	3029	131	4.32%
Estonia	4969	387	7.79%	Kazakhstan	18,230	1,250	6.86%
Hungary	45,633	1692	3.71%	Kyrgyz Rep	1303	(2.4)	-0.18%
Latvia	7150	407	5.69%	Moldova	1285	128	9.96%
Lithuania	11,313	379	3.35%	Russia	251,105	2,713	1.08%
Macedonia	3573	175	4.91%	Tajikistan	991	24	2.42%
Poland	157,738	9342	5.92%	Turkmenistan	4403	NA	NA
Romania	36,718	1025	2.79%	Ukraine	31,791	595	1.87%
Slovakia	19,120	2052	10.73%	Uzbekistan	7666	100	1.30%
Slovenia	18,128	175	0.97%				
CEE total	**389,901**	**22,289**	**5.72%**	**CIS total**	**356,936**	**5299**	**1.48%**

All figures in US$ millions.
Data Source: World Bank (2002).

The Role of Geography on Economic Factors

The role played by geography in explaining economic policy choices, institutional development and, in turn, in explaining differences in levels of economic development has recently attracted significant attention among economists. The two strands of research focus on two different hypotheses: the "geography hypothesis" argues that geography plays a direct role in determining economic

policies and development, whereas the "institutions hypothesis" argues that institutions instead play a more significant role in the process, with geography playing a possible indirect role.

Gallup, Sachs, and Mellinger (1999) argue for the geography hypothesis, finding that location and climate, through their effects on economic policy choices as well as agricultural productivity, transportation costs, diseases, among other things, significantly affect the income levels and growth rates. Along these lines, in explaining the underdevelopment of the tropics, Sachs (2001) argues that physical geography and climate, which he proxies by the distance from the equator, influences several dimensions that are important determinants of economic development.

In testing the relative significance of geography, Easterly and Levine (2002) find that geography, proxied by tropics, germs and crops, affects economic development through institutions rather than directly. Their study provides more support to the institutions hypothesis, where geography matters for institutional development and institutions matter for economic indicators. Acemoglu, Johnson, and Robinson (2001) distinguish between the role geography and institutions could play in explaining the income differences across countries, arguing that if the former is more significant, the income differences between countries should be persistent. If the latter factor, however, is more significant, the income distribution could change as institutional "reversal" occurs. The authors find supporting evidence for the "institutions hypothesis" by documenting the reversal of relative incomes during the 19th century. The following analysis does not take a position on the relative importance of geography and institutions, but rather allows the data to sort the countries on several dimensions, excluding geography, to see the correlation of this sorting with the geographic location of a country.

Fisher, Sahay, and Végh (1998) undertake a parallel analysis of geography and economic conditions for several transition economies that our analysis focuses on. Although not explicitly testing the geography versus institutions hypothesis, they conclude that economic variables that measure potential and likely timelines for convergence of the Central European countries to EU resemble physical distance from Brussels. They find that economic and physical distances are analogous, and that both affect the rate of convergence within the European continent.

The present paper, along the lines of Fischer et al. (1998), explores the relationship between geographic location and socio-economic indicators. It expands the analysis to include 25 transition economies and a wide range of social and economic indicators, including FDI. It should be noted, however, that the analysis below provides insight into the relative distance, both economic and physical, *within* the group of transition economies. Unlike Fischer et al. (1998), it does not explicitly discuss distance from the EU. The analysis that follows adds, however, to the discussion by employing a novel technique of data visualization in searching for a comparison of physical and economic distance.

Specifically, the analysis employs Kohonen maps to seek structures in a set of 21 macroeconomic indicators to provide better insights on the statistical clusters among countries, and then relates the resulting structures to the physical location of each country. We describe the methodology in more detail below.

Kohonen Self-Organizing Maps

A Self-Organizing Map (SOM) is an exploratory data analysis technique that projects a multi-dimensional data set onto a space with a small dimension (typically a two-dimensional plane). An SOM thus allows for a convenient visualization of groups in the data set that share similar characteristics.

The most commonly used algorithm for constructing SOMs is due to the work of Kohonen. Kohonen maps are known to be "topology preserving", so that observations with similar multi-dimensional vectors of variables find themselves positioned close to each other on the map. In addition a Kohonen map is "self-organizing", so that the variables tend to vary along the map in a meaningful way.

Kohonen (2001) provides a comprehensive overview of SOM methods and case studies. The key demonstration of Kohonen's technique is detailed in his extensive comparison with other methods of visualization and summarized by Kaski and Kohonen (1996). Kohonen first used SOMs to build a biological model of the brain, reflecting the behavior of neural cells (Kohonen 1982, 1984). Since the introduction of SOMs by Kohonen (1982, 1984) in the 1980s, several researchers have contributed to the technique's entry into the mainstream of disciplines ranging from statistics to economics and finance (Deboeck, 1998; Oja and Kaski, 1999).

The Kohonen algorithm has also attracted the attention of probabilists concerned with establishing rigorous proofs of some of its properties. Among such theoretical advances, we note the work of Fort and Pages (1996), who provide a very useful summary of current results on one-dimensional Kohonen maps, such as convergence and self-organizing properties of the algorithm, and extend these results to the case of two-dimensional maps.

Burton and Faris (1996) show that in the algorithm "all memory of the random initial state is lost" eventually, in some rigorous sense explicited in detail in the paper, and thus "the environment alone determines future history". Of course in finite time it is possible that different maps might arise from different initial random states, an issue touched upon by Varfis and Versino (1992).

Deboeck (1998) provides a brief but instructive discussion of SOM applications in finance, economics, and marketing. The author supports this discussion with a financial data application of how mutual fund maps can be created and interpreted using the same grayscale lattice specification chosen for the present analysis. Deboeck stops short of providing substantive analysis of his dataset, but makes constructive suggestions for wider study and supplicates interdisciplinary exploration of the technique.

Varfis and Versino (1992) compare SOMs to the well-established techniques of principal component analysis and hierarchical clustering; sequentially mining the same data set with each technique, they find that Kohonen maps are a worthy alternative to the older techniques. In their Kohonen map of European Statistical Territorial Units (given a set of socioeconomic data), the authors discover that statistical units on the map cluster into distinct European geographic areas. The paper focuses mainly on the techniques themselves, and, unfortunately, the authors do not go into detail on the clustering of specific observations.

For the purpose of the present study, the most instructive application of Kohonen maps was made by Kaski and Kohonen (1996) in an effort to identify structures of welfare and poverty among countries of the world. Using 39 welfare indicators, the authors map the countries on a two-dimensional grid. The self-organizing property of their map is clearly evidenced by the fact that the horizontal axis seems to indicate wealth, and the countries become poorer as one moves from left to right on the map. Among the remarkable findings of this article is the fact that although no geographical information was entered into the computations, the organization of the countries on Kaski and Kohonen's map approximates their geographical organization.

Ponthieux and Cottrell (2001a) use the Kohonen algorithm to group households according to living conditions, focusing on the technique's topology conservation property. Using the French section of the European Households Panel, they find that living conditions tend to deteriorate as family size increases, as well as in cases where adults live alone.

In an expansion of their previous work (2001a), Ponthieux and Cottrell (2001b) use Kohonen maps both to identify groups of variables (characteristics of living standards) and to perform a series of classifications of households. First, the authors identify a group of households lacking in both comfort level and in the most basic consumer goods. Second, the authors conclude that living conditions are best thought of as a complex system of factors rather than simply "levels" that can be defined quantitatively by such variables as income. They illustrate this argument through the inclusion of detailed frequencies of responses combined with the Kohonen maps, elucidating the complexity that is often overlooked when classifying households according to simple categories.

Methods and Data

A self-organizing map (SOM) is a special case of a neural network that simultaneously summarizes a set of variables and clusters observations. It can be viewed as a principal components analysis combined with a cluster analysis, with both procedures influencing each other in the algorithm.

The algorithm begins by assigning to each position i in a grid an arbitrary (random) vector $m_i(0)$ with as many components as there are variables under study. At each time t the vector of variables $x(t)$ corresponding to one of the

observations updates the current vectors $m_i(t)$ according to the formula $m_i(t+1) = m_i(t) + h_{ci}(t)(x(t) - m_i(t))$, where $c = \arg\min_i(\|x - m_i\|)$ and $h_{ij}(t)$ is a function of t and of the geometric distance on the lattice between position i and position j. Typically $h_{ij} \longrightarrow 0$ with increasing distance between i and j and increasing time. So the vector $x(t)$ is allowed to update the vector $m_c(t)$ it is closest to as well as some neighboring vectors $m_i(t)$. When the algorithm converges, the m_i tend to be ordered along the lattice in a meaningful way.

SOMs are superior to typical clustering techniques in that they preserve the topology of the initial high-dimensional data set by locating data points which have similar vectors of variables near one another on the map. Following Kaski and Kohonen (1996), our map is computed to illustrate the structures in a complex set of socioeconomic variables. Here, we use a data set resembling one collected and analyzed elsewhere (Deichmann et al., 2003) in a study of determinants of foreign direct investment in the same 25 countries.[3] In the present study, the measure of FDI is the value of FDI[4] per capita, averaged over 1993-98.

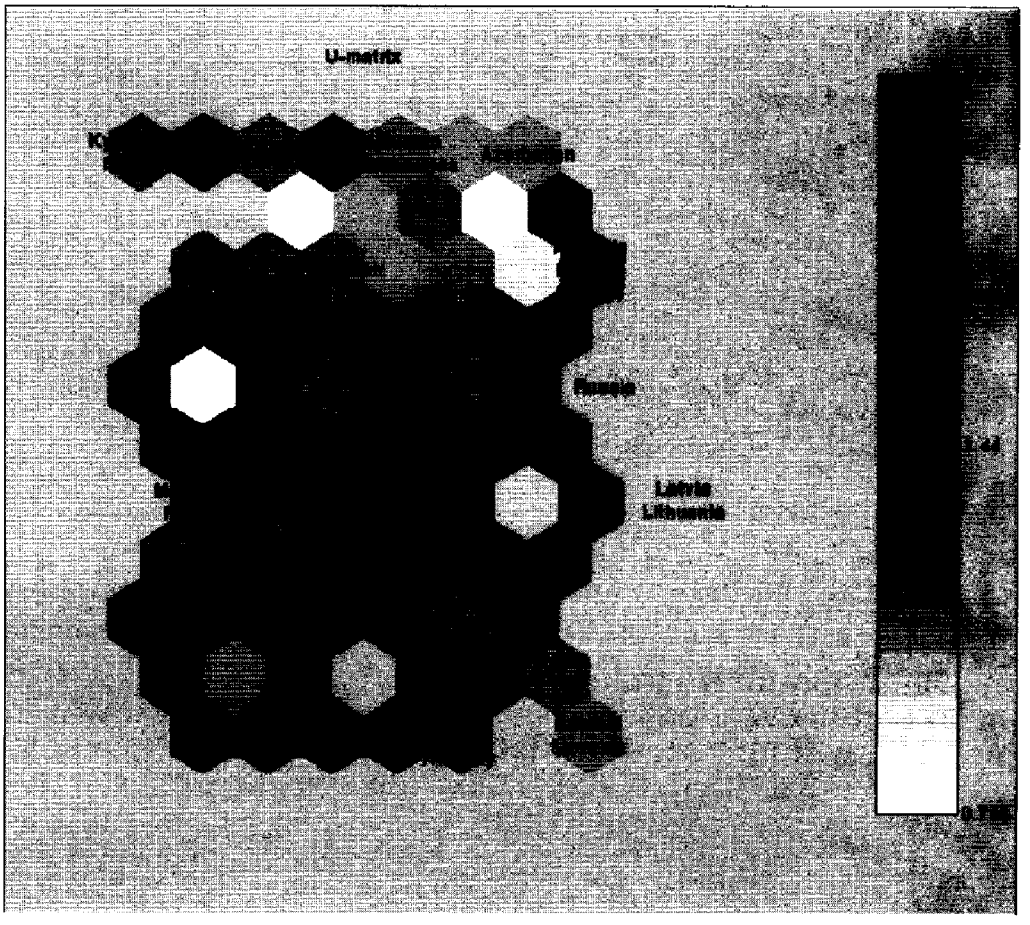
The data set includes 22 variables including FDI and is assembled from and justified by several other studies: DeMelo et al. (1997), EBRD (1999); Kaufmann, Kraay, Zoido-Lobaton (1999); and the World Bank (1999, 2000). This inquiry differs from these investigations and Deichmann et al. (2003) in that its purpose is to identify groups of countries that share a similar set of characteristics on the basis of their location and relative proximity on a two-dimensional hexagonal lattice.

Analysis

The Kohonen analysis yields several plots, a U-matrix (Figure 1), and component plots (Figure 3). Each country on the U-matrix is positioned in that cell of a 4 x 6 grid whose estimated vector is closest to its own vector of indicators. Intervening cells represent the distances between adjoining cells; for example, the mid-grey cell between Kyrgyzstan/Moldova and Kazakhstan represents the distance between these two cells. Shades of grey on the 4 x 6 cells represent the average distance between each cell and its neighbors; for example the dark grey for Russia (with its label to the right for readability) represents the average distance between Russia's cell and its neighboring cells on the 4 x 6 grid. Dark cells indicate large distances or "walls" between observations, while light cells indicate close proximity between observations. The bars to the right of the map provide the values estimated by the Kohonen algorithm that correspond to the shade assigned to each of the countries.

Perhaps the most outstanding feature of this graphic is the nature of clustering of the specific countries on the map, clearly reflecting spatial patterns in the vectors of non-spatial variables (see Figure 2 for geographic locations). The distance between Central Asian republics vis-à-vis those in Central (Czech Republic, Hungary, Slovenia) and Northern Europe (Latvia, Lithuania, Estonia)

Figure 1
U-matrix of vectors based upon 21 variables and non-intervening FDI



is clearly indicated by the distance between cells (as indicated by scales of gray). Following work by Kaski and Kohonen (1996), no geographical data were entered into the model, but remarkably, the countries organized themselves in a manner that mimics their relative geographical locations. Importantly, geographical neighbors and adjacent country pairs are often located on identical cells in the U-matrix, revealing very similar conditions. Such pairs include Tajikistan and Turkmenistan, Belarus and Ukraine, and the Czech Republic and Hungary.

Following a careful inspection of Figure 1, five discrete and relatively homogeneous clusters can be identified. First, and in agreement with mainstream conceptualization of the region, Group I includes the regional leaders of Estonia, Slovenia, Czech Republic, Hungary, and Slovakia. These countries approximate the grouping of Visegrád states; all enjoyed strong linkages to Western Europe prior to World War II. In addition, Croatia, with small distances to both the Slovakia and Czech R./Hungary cells, can be considered as part of this

Figure 2
Selimitations of Area Under Investigarion

group. In this context, it is interesting to note that Croatia quickly turned its back on the complex difficulties with Serbia and re-oriented itself toward the west, creating democratic governments and re-establishing economic linkages with Italy, Austria, and Hungary during the early 1990s.

While Poland's position outside the group and Slovakia's favored relative position may appear surprising, it should be remembered that most of the variables under consideration here are economic — rather than political — in nature. Poland's *democratic* transition has driven its *economic* progress leaving considerable economic stagnation in the isolated eastern reaches of the country, while the opposite holds true for Slovakia. Estonia's membership in this group also merits some attention because its geographical neighbors are found in other clusters. Estonia enjoys deep and longstanding cultural and economic relationships with its neighbor Finland (Huang, 1999), and these appear to play out in Figure 1.

Group II includes Poland, Albania, Bulgaria, Romania, and Macedonia. Although it features a high level of heterogeneity, this group roughly approximates the membership of Michalak's (1995) "Balkan" region, which is characterized by a slow but steady transformation and widely viewed as a pool for second-round EU expansion (Barnard, 2000; EU, 2002). However, our map reveals several exceptions to a uniform Balkan region. As noted earlier, Croatia tends to cluster with Group I. An additional caveat to representing Group II as a Balkan region is Poland, where enduring agrarianism, deficient infrastructure, and lagging human capital contribute to plausible separation from Group I. Finally, Albania's membership in this group may surprise some observers. From a very

slow starting point in 1990 when its authoritarian government ended, Albania has made great strides in terms of reforms and growth, areas that can be examined more closely given the maps of specific indicators in Figure 3.

Group III includes Latvia and Lithuania. Toward an explanation of the clear separation between this group and Estonia, it is worthwhile reiterating Estonia's longstanding cultural and trade linkages with Finland (Laakso, 2001). Although both are among the next-round candidates for EU accession, neither Latvia nor Lithuania enjoys the geographic proximity to the EU that Estonia possesses opposite Helsinki on the Gulf of Finland. Moreover, neither of these states has the cultural proximity to any EU member state that Estonians have as part of the Finno-Ugrian language group.

Russia (Group IV) is distinct from Group III by virtue of its superior resource endowments, less successful reforms, higher inflation, and inferior infrastructure. Russia's outlier status is attributable to its geographical magnitude, which spans eleven time zones. Moreover, its internal heterogeneity confounds membership in groups of smaller, homogeneous states.

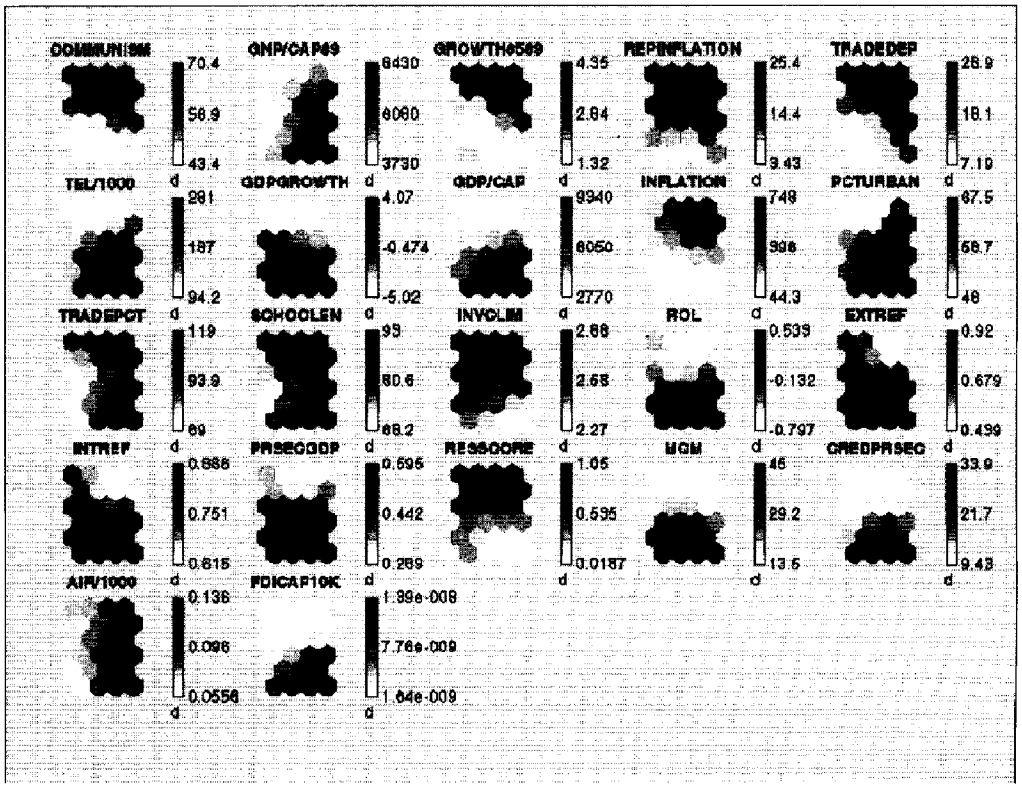
Group V is a relatively homogeneous group in so far as it includes all of the remaining former Soviet Republics. Nearly all of the indicators make it clear that these states remain laggards ten years after the demise of the USSR. There are a few exceptions, such as plentiful resources in the Caucasus and Central Asia, and countries with positive investment climates and far-reaching reforms that have yielded few acquisitions to date, including Moldova and Kyrgyzstan. Certainly, this cluster offers great potential given its substantial resources and reforms. However, as former republics of the USSR these countries suffer from the deepest impacts of the communist legacy, including trade dependency (in particular, remaining linkages with Russia), high inflation, and negligible capital availability.

Figure 3 provides a closer look at the values — estimated by the Kohonen algorithm — of the 21 components corresponding to each of the individual variables (as defined in Appendix) at each position on the 4 x 6 grid.

As a demonstration of Kohonen map interpretation, the number of years under communism ("COMMUNISM") is highest toward the top of the map. Toward an explanation of the scale to the right of each variable, the range in values of COMMUNISM is 43.4 (Soviet liberation and occupation of Central European states after World War II) to 70.4 (Russia, the hub of the Soviet Union itself). Again, the numbers vary slightly from the actual values in the dataset since they are estimated by the Kohonen algorithm. Similarly, GNP per capita prior to 1989 increases as we move along diagonals from the top left to the bottom right. An inspection of our dataset confirms that weak personal earnings in Central Asian Republics predate the razing of the Iron Curtain in 1989 and the demise of the USSR in 1991.

Figure 3 reveals insights on the positions of individual countries in clusters that do not match our expectations. The location on Figure 1 of four countries in particular appears contrary to widely held notions of socioeconomic status in the

Figure 3
The values of the 21 indicators, visualized on the self-organizing map



The acronyms and definitions of the variables are provided in Appendix. The "d" underneath each legend indicates that the variable has been de-normalized for the purpose of visual analysis.

region. These outliers are Slovakia, Poland, Albania, and Russia.

Slovakia's membership in Group I reveal the economic bias of our set of variables. Since the demise of the Czechoslovak Federation in 1993, Slovakia's progress toward democratization has lagged behind those of its Visegrád neighbors, despite declared intentions of simultaneous EU accession (Bernard, 2000; Visegrád, 2002). However, in terms of economic measures Slovakia has fared better than politically. The only *economic* variable clearly separating Slovakia from the rest of the group is trade as a percentage of national income (TRADEPCT). In other respects, Slovakia performed well, and based on our predominantly economic variables it remains clustered with its declared Visegrád allies. It is worth noting that with the defeat of xenophobic and nationalistic Premier Vladimir Meciar in September 2002, the *political* obstacles that separated Slovakia from the other countries in Group I (and stood in the way of NATO and EU membership) have, for all intents and purposes, fallen away (Rebac, 2002).

Leading the region in FDI, Poland is widely heralded as a darling of Western (particularly US and German) investors and an obvious candidate for the next round of EU expansion (Deichmann, 2001; EU 2002). However, it is separated

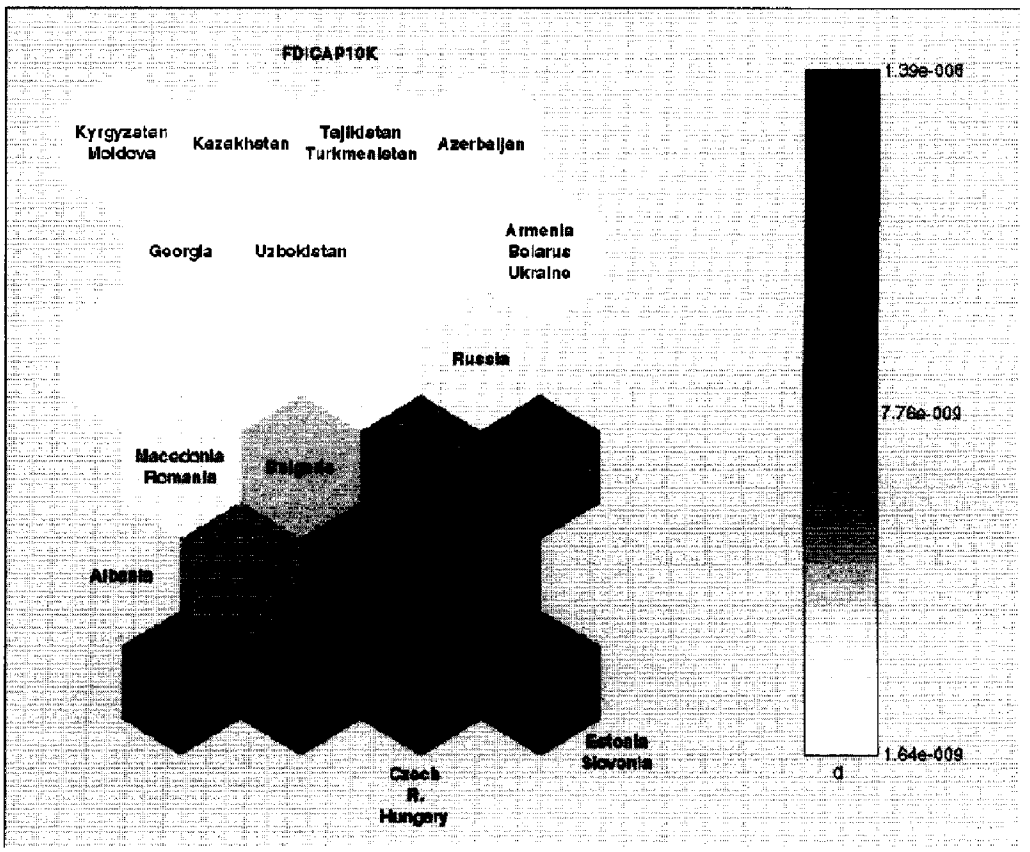
from Group I (Estonia, Slovenia, Czech Republic, Hungary, Slovakia and Croatia) by a distinct wall in the U-matrix. In Figures 1 and 2, Poland appears in the lower left-hand hexagon. An inspection of Figure 3 reveals that Poland differs substantially from Group I on the following variables: GNP/CAP89, TRADEDEP, TEL/1000, TRADEPCT, SCHOOLEN, RESSCORE, and AIR/1000. These variables roughly correspond to broader categories of infrastructure, human capital, and international linkages, all of which have been demonstrated as important considerations for investors (Deichmann, 2001; Deichmann et al. 2003).

In contrast to Poland, Albania's position on the U-matrix might be considered as surprisingly strong by regional experts. Albania has traditionally been viewed as Europe's poorest state, especially under the stranglehold of Enver Hoxha and immediately following his death in 1985 (Vickers and Pettifer, 2000). While corruption, gangsterism, and drug trafficking continue to confound efforts at stabilization, Albania performs fairly well on many of our economic variables, especially those quantifying external and internal reforms and economic growth. In Figure 3, the characteristics appear under EXTREF, INTREF, and GDPGROWTH, respectively.

Figure 3 also sheds light on Russia's position in the U-matrix of Figure 1. Spanning eleven time zones and nearly half of the meridians in the northern hemisphere, Russia possesses unmatched internal diversity within the group of countries included in this investigation and therefore confounds classification as a single country that either is or is not suitable for investment, EU expansion, or any other conceptual category. In spite of its internal heterogeneity, Russia clearly varies from Groups I, II, and III in several indicators. These include Russia's 74 years under Communism, and to a lesser extent rule of law rating (ROL), investment climate (INVCLIM, handicapped by organized crime and the informal sector), relatively shallow financial markets (MQM), and lagging reforms (EXTREF, INTFER, PRSECGDP). Only Russia's vast resources (RESSCORE) and suitable accessibility (AIR/1000) separate it from many of the other former republics of the USSR.[5]

The discussion so far has concentrated on the socio-economic factors that are possible candidates in explaining the FDI distribution among the transition economies. In order to analyze the possible relationship between FDI and these socio-economic factors governing the distribution of FDI, we continue by mapping FDI in the region. Figure 4 shows in detail the enlarged map of FDI per capita, divided by ten billion in order to decrease the variable's variance relative to the other standardized variables. As suggested by Kaski (2002), this is a useful technique for plotting a dependent variable on the map while preventing its intervention in the construction of the U-matrix. Figures 1, 2 and 3 support the hypothesis that FDI is distributed, as are the socio-economic factors that influence the FDI decisions, so that the Kohonen maps of FDI and its determinants mirror the regional physical map. This finding suggests a strong relationship, without any directional causality, between geography, policies and policy-driven FDI flows.

Figure 4
Structured diagram of foreign direct investment
(enlarged from Figure 3 with countries labeled)



Conclusions and Implications

There is no question that foreign direct investments have been instrumental in accelerating the Schumpeterian "creative destruction" of old structures in emerging markets of the region. They have played a critical role in transferring the much needed technological, marketing and management know-how, increasing the international competitiveness of the host countries and facilitated their access to world markets. (Bod, 1998).

In this paper we not only reaffirm the importance of spatial proximity in determining the distribution of FDI in the region, but we also demonstrate the application of Kohonen Self-Organizing Maps to gain a better understanding of the patterns of FDI in Central and Eastern Europe. More specifically, our Kohonen maps provide an intuitively useful method of visualizing a dataset that is otherwise too amorphous and complex to conceptualize. Most interesting is the fact that while no spatial variables, per se, were entered into the clustering exercise, the relative geographic location of many of the countries "re-appears" in our

maps, and clearly distinguishes between the Central European and Baltic States and the laggard states of the former Soviet Union.

The first cluster includes Slovenia, Estonia, Hungary, the Czech Republic, Slovakia and Croatia. All of these countries, including Slovakia following the elections of September 2002 (Rebac, 2002), are generally viewed as reform-progressive, and are in most cases geographically proximate. Poland is positioned as the "leader" of Group II, which also includes Albania, Romania, Bulgaria, and Macedonia, representing economies that remain relatively more agrarian and struggle to establish the rule of law, carry out economic reforms, and address capital availability. Latvia and Lithuania form a group of their own (Group III), isolated from Group I on the one hand, but also distinct from Russia (Group IV). The westernmost portion of Russia enjoys many favorable conditions, but the rest of the country continues to struggle with the challenges of transition. Group V includes all of the non-Baltic former Soviet Socialist Republics that cluster together with no substantial intervening walls.

These clusters of states are remarkably consistent with the projected sequence of entry into the European Union (EU, 2002), and categories established for the purpose of transition analysis by others (Michalak, 1995). Moreover, these observations are in accordance with geographic patterns identified in other projects undertaken without explicitly adding geographical variables (Varfis and Versino, 1992; Kaski and Kohonen, 1996). It is no surprise that the position of the countries on the maps (see Figure 1, the U-matrix) resembles the relative attractiveness of the countries in transition as indicated by their ability to attract foreign capital (Deichmann, 2001; Deichmann et al., 2003).

In short, the clustering exercise provides two significant results. First, the mapping of FDI mimics the mapping of the "combined economic policy and development" indicators; and second, the positioning of country groups in Kohonen mappings of both FDI and the economic characteristics of the transition economies are representative of the geographic location of the region. These two results suggest that, with no indication of causation, FDI, economic characteristics, and the geographic location of the host country, are interlinked. As discussed below, theses findings have several implications for government policy and MNE investment decision making.

From the policymakers' perspective these findings once again reiterate the importance of improving economic conditions in attracting FDI. Due to the perceived benefits of FDI, which include provision of additional capital, creation of employment, and positive spillovers from the advanced technology and managerial skills inherent in the act of FDI, several transition economy governments have liberalized capital restrictions and even gone further by providing special incentives to foreign firms, to entice them to open shop within their economies'.[6] Our results confirm to the policy-makers that such efforts to attract FDI should include improving their domestic economic conditions. As the Kohonen mapping of 21 variables suggests, these indicators include those related to economic development, infrastructure, macroeconomic stability, and reforms.

The case of Poland provides supporting evidence that economic policies and FDI are related. According to our analysis, despite their geographic proximity, when considering Hungary, the Czech Republic and Poland as possible investment sites MNEs prefer the former two transition economies to Poland. In other words, in our analysis Poland places in Group II, the second most attractive group, because its infrastructure, macroeconomic policies and internal reforms lag behind Hungary and the Czech Republic.

Moreover, our results show that aspects of geography (proximity to export markets, access to open seas, access to EU, cultural and linguistic ties, etc.) play an instrumental role in attracting FDI. This is especially true where MNEs pursue an export market-oriented as opposed to domestic market-oriented strategy. A case in point is Estonia. As stated above, cultural and linguistic ties and convenient access to EU markets have played an important role in placing Estonia in Group I, while its neighbors, Latvia and Lithuania, lag behind in Group III. Nation states blessed by favorable geography can leverage their favorable location to attract MNEs that pursue export-market-oriented strategies on a regional or pan-European basis, strengthening the competitive advantage they already have among the nations in the region.

On the other hand, the challenge for countries suffering from unfavorable geographic locations, such as landlocked nations, is to compensate for their location disadvantage by instituting domestic policies that are particularly hospitable to domestic market-oriented MNEs. Alternatively, these nations can leverage their unique advantages in certain factors of production, such as skilled labor force, to attract MNEs that seek to rationalize their operations on a global basis.

Finally, our research allows for scrutiny of widely accepted groupings of the Central and Eastern European countries such as the multi-tiered sequence for EU accession. The European Commission finalized this list of new-round candidates on 9 October 2002, with a timeline of just over one year for accession (EU, 2002). Remarkably, the first cluster of states identified here (the Czech Republic, Hungary, Slovenia, Estonia, Slovakia and Croatia), with the exception of Croatia, falls completely within the list of most favored states for the first round of states for accession. Moreover, the positioning on the hexagonal lattice of nearly all countries seems to depend upon their distance from Brussels, in agreement with findings by Fischer et al. (1998).

Future Research

After deliberately removing geographical variables, we discover that our clustering of states generally mimics the geographical location of states, leading us to conclude that geography and socioeconomic indicators are highly correlated in the context of the Central and Eastern European countries. However, additional confirmation of the link in other regions of the world is needed before we can reach a definite conclusion.

Second, as the first regional application of Kohonen maps, this paper demonstrates the utility of Kohonen maps in understanding the overall pattern of foreign direct investments. Past research has shown that geographic, economic, cultural, and administrative distances affect different industries in different ways (Ghemawat, 2001). Therefore, an interesting application of Kohonen maps would be to examine the sectoral distribution of FDI in the region and its linkage with geography.

Finally, while Kohonen clusters of economic characteristics of these countries closely mimic their geographic locations, there were some exceptions, notably Poland and Estonia. This raises the possibility that other types of distances such as cultural and administrative may influence country classifications. Therefore, future research should include distance measures in addition to economic variables as explanatory variables in mapping the distribution of FDI using the Kohonen algorithm.

References

Acemoglu, D., Johnson, S., & Robinson, J.A. (2001). Reversal of fortune: geography and the institutions in the making of the modern world income distribution. National Bureau for Economic Research. Working paper, no. 8460.

Barnard, B. (2000). EU enlargement: continent could be unified by the end of the decade. Europe. June.

Bod, P. A. (1998). The social and economic legacies of direct capital inflows: the case of Hungary, in J. Bastian (Ed) The political economy of transition in Central and Eastern Europe: the light(s) at the end of the tunnel. (pp.13-44). Ashgate: Brookfield, U.S.A.

Burton, R. & Faris, W. (1996). A self-organizing cluster process. The Annals of Applied Probability. 6, 1232-1247.

Deboeck, G. (1998). Financial applications of self-organizing maps. American Heuristics Electronic Newsletter. January, 1-7.

Deichmann, J. (2001). Distribution of foreign direct investment among transition economies in Central and Eastern Europe. Post-Soviet Geography and Economics. 42, 2.

Deichmann, J., Eshghi, A., Haughton, D., Sayek, S., & Teebagy, N. (2003). Foreign direct investment in the Eurasian transition states. Eastern European Economics. Forthcoming. 41, 5-34.

DeMelo, M., Denizer, C., Gelb, A., & Tenev, S. (1997). Circumstance and choice: The role of initial conditions and policies in transition economies. Washington D.C.: World Bank.

Easterly, W. & Levine, R. (2002). Tropics, germs, and crops: how endowments influence economic development. NBER. Working paper, no. 9106.

European Bank for Reconstruction and Development. (1999). Transition report: ten years of transition. London: EBRD.

European Union. (2002). Towards the enlarged union: strategy paper and report of the European Commission on the progress toward accession by each of the candidate countries. Internet: http://europa.eu.int/comm/enlargement/report2002/strategy_en.pdf. Accessed 9 October 2002.

Fisher, S., Sahay, R., & Végh, C. (1998). How far is Eastern Europe from Brussels? International Monetary Fund. Working paper no. 98/53.

Fort, J-C. & Pages, G. (1995). On the convergence of the Kohonen algorithm with a general neighborhood function. The Annals of Applied Probability. 5, 1177-1216.

Freudenberg, M. & Lemoine, F. (1999). Central and Eastern European countries in the international division of labour in Europe. CEPII. Working paper nos.1999-05.

Gallup, J.L., Sachs, J., & Mellinger, A.D. (1999). Geography and economic development. Center for International Development (CID). Working paper. no 1.

Garibaldi, P., Mora, N., Sahay, R., & Zettelmeyer, J. (2001). What moves capital to transition economies? IMF Staff Papers, 48, 109-145.

Ghemawat, P. (2001). Distance still matters: the hard reality of global expansion. Harvard Business Review, September, 3-11.

Huang, M. (1999). The amber coast: a nordic freeze? Central Europe Review. 28, 6.

Kaski, S. (1997). Data exploration using self-organizing maps. Mathematics, Computing and Management in Engineering. Finish Academy of Technology Series. 82, 1-57.

Kaski, S. (2002). Email correspondence with Dominique Haughton, co-author. 2 May.

Kaski, S. & Kohonen, T. (1996). Exploratory data analysis by the self-organizing map: structures of welfare and poverty in the world. Proceedings of the third international conference on neural networks in the capital markets, London, England. (pp.498-507). Singapore: World Scientific.

Kaufmann, D., Kraay, A., & Zoido-Lobaton, P. (1999). Governance matters. World Bank Policy Research Paper. No. 2196.

Kohonen, T. (1982). Analysis of a simple self-organizing process. Biological Cybernet. 44, 135-140.

Kohonen, T. (1984). Self organization and associative memory. (3rd ed.). Berlin: Springer-Verlag.

Kohonen, T. (1995). Self organizing maps. Berlin: Springer-Verlag.

Kohonen, T. (2001). Self-organizing maps. (3rd ed.). Berlin: Springer-Verlag.

Laakso, J. (2001). Frequently asked questions about Finno-Ugrian Languages. Internet: http://www.helsinki.fi/~jolaakso/fufaq.html. Accessed 16 October 2002.

Lansbury, M., Pain, N., & Smidkova, K. (1996). Foreign direct investment in Central Europe since 1990: an econometric study. National Institute Economic Review.

Lemoine, F. (1998). Integrating Central and Eastern Europe in the regional trade and production network, CEPII. Working paper.

Michalak, W. (1995). Foreign aid and Eastern Europe in the "new world order." Tijdschrift voor Economische en Sociale Geografie. 86 , 260-277.

Oja, E. & Kaski, S. (1999). Kohonen maps. Amsterdam: Elsevier Science.

Ponthieux, S. & Cottrell, M. (2001a). Neural classification and "traditional" data analysis: an application to households' living conditions. Proceedings of the International Work-Conference on Artificial and Natural Neural Networks (IWANN), Grenada, Spain, June 13-15. 738-745.

Ponthieux, S. & Cottrell, M. (2001b). Living conditions: classification of households using the Kohonen algorithm. European Journal of Economic and Social Systems. 15, 69-84.

Rebac, L. (2002). Pro-EU parties win in Slovakia. The Prague Post. Internet: http://www.praguepost.com/P03/2002/Art/0925/news3.php. Accessed 8 October 2002.

Sachs, J. (2001), Tropical underdevelopment. NBER Working Paper. no. 8119.

Tøndel, L. (2001). Foreign direct investment during transition: determinants and patterns in Central and Eastern Europe and the former Soviet Union. Michelsen Institute. Bergen, Norway.Working paper nos. 2001-9.

Varfis A. & Versino, C. (1992). Clustering of socio-economic data with Kohonen maps. Neural Network World. June, 813-834.

Vickers, M. & Pettifer, J. (2000). Albania: from anarchy to a Balkan identity. New York: New York University Press.

Visegrád Group. (2002). Visegrád Group official homepage. http://www.visegradgroup.org/ Accessed 19 September 2002.

Wheeler, D. & Mody, A. (1992). International investment location decisions: the case of US firms. Journal of International Economics. 33, 57-76.

World Bank. (1999). World Development Report. New York: Oxford University Press.

World Bank. (2002). The 2002 World Development Indicators CD-ROM, Washington, DC: World Bank.

Appendix

Variables (except where otherwise specified, values represent means for years 1993-98)

Variable	Description	Data Source
COMMUNISM	Number of years under central planning	DeMelo et al. 1997
GNP/CAP89	1989 per capita GNP (in 1989 PPP US $)	World Bank 2002
GROWTH8589	GDP growth (annual %, mean 1985-89)	World Bank 2002
REPINFLATION	Repressed Inflation (1987-90)[7]	DeMelo et al. 1997
TRADEDEP	Trade Dependence, 1990 (% of GDP)[8]	DeMelo et al. 1997
TEL/1000	Telephone mainlines per 1,000 people	World Bank 2002
GDPGROWTH	GDP growth (annual %)	World Bank 2002
GDP/CAP	GDP per capita, PPP (current international $)	World Bank 2002
INFLATION	Inflation, consumer prices (annual %)	World Bank 2002
PCTURBAN	Urban population (% of total)	World Bank 2002
TRADEPCT	Trade as a percentage of GDP	World Bank 2002
SCHOOLEN	School enrollment, secondary (% gross)	World Bank 2002
INVCLIM	Aggregate investment climate score	EBRD 1999
ROL	Rule of law (scale of -2.5 to 2.5)[9]	Kaufmann et al. 1999
EXTREF	External Reforms	World Bank 1999
INTREF	Internal Reforms (price reforms)	World Bank 1999
PRSECGDP	Private sector share of GDP	EBRD 1999
RESSCORE	Natural resources (0=poor, 1= moderate, 2=high)	Index using EBRD (1999) categories.
MQM	Money and quasi money (M2) as % of GDP	World Bank 2002
CREDPRSEC	Credit to Private Sector (% of GDP)	World Bank 2002
AIR/1000	Number of air passengers per capita	World Bank 1999
FDICAP	Mean annual FDI value per capita, 1993-98	World Bank 2002

[1] Selin Sayek is affiliated with the International Monetary Fund (IMF). The views expressed in this paper are those of the authors and do not necessarily represent those of the IMF or IMF policy.

[2] Cyprus and Malta are not included in the present analysis because they were not part of the Soviet sphere of influence and are not considered to be among the transition economies.

[3] The explanation of the variables and the data sources are detailed in the Appendix.

[4] Foreign Direct Investment (FDI) is defined as "net inflows of investment to acquire a lasting management interest (10 percent or more of voting stock) in an enterprise operating in an economy other than that of the investor. It is the sum of equity capital flows, reinvestment of earnings, other long-term capital flows, and short-term capital flows as shown in the balance of payments" (World Bank 1999, p.286)

[5] While it is beyond the scope of this paper, it would be worthwhile to employ a similar methodology to have a closer look at Russia's 89 complex internal entities, which include the two autonomous and westernized major cities of St. Petersburg and Moscow, 21 republics, 11 okrugs, 49 provinces (oblasts), and six territories (krays).

[6] According to Hanson (2001), between 1998-2001, 103 countries had offered special tax concessions to foreign corporations to set up production or administrative facilities within their borders.

[7] Repressed Inflation is calculated as percent change in real wage minus the percent change in real GDP over 1987-1990 (DeMelo et al. 1997).

[8] Trade dependence is defined as the ratio between the average of exports and imports and GDP. (DeMelo et al. 1997)

[9] Kaufmann et al. (1999) based these scores on survey responses from non-governmental organizations, commercial risk rating agencies, and think tanks. Higher scores represent better governance.

Joel I. Deichmann is Assistant Professor of Geography in the International Studies Department at Bentley College. He received a Ph.D. from SUNY Buffalo in 1999. His research interests include flows of foreign direct investment, international trade, and tourism, with a regional specialization in Central and Eastern Europe. He teaches regional geography courses as well as geographic information systems.

Abdolreza Eshghi is Associate Professor of Marketing and Director of the Center for Marketing Technology in the Marketing Department at Bentley College. He received his Ph.D. from the University of Illinois in marketing/international business in 1984. Professor Eshghi's current research interests include customer relationship management and marketing strategy in Eastern and Central European countries. His most recent articles have appeared in the Eastern European Economics and Journal of Interactive Marketing.

Dr. Dominique Haughton is Professor of Mathematical Sciences at Bentley College. She received her Ph.D. in mathematics from MIT in 1983. Her research interests include model selection, statistics applied to marketing, data mining and the analysis of living standards surveys. She has published over twenty-five articles in refereed journals and has won prizes for research and for innovation in teaching. She is a co-editor on two books on the analysis of living standards in Vietnam.

Selin Sayek is an Economist at the International Monetary Fund. Prior to joining the IMF she was Assistant Professor at the Department of Economics, Bentley College where she was a member of the Data Analytics Research Team (DART). She received her Ph.D. in Economics from Duke University in 1999, and holds a B.Sc. degree in economics from the Middle East Technical University. Her research focuses on the determinants and effects of international capital flows.

Nicholas Teebagy is Associate Professor of Mathematical Sciences at Bentley College. He previously taught at Northeastern University and Babson College. He received his Ph.D in applied statistics from Boston University. Dr. Teebagy has published in such journals as Decision Sciences, Sankhya, Journal of Interactive Marketing, and Eastern European Economics, and presented papers at various professional meetings. He is the authored of The Math Behind Wall Street.

Heikki Topi is Associate Professor of Computer Information Systems at Bentley College. His current research focuses mostly on usability in Web search and other areas of information systems, but he is also interested in a variety of topics related to the utilization of information technology resources and patterns of global change. His research has been published in several information systems and information science journals, and he is the co-editor of the Auerbach's IS Management Handbook.