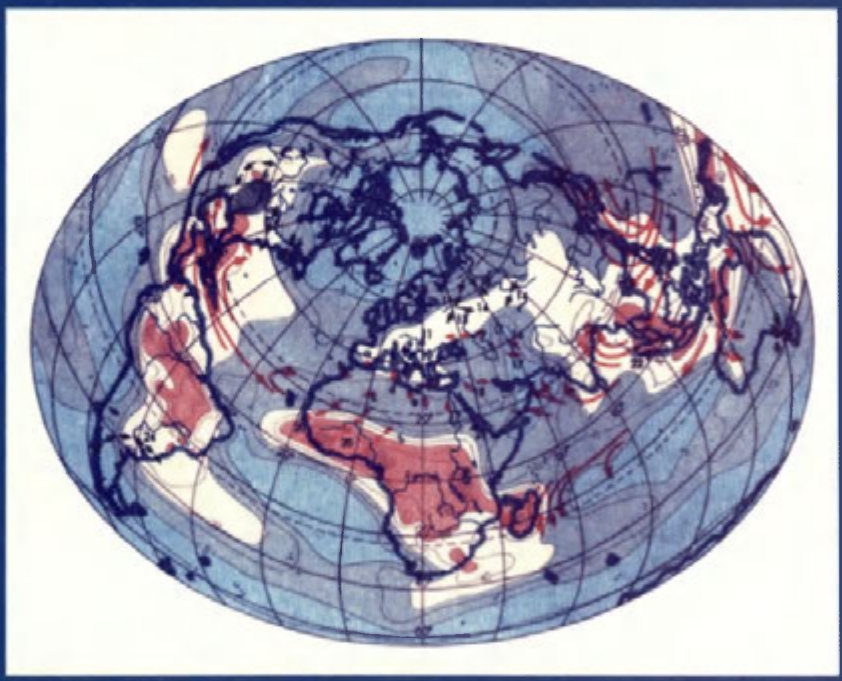
Vol. 73 No. 1 Spring 2000

GEOGRAPHIA POLONICA



POLISH ACADEMY OF SCIENCES
INSTITUTE OF GEOGRAPHY AND SPATIAL ORGANIZATION

COVER: Map of Storms and Winds from Atlas Świata (World Atlas), 1962,
reproduced courtesy of the Publisher: Wydawnictwo Naukowe PWN
(Polish Scientific Publishers – PWN), Warszawa.

Vol. 73 No. 1 Spring 2000

GEOGRAPHIA POLONICA

POLISH ACADEMY OF SCIENCES
INSTITUTE OF GEOGRAPHY AND SPATIAL ORGANIZATION

WARSZAWA

PUBLISHED by the Institute of Geography and Spatial Organization,
Polish Academy of Sciences, Warsaw, Poland

EDITORIAL OFFICE:
Twarda 51/55, 00-818 Warsaw, Poland,
Tel. (48 22) 6978-841, Fax (48 22) 620-62-21

SUBSCRIPTION INFORMATION: Geographia Polonica (ISSN 0016-7282)
is published twice a year in Spring and Autumn by the Institute of Geography
and Spatial Organization, Polish Academy of Sciences

CONTRIBUTIONS and CORRESPONDENCE related to editorial matters should be sent
to the Editor or e-mailed to Ewa Nowosielska: e.nowos@twarda.pan.pl

OPINIONS expressed in individual papers are the sole responsibility of the authors

SUBSCRIPTION ORDERS for Geographia Polonica can be placed with:
ARS POLONA, Krakowskie Przedmieście 7, 00-068 Warsaw, Poland
ORPAN BOOKSHOP, Twarda 51/55, 00-818 Warsaw, Poland
(including single past and current issues)

Information on the content (from 1999) and forthcoming issues are available at:
www.igipz.pan.pl/wydaw/polonica.htm

ISSN 0016-7282

PREPARED FOR PRINT BY
Continuo

Czackiego 46/1, 51-607 Wroclaw, Poland
tel./fax 0048 71 34 390 18 ext. 223

CONTENTS

GEOGRAPHIA POLONICA, 73, 1, Spring 2000
PL ISSN 0016-7282

LABOUR MARKET REFORM AND YOUTH UNEMPLOYMENT IN POLAND

MIKE INGHAM
European Studies Research Institute, University of Salford,
Salford M5 4WT, UK

HILARY INGHAM
Department of Economics, The Management School, Lancaster University,
Lancaster LA1 4YX, UK
e-mail: h.ingham@Lancaster.ac.uk

ABSTRACT: Widespread youth unemployment has been an unwelcome, if predictable, cost of Polish labour market reform. Yet even though young people everywhere are over-represented in the unemployment pool, the burden of the early shock was not borne equally across the country. This paper charts the emergence of the problem and its distribution across space while examining the possibility that the early regional inequalities were transitory. Although there was a trend towards convergence during the upswing in unemployment, it was neither achieved fully, nor is the process continuing. It is, of course, possible that persistent disparities in the seriousness of youth unemployment may be tolerable in an environment in which unemployment generally is now falling. However, the paper urges that recent signs of some improvement in the situation of young people on the labour market should not be taken to imply that active policy interventions have thereby been rendered redundant. The extent of long-term unemployment amongst the young, the demographic trends, and the inevitable future restructuring of large parts of the economy suggest that unemployment may be the lasting legacy of the shock-therapy adopted by the authors of the Polish transformation.

KEYWORDS: Poland, youth, unemployment, convergence, regions.

INTRODUCTION

Pursuit of systemic transformation has been accompanied by many costs, some of which have proved amenable to ameliorative action, others resilient to intervention, still others left unaddressed in any coherent manner at all by policy-makers. Until the intention to undertake a radical overhaul of the education system was announced recently, the problem of the explosion in youth unemployment must be considered to have fallen largely into the third category (Kabaj 1996). However, it would be erroneous to believe that educational reform will of itself prove a panacea for what might be seen as a very serious problem, and it is certain that it will do little for those young people who have already abandoned their scholastic careers. At the same time, it is inappropriate to view youth

joblessness as simply an economic issue; important social and moral consequences could also follow, and these may reverberate for many years into the future. The Polish authorities need look no further than their western neighbours for confirmation of the dangers inherent in allowing large numbers of young people to enter the labour market with little prospect of work (Roberts 1995; Ryan et al. 1991; Freeman and Wise 1982).

None of this is to say that economic revival of itself will not help to contain the problem. Youth unemployment, in common with unemployment in general, has fallen in recent years. However, those aged less than twenty-five, the definition of youth adopted in this paper, continue to have an unemployment rate more than twice that of any other age group, and indications of any increase in their employment are muted. This said, the labour-market prospects of the young have exhibited a marked spatial dimension throughout the current decade. In these circumstances, it is somewhat surprising that there have been few analyses of, as distinct from comments upon, the passage of young Poles into the world of work.[1] The intention in this article is to focus upon the emergence of the problem of youth unemployment in the economy at large and to consider certain important aspects of its regional distribution.

In pursuit of the goals of the paper, the next section will examine the emergence of the youth labour market 'crisis', provide a certain amount of comparative context, and discuss the data series that are available to enable light to be shed upon the problem. This is followed by an examination of the evolution of youth unemployment, as recorded in the official registration statistics, at the level of the voivodship. The evidence relating to the existence of convergence in the unemployment of young persons across spatial labour market is then analysed. The recent reduction in youth unemployment in the light of anticipated future developments in the Polish socio-economic environment follows, while Section V sets out a policy-oriented conclusion closes the paper.

AN INEVITABLE COST OF TRANSITION

Warnings that shock-therapy would bring widespread youth unemployment, unless countered by policy intervention, are to be found throughout the Western world. According to one line of argument, young people enter the labour market as 'outsiders' and, as such, are particularly vulnerable in times of recession and reduced hiring (Lindbeck 1991; Lindbeck and Snower 1988). In addition, in all but the most casual of occupations, young workers have only limited seniority and are therefore susceptible to dismissal under redundancy schemes that offer a certain amount of protection to those employees with the longest service (so--called 'last-in, first-out' (LIFO) policies). There is nothing perverse about firms

[1] An exception is Kowalska (1996).

that engage in such practices, which have solid economic underpinnings, albeit capable of imposing heavy social costs.

As is now legend, open unemployment in Poland began an inexorable rise from the moment that the Balcerowicz Plan was launched formally. Two major sources of information with which to study the behaviour of this phenomenon are available: labour-office unemployment registration data, and the Labour Force Survey (LFS) series. It needs to be stressed that these sources do not measure the same thing; in particular, the definition of unemployment that they apply differs (Reise 1993). Yet, whatever their respective failings, the general picture they paint is broadly similar and each is invaluable for studies of the labour market. A major advantage of the registration data is that they are available monthly from January 1990, though not all currently available disaggregations of the statistics were introduced immediately. Also, information is produced at the levels of both the voivodship and the locality, though only that for the former will be utilised here. The LFS, on the other hand, employs internationally accepted definitions of labour-market status (including unemployment), and provides information on a rich variety of topics concerning the relationship of respondents to the world of work. Less favourably, the survey did not commence until May 1992, is undertaken only quarterly, and is unreliable for disaggregations below the level of the macro-region. In this paper, recourse will be made to data from both registration and LFS sources, with the choice being determined by the question under consideration.

Figure 1 depicts the time-path of national unemployment over the years 1990–1997, both in aggregate and for young people aged 15–24, as derived from both the registration count and the LFS. The Polish experience is somewhat unusual in an international context in that the total for the former has always exceeded that for the latter, although the difference has diminished steadily over time. Both series show unemployment rising to a peak in 1994; in July according to the registration data and in February according to the LFS. Thereafter, the jobless tally trended downwards, slowly at first, but more recently with increased momentum. At its height, aggregate Polish unemployment greatly exceeded of the EU average and was comparable to that prevailing in Finland and Ireland. More recently, however, Poland's unemployment has converged to the EU average (GUS 1998:1).

The impact of the transition-induced recession on young people is startlingly evident. By June 1991 there were over half a million persons under 25 years of age registered as unemployed, a figure which was to exceed one million during three periods in the year' September 1993 – September 1994. While the registration figures do not provide unemployment rates by age, young people consistently formed more than 30% of the total pool over the period examined in this paper.[2] This may be placed in context by noting that those aged between 15 and

[2] The one exception was in March 1997, when the figure dipped to 29%.

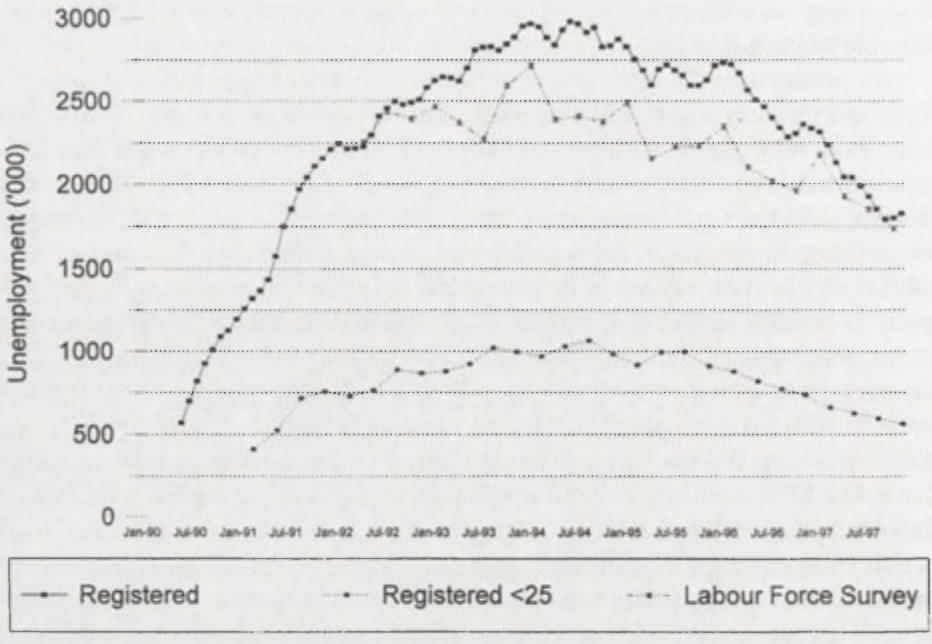


Figure 1. Unemployment in Poland
January 1990 – December 1997.

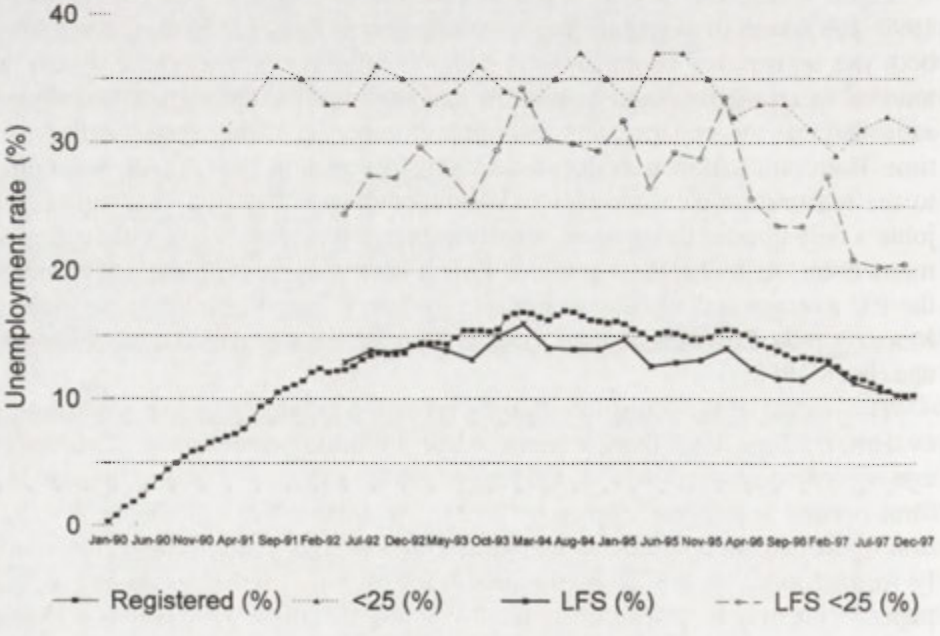


Figure 2. Unemployment rates in Poland
January 1990 – December 1997.

24 typically accounted for little more than 10% of the economically active population during this time. On the other hand, unemployment rates by age are calculable from the LFS and, as shown in Figure 2, the rates for those less than 25 years of age remained in excess of 23% even in November 1997, having been as high as 36% in February 1994.

Beyond its initial explosive growth and recent, more tentative fall, the most obvious characteristic of relative youth unemployment according to its the registration data is marked seasonal pattern. This presumably reflects the pattern of school-leaving. At the same time, the youth unemployment rates calculated from the LFS, while more sensitive to the cycle than the equivalent overall rate, bear a highly significant correlation of 0.909 ($p = 0.00$) with the total figure. As with overall unemployment, however, there is a much richer vein of data to explore, once the figures for youth joblessness are disaggregated spatially, as is drawn out more fully in the following section.

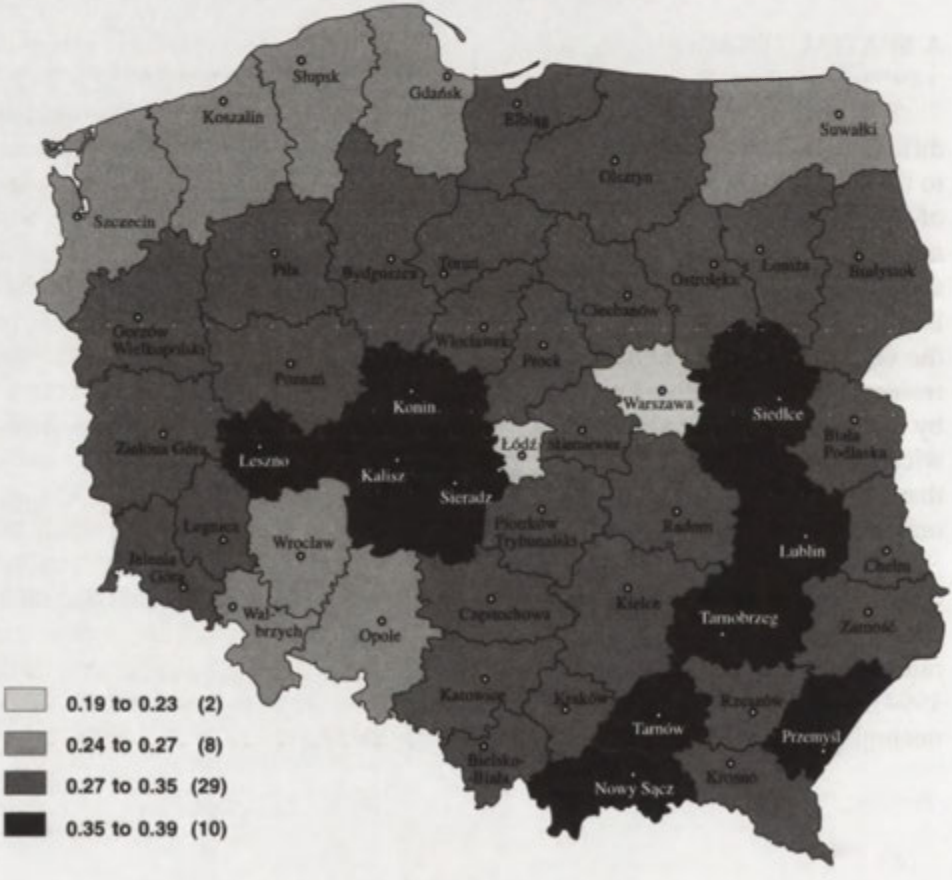
A SPATIAL DISAGGREGATION

The process of transformation has revealed the stark reality of the spatial differentiation inherent in the Polish economy but for so long disguised – at least to the outside world – by the communist authorities. To note but a few examples of the disparities now evident within the country, estimated GDP per capita is more than three times greater in Płock than in Suwałki and Zamość (Saganowska and Śmiłowska 1996); Warsaw continues to attract the lion's share of Foreign Direct Investment (Ingham and Ingham 1997); there are enormous differences in the educational opportunities available to, and attainments of, rural and urban residents (Gorzelak 1996); and the quality of the natural environment, and thereby levels of morbidity and mortality, vary greatly across the country (Węcławowicz 1996). Here though attention focuses upon the unemployment differentials that have emerged in the current decade across the voivodships.[3] Thus, to note only a few examples, by the end of 1990 the rate of registered unemployment in Suwałki was over five times that recorded in Warsaw. At the end of 1997, which coincides with the end of the data period currently under consideration the ratio between the highest (Suwałki) and the lowest (Warsaw) recorded unemployment rates was over 7.5. At the time of maximum registered unemployment (July 1994), this ratio was just under four. As now described, the problem of youth unemployment has a similarly marked geographic dispersion.

[3] As intimated above, even more dramatic differences can be found if the spatial disaggregation of the data is extended.

The Spatial Distribution of Youth Unemployment

Spatial analysis of the age-delimited unemployment impact of the transition is bedevilled by the lack of an appropriate denominator for the registration data. This paper adopts two approaches to overcome the problem. The first is to consider the proportion of unemployment in each voivodship accounted for by those less than 25 years of age ($RYU = U_{25}/U_T$). Simple correlation tests between this measure in March 1992 and December 1997 deliver a Pearson value of 0.846 ($p = 0.00$) and a Spearman of 0.856 ($p = 0.00$). Whatever the root causes behind particular youth unemployment blackspots might be, they would appear to have been persistent through time. In the face of the general stability exhibited by the data, it is assumed that Map 1, depicting the situation in December 1997, will suffice to characterise the comparative situation that has obtained throughout the present decade. It is though worth noting that in Bielsko-Biała, Jelenia Góra,



Map 1. Relative youth unemployment.

Katowice, Legnica, Łódź, and Zielona Góra voivodships the relative situation of young people deteriorated over the period. However, only in Chełm, Konin and Radom voivodships did the absolute number of young unemployed actually increase.

In December 1997, youths comprised 31% of the pool of registered unemployment. Across the 49 voivodships, the average proportion of young people aged less than 25 in the registration stocks was also 0.31, with a standard deviation of 0.04. These latter properties of the data have been used in compiling the grid for Map 1.[4] Two distinct spatial clusterings of voivodships with youth unemployment relative values more than one standard deviation greater than the mean are apparent immediately: one central grouping around Kalisz and one east-south-eastern banana, to borrow a term from Gorzelak (1996), pivoting around Tarnobrzeg. At the other end of the spectrum, only Warsaw and, surprisingly perhaps, Łódź had relative values lying more than two standard deviations below the average.

A reasonable reaction to the distribution depicted in Map 1 might be to assume that it is just one more symptom of multiple labour market deprivation and, as such, a further ill for which the same one cure will suffice. In fact, there is little evidence to support the leading proposition in this hypothesis, with the pair-wise correlations between relative youth unemployment and voivodship unemployment rates, proportions of long-term (more than 12 months) jobless, and unemployment-vacancy ratios not being even remotely significant. However, it is worthy of note that the taxonomy developed by Scarpetta and Huber (1995), which was based on purely structural, as opposed to labour-market, characteristics, identifies as weaker regions eight of the ten voivodships classified as blackspots in Map 1. The two 'errors' are Kalisz and Tarnów, both of which they classify as developed and thereby possessing relatively healthy labour markets.

The foregoing description of the spatial distribution of youth unemployment is useful, insofar as it denotes the likelihood that any unemployed person encountered in a particular voivodship will be less than 25 years of age. However, in a comparative context, the figures are potentially misleading. In particular, the unemployment relative can be the same in a low-unemployment area as in one with a very high overall rate of joblessness. For example, the ratio of youth to total unemployment (after rounding) was 0.30 in Częstochowa, Elbląg, Gorzów Wielkopolski, Radom and Zielona Góra, yet their overall unemployment rates were 8.2%, 18.0%, 12.4%, 14.4% and 12.3%, respectively. Presumably, one would consider that a young person domiciled in Częstochowa faced better life chances than one resident in Elbląg. Similarly, the unemployment rate in Chełm was 10.2%, with young people forming 35% of the total, while, in Wrocław the unemployment rate was 10.1%, with those less than 25 comprising 27% of the

[4] The shift in the overall distribution of youth unemployment relatives is discussed later in the paper.

total stock of those without work. Without the data necessary to compile exact unemployment rates for the young, ambiguity cannot be avoided. However, it is possible to compute an alternative indicator to the simple relative unemployment statistic described in Map 1 when attempting to identify spatial variations in the plight facing the young.

In particular, the relative youth unemployment ratios (RYU) can be standardised by the prevailing voivodship employment rates (defined as the complement of their unemployment rates). That is,

$$SRYU = [RYU/(100 - UR)] *100, \quad (1)$$

where UR is the overall voivodship rate of registered unemployment.

This measure ensures that voivodships with equal values of RYU, but different degrees of overall labour market tightness, are distinguished in a way that appeals to intuition. However, one should not be blind to the consquence that the SYRU's of labour markets with differing degrees of relative youth unemployment are thereby equalised. For example, Gorzów Wielopolski, Jelenia Góra, Olsztyn, Poznań and Suwałki all had SRYUs equal to the national figure of 0.34, but RYUs that ranged from 0.27 in Suwałki to 0.33 in Poznań. At the same time, their overall unemployment rates lay between 3.4% in Poznań and 21.2% in Suwałki. Having injected this caution, one can note that there is again strong temporal stability to the data, with SRYU in December 1997 having both Pearson and Spearman coefficients of 0.85 ($p = 0.00$) for the correlation with the comparable measure for March 1992. Map 2, in which the standard deviation of the data (0.05) is once again used to determine the bands around the sample mean (0.36), is therefore taken as an adequate characterisation of the comparative situation in the present decade.[5]

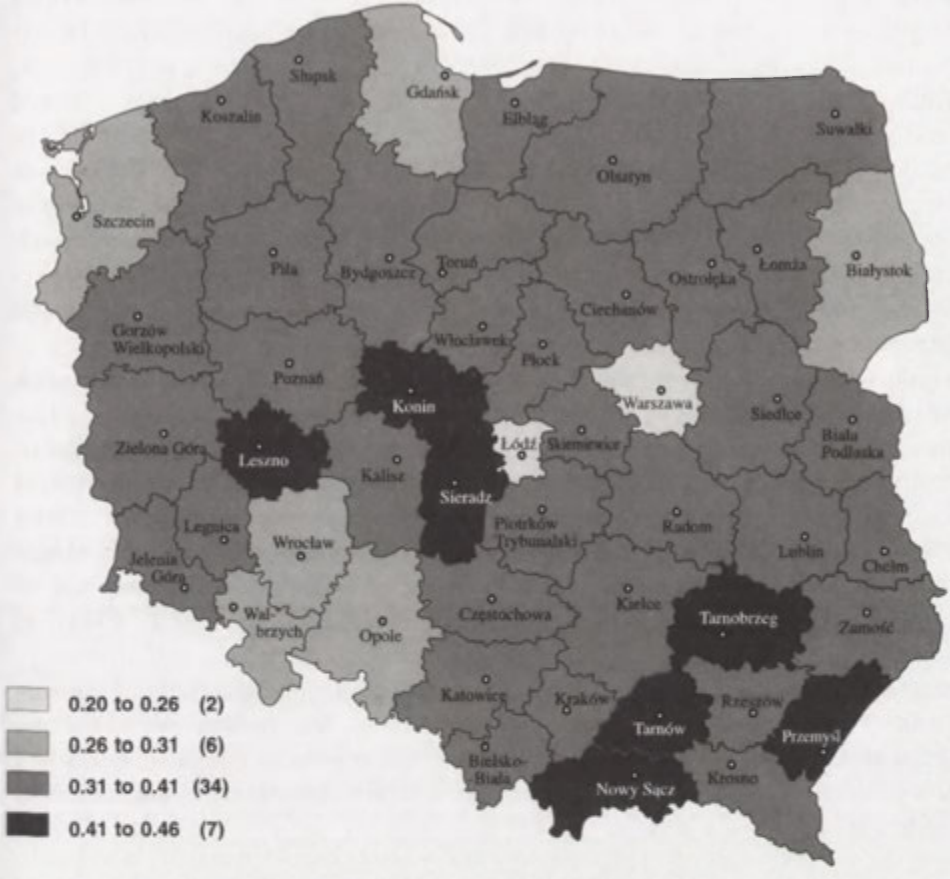
Standardisation leaves unchanged the finding that Łódź and Warsaw, with SRYU's of 0.2 and 0.24 possessed what, on this measure at least, were the most favourable youth labour market environments. Indeed, no other voivodship had a standardised youth unemployment relative less than 0.29. At the opposite end of the spectrum, however, the effect of standardisation was to reduce the number of voivodships with ratios more than one standard deviation greater than the mean. In particular, Lublin, Kalisz and Siedlce regressed towards the average once the correction was made.[6]

CONVERGENCE

The previous discussion has indicated remarkable stability in the spatial distribution of the burden of youth unemployment. In this section, the intention is to

[5] Changes in absolute differences are discussed below.

[6] Kalisz was one of the two voivodships that did not agree with the Scarpetta and Huber (1995) taxonomy when RYU was the measure of youth unemployment.



Map 2. Standardised relative youth unemployment.

provide a rather fuller treatment of the problem by relating it to the recently-emerging body of literature examining questions pertinent to the existence, or otherwise, of tendencies for national economies, and regions within nations, to converge to a common value on a wide variety of economic performance indicators (Baumol 1986; Barro and Sala-I-Martin 1991; Quah 1996). In most instances, the underlying question reduces to one of whether initial laggards (or late-starters) tend to catch up with the early leaders. Typically, as in the case of GDP per head (Durlauf and Quah 1998), there exists at least one theory which suggests that such a process might be expected to occur and one or more pointing to just the opposite conclusion. The same situation obtains in relation to regional differences in Poland's youth unemployment.

In the case of a system of regional labour markets, appeal could be made to the operation of a number of theoretically plausible price and quantity adjust-

ments in generating an equilibrium that is balanced, if perhaps still characterised by 'excessive' overall unemployment. On the one hand, workers might be expected to flow from high-to-low unemployment regions in their search for work, while capital – attracted by the impact of excess supply on wage costs – should tend to flow in the opposite direction. However, in the case of the Polish labour market there are a number of standard objections to such reasoning. The parlous state of the housing market deters worker migration, for example, while the severely under-developed infrastructure in many regions serves to inflate considerably the costs of capital movement to areas of relative labour surplus. Nevertheless, in the case of young workers at least, there might still appear to be grounds for expecting some degree of regional equalisation to be observed.

On the one hand, the housing demands of unmarried young people are much more limited than those of whole families, while arguments that people are less risk-averse when young are commonplace. In any event, the pay-off period to costly initial investments (in terms of either in finance or physical discomfort) is longer. On the other hand, many youngsters have an obvious alternative to labour market participation, insofar as they can prolong their educational careers. While remaining at school and potentially over-investing in human capital may not, of itself, be optimal from a social perspective (Lindley 1996; OECD 1977), in a statistical sense it certainly provides a plausible reason why youth unemployment rates across regions might be observed to converge. The issue of convergence is far from simple (Durlauf and Quah 1998), but the analysis to follow, although addressing by no means all of the theoretical complexities, considers two possible types of convergence that have received widespread discussion in the literature.

δ *Convergence*

Attention focuses first upon the extent (if any) to which the distribution of registered youth unemployment across voivodships and macro-regions has converged over time. While such questions are often addressed by examining the behaviour of the standard deviation of the variable of interest across the units of observation (hence δ *convergence*), various other measures of dispersion are also presented here. In particular, information on the mean and range of the observations, as well as the coefficient of variation (cv = $\delta/mean$), is provided. Also, movements in the dispersion are explored across different time periods to order that account might be taken of both potentially-differing seasonal patterns in youth unemployment and the fact that joblessness has assumed two distinct trends in the current decade.

The first section of Table 1 examines the dispersion of RYU and SRYU over the entire period for which relevant data are available, up to the end of 1997. This gives a series running from March 1992 to December 1997 for which there was an unambiguous reduction in dispersion (on all measures and both definitions of

Table 1. Cross-sectional measures of youth-unemployment dispersion.

	Mean	Range	St. Dev't'n	Coeff Var
		Mar 92/Dec 97		
RYU	0.34/0.31	0.34/0.20	0.06/0.04	0.18/0.13
SRYU	0.39/0.36	0.37/0.26	0.07/0.05	0.18/0.14
		Dec 92/Dec 97		
RYU	0.36/0.31	0.38/0.20	0.07/0.04	0.19/0.13
SRYU	0.42/0.36	0.38/0.26	0.07/0.05	0.17/0.14
		Mar 92/Sept 94		
RYU	0.34/0.38	0.34/0.26	0.06/0.05	0.18/0.13
SRYU	0.39/0.46	0.37/0.30	0.07/0.06	0.18/0.13
		Sept 94/Dec 97		
RYU	0.38/0.31	0.26/0.20	0.05/0.04	0.13/0.13
SRYU	0.46/0.36	0.30/0.26	0.06/0.05	0.13/0.14

Notes: 1. RYU = U_{25}/U_T. 2. SRYU = [RYU/(100 – UR)] *100.
Source: *Bezrobocie rejestrowane w Polsce* (GUS), various years.

youth unemployment). However, because the beginning and end of the foregoing data series occur in different months of the year, it is conceivable that the results might be influenced by seasonal effects that differ across voivodships. This possibility is explored in the second section of the Table, in which the data series runs from December 1992 to December 1997. In the event, the modification is seen to have no substantive impact on the previous findings so the original conclusion is unaffected.

Even having accounted for different beginning and end months of observation, it is still possible to argue that the comparisons are in some sense flawed, because unemployment was first rising but has more recently been downward--trended. In order to consider this possibility, the series are disaggregated to examine the two epochs separately, though even this exercise is not without ambiguity, as all voivodship unemployment rates did not peak in the same month, and neither did youth and aggregate unemployment totals. Furthermore, national unemployment was at a maximum in July 1994, while observations on youth unemployment are available for the ends of quarters only. Nevertheless, inspection of the data indicates that, in the large majority of cases, each of the variables considered here attained peak values between the end of the first and the third quarters of 1994.

The third and fourth sections of the table take September 1994 as the approximate peak and consider movements prior and subsequent to that date, respectively. In the event, imposing this break has little impact on the earlier findings regarding dispersion. The distribution of youth unemployment became progressively more concentrated over the data period. However, note should be made of the failure of the coefficient of variation for both RYU and SRYU to fall beyond September 1994, an issue discussed further below.

β *Convergence*

The concept of β convergence relates to the rate of change of a particular variable through time and, in particular, to whether this is negatively related to the initial value of the variable itself. Investigations of convergence have explored the behaviour of a range of magnitudes both nominal (e.g. interest rates, inflation rates, exchange rates) and real (e.g. income per capita, productivity growth). Also to be included in the latter group is the limited number of studies that have examined the movements in spatial unemployment patterns (e.g. Ingham and Grime 1994). Put formally, if β in the regression

$$\ln(RYU_t/RYU_0) = \alpha + \beta RYU_0 + \mu_t \qquad (2)$$

is negative, this is necessary, although not sufficient (Chatterji and Dewhurst 1996; Chatterji 1992) evidence for the existence of convergence in youth unemployment.[7] In (2), subscript t relates to some terminal observation date and subscript 0 to an initial starting point, while μ_t is a standard error term.

The findings presented in the first rows of Table 2 provide significant evi-

Table 2. Time-series convergence of relative youth unemployment.

	α	β	Adjusted R^2	D.W.
Mar 92/Dec 97				
RYU	0.343	–1.235	0.45	1.73
	(5.05)	(6.31)		
SRYU	0.366	–1.126	0.47	1.69
	(5.33)	(6.57)		
Dec 92/Dec 97				
RYU	0.257	–1.078	0.44	1.75
	(4.04)	(6.17)		
SRYU	0.276	–0.961	0.45	1.65
	(4.23)	(6.30)		
Mar 92/Sept 94				
RYU	0.475	–1.074	0.53	1.71
	(9.41)	(7.38)		
SRYU	0.505	–1.004	0.58	1.63
	(10.33)	(8.23)		
Sept 94/Dec 97				
RYU	–0.105	–0.218	0.02	1.62
	(1.85)	(1.46)		
SRYU	–0.113	–0.160	0.01	1.64
	(1.87)	(1.24)		

[7] Of course, the same conclusion holds for an equivalent regression involving SRYU, and both will be examined in what follows.

dence that convergence did in fact take place, over both the whole of the period 1992–1997 and the early sub-period during which unemployment was continuing to rise. Furthermore, the estimated β coefficients satisfy the equivalent of Chatterji's (1992) criterion for strong convergence; that is, for convergence to an equalised steady-state value of relative youth unemployment across voivodships. This point is returned to below. However, beyond the unemployment peak in 1994, the relationship appears to break down, with the estimated values of β falling well short of conventional statistical significance levels.

Two standard questions arise in the analysis of such difference equations as those described in Table 2. The first concerns the existence of equilibria (at which, for example, $RYU_t = RYU_{t-1} = RYU^*$), while the second focuses upon their stability. In this instance, the non-linear equation (z) has two roots; the origin, or $RYU^* = 0$, and $RYU^* = -\alpha/\beta$. The character of the equilibria described by these roots depends upon the values of the parameters α and β. Generalising from the RYU, SRYU notation to simply y, replacing subscripts t and 0 by t and t–1, and assuming $\alpha > 0$, $-1/RYU^* < \beta < 0$, the relevant phase diagram is as depicted in Figure 3. Within this particular class of cases, which covers the first

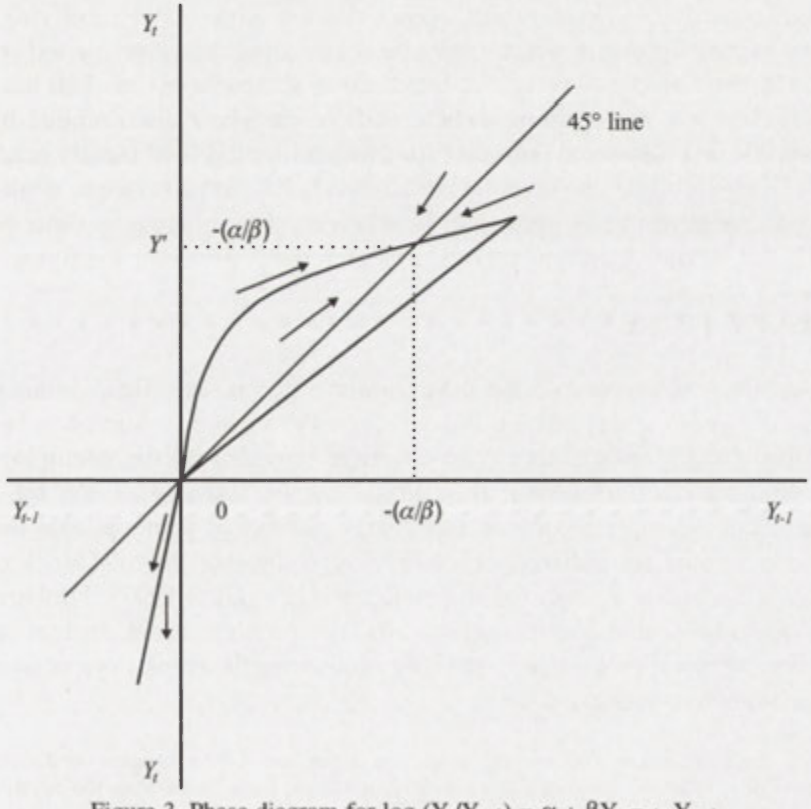


Figure 3. Phase diagram for $\log (Y_t/Y_{t-1}) = \alpha + \beta Y_t = {}_1 - Y_{t-1}$.

six regressions in Table 2, the equilibrium at Y is stable, while that found at the origin is not.

Table 3 provides the non-zero equilibria associated with the estimated model for the variables and time-periods considered above. The regressions for the whole period 1992–1997 imply stable equilibria for RYU and SRYU of approximately one-quarter and 0.3, respectively. These figures are of course considerably in excess of the share of the total labour force accounted for by young people. However, for the period 1992–1994, covering the upswing of Polish unemployment, the steady-state values implied by the model are of the order of one-half, an indication of just how serious an impact the transition recession actually had upon young people.

Table 3. Steady state values of relative youth unemployment.

	RYU	SRYU
Mar 92 – Dec 97	0.28	0.33
Dec 92 – Dec 97	0.24	0.29
Mar 92 – Sept 94	0.44	0.50
Sept 94 – Dec 97	–0.48	–0.71

As discussed above, the model breaks down beyond 1994 signifying that, contrary to the impression which might be drawn from the whole period regression, no further, simple convergence trend across spatial labour markets has been at work. As such, the apparent radical shift of the phase line implied by the (non-significant) parameter estimates for this later period is of merely academic interest.[8] In view of the break in the earlier trend, and the subsequent, if spatially seemingly random, overall reduction in relative youth joblessness described in Table 1, the following section seeks to examine the portents for the future.

A CRISIS AVERTED?

Given the concurrence of the data from the two most official sources, the reduction in youth joblessness in Poland since 1994 can be assumed to be real, rather than simply an artefact of the coverage provided by the unemployment registration statistics. However, this should not be taken as reason for complacency. There remain an unacceptably large number of young people without work who account for a disproportionately large share of the total stock of the unemployed, at least in terms of EU member states (GUS 1997). Furthermore, the evidence presented above suggests that the problem has a distinct spatial dimension, which appears to have stabilised following the breakdown of an initial tendency to convergence.

[8] The phase line has, in fact, twisted around the origin and, if the regression coefficients had been statistically different from zero, the equilibrium located there, rather than the negative one identified in Table 3, would be stable.

It is of course possible to argue, on the basis of recent experience in western Europe, that youth unemployment rates are typically higher than those for other groups because of the higher turnover rates that accompany the initial labour-market induction period. While it is difficult to explain wide regional differentials by this argument, if such reasoning has any substance then youth unemployment is more the result of a process of job experimentation, interspersed with short periods without work, which tempers emphasis on the phenomenon as a problem. In fact, however, 27.7% of the 537,000 youths recorded as unemployed in the August 1997 LFS had been without work for over one year, while the figure rose to 32.6% for those aged 20–24. Furthermore, very nearly half of the unemployed aged less than 25 had been so for over six months (GUS 1997a). Such figures hardly give much credence to the job experimentation thesis. Also, youth unemployment rates in excess of those for adults have not always been the norm, at least in Britain (e.g. Deakin 1996:55).

To the evidence on unemployment duration might also be added data on the economic activity and employment of the young in the present decade. This reveals that the decline in the rate of youth unemployment from 1994 until the end of 1997 was accompanied by a reduction in the proportion of the population aged 15–24 who were economically active, but by no appreciable increase in their employment rate (GUS 1998a:XXXII). The implied growth in the residual, non-economically active category of young people has translated into an increase in schooling rates over recent years (GUS 1997b:228).

The clear suggestion is that the reduction in youth unemployment has come about in large part because of the operation of a strong 'discouraged-worker' effect. This reflects an understandable reaction on the part of the individuals concerned and might be preferable on social grounds to the psychologically destabilising impact of rejection on the labour market that might otherwise confront school-leavers. However, one must not lose sight of the widespread belief that the Polish education system is unsuited to the needs of a modern economy (OECD 1996; Krajewska 1995) or of the fact that the government has recently pledged to undertake radical reform of its structure. The possibility must therefore exist that, all else equal, the threat of unemployment is merely being postponed rather than avoided. At the same time, some 17% of the young who were unemployed in August 1997 possessed primary-level schooling at most, while a further 46% had no more than secondary vocational education.

Furthermore, future developments may conspire to increase the problems confronting youths on the labour market. First, the 'baby-boom' of the years 1977–1983 means *ceteris paribus* an inevitable increase in the population of working age in the years to come (Kalaska and Witkowski 1996). Indeed, the demographic projections indicate no reduction in the size of the population aged 15–24 before the year 2005 (GUS 1996:94–95). Second, there remains the problem of hidden unemployment, most particularly in the agricultural sector; an issue that is worthy of brief further comment.

At least in a contemporary western setting, Poland is unusual in the size of that part of its workforce classified as 'unpaid family workers' under the conventions of the LFS. In 1997, there was still an average of over three-quarters of a million such workers in the economy (GUS 1998a:XX); around 90% of these were in agriculture, and nearly 30% were under the age of 25 (GUS 1997a:10). Such employment can be viewed as a traditional way of life – albeit one that is likely to be ill-suited to the ways of the EU to which Poland seeks accession, or else as nothing more than a buffer against overt unemployment (Ingham et al. 1998). Either way, further reform will undoubtedly imply that the youngsters involved, or perhaps those involved in the future, will come to be subjected to the demands of the external labour market. While clearly willing workers, they are highly unlikely to possess the kinds of skills that the new economy requires.

The question of agricultural reform is one of the major regional issues confronting Polish economy in the new millenium. Yet far from shrinking gradually, recorded agricultural employment has actually started to increase again (GUS 1997:28–30), and undoubtedly represents a sign of increased hidden unemployment. Somewhat ominously, Table 4 reveals that only one of the five voivodships with the highest concentrations of agricultural employment currently exhibits a particularly high relative youth jobless ratio. Indeed, only four of the ten voivodships with the largest agricultural workforces simultaneously figure in the corresponding list for youth unemployment.

Table 4. Highest agricultural employment and RYU, 1997.

Voivodship	% Ag Emp 1997	Voivodship	RYU 1997
Zamość	62.9	Tarnów	0.3928
Łomża	60.3	Konin	0.3924
Siedlce	55.1	Przemyśl	0.3864
Biała Podlaska	54.9	Leszno	0.3851
Ostrołęka	54.7	Sieradz	0.3846
Tarnobrzeg	53.7	Siedlce	0.3736
Przemyśl	52.2	Nowy Sącz	0.3666
Ciechanów	50.8	Tarnobrzeg	0.3600
Chełm	50.0	Kalisz	0.3537
Sieradz	49.8	Lublin	0.3528

Source: Computed from GUS (1998b) and GUS (1997c).

The upturn in economic activity witnessed in Poland following the ravages of the early recession associated with shock-therapy is, in itself, a healthy sign that the process of economic transformation is heading towards its target. However, the accumulated evidence indicates that this growth has not been translated into an equivalent increase in employment. The implied improvement in labour productivity is an important success, although the youngest cohorts of the Polish working population could be one of its major casualties. Until the impacts of

educational reform are translated into a more widely-dispersed distribution of marketable skills, and current regional imbalances – most particularly in the context of agriculture – are addressed firmly, the future of the country's young workers, wherever they live, cannot be assured.

CONCLUSION

In much the same way that equality and liberation were supposed to be found through participation in economic activity under the rhetoric of communist doctrine, so inclusion in democratic civil society is usually to be found through meaningful engagement with the world of work. Poland's future lies in the hands of its young people and it is imperative that this future is not placed in jeopardy by the unwitting creation of a lost generation. The recent reduction in the share of total unemployment accounted for by young people is therefore to be welcomed, although the number of potential workers aged under 25 who are without jobs is uncomfortably high by international standards (GUS 1997:208; Kabaj 1996:12). Furthermore, a number of factors exist which suggest that, without intervention, the future may not witness a virtuous improvement in the outlook for young workers.

What is required is strategic thinking, rather than reactive policy enactments, and this is most likely to evolve in the context of tripartite understandings between government, employers and trade unions. Expecting the European Social Fund (ESF) to be a palliative for the problem of youth unemployment, as well as a number of other ills, is not a policy at all. The ESF will be reformed before the EU expands to the east: an inevitability in view of the dilution of overall Union resources that widening will entail. Even if this were not the case, it is unrealistic to believe that appropriate policies can be devised from afar; this is a domestic task. Nevertheless, the experience of EU members, gained in over three decades of attempting to combat the problem of youth unemployment, could be invaluable and should be canvassed.

This said, in its desire to join the European mainstream, Poland would probably be well advised to avoid some of the more controversial EU initiatives at this stage in its development. In the case of the youth unemployment problem, this must surely involve the rejection of minimum wage directives that serve merely to price young workers out of the market. The experience of German Unification (Sinn 1995) should be sufficient to warn against such a policy. At the same time, it is imperative that no heed is paid to arguments which imply that service-sector jobs, notwithstanding their low pay rates, are inherently bad. They frequently provide young people with their first real position on the labour market and, furthermore, may be the ultimate source of salvation for many of Poland's agricultural regions.

With the wide-ranging spatial discrepancies in labour-market performance

that currently confront the country and the consequent variations in the opportunities available to young people, it is likely that locally-driven initiative is the primary requirement. While likely to be found wanting on several other counts, it is quite possible that the new poviat tier of local government will be the appropriate location of responsibility for the stewardship of the transition of young people into the world of work, provided that the appropriate personnel are appropriately empowered and trained. Given the budgetary constraints that undeniably still confront Poland, this may be the best way of avoiding the deadweight and substitution costs that almost invariably accompany government programmes of market intervention while, at the same time, ensuring a tangible pay-off to both the country's youth and its own future prosperity.

REFERENCES

Barro R., Sala-I-Martin X., 1991, *Convergence Across States and Regions*, Brookings Papers on Economic Activity, 1, 107–128.

Baumol W., 1986, *Productivity Growth, Convergence and Welfare: What the Long-Run Data Show*, American Economic Review, 76, 5, 1072–1085.

Chatterji M., 1992, *Convergence Clubs and Endogenous Growth*, Oxford Review of Economic Policy, 8, 1, 57–69.

Chatterji M., Dewhurst J., 1996, *Convergence Clubs and Relative Economic Performance in Great Britain, 1988–1991*, Regional Studies, 30, 1, 31–40.

Deakin B. M., 1996, *The Youth Labour Market in Britain: The Role of Intervention*, Cambridge University Press, Cambridge.

Durlauf S., Quah D., 1998, *The New Empirics of Economic Growth*, Centre for Economic Performance Discussion Paper 384, March, London School of Economics, London.

Freeman R., Wise D. (eds), 1982, *The Youth Labor Market Problem: Its Nature, Causes and Consequences*, University of Chicago Press, Chicago.

Główny Urząd Statystyczny, GUS, 1998, *Bezrobocie rejestrowane w Polsce. II kwartał 1998*, Warsaw.

– 1998a, *Aktywność ekonomiczna ludności Polski. Maj 1998*, Warsaw.
– 1998b, *Pracujący w gospodarce narodowej w 1997 r.*, Warsaw.
– 1997, *Rocznik Statystyczny Pracy 1997*, Warsaw.
– 1997a, *Aktywność ekonomiczna ludności Polski. Sierpień 1997*, Warsaw.
– 1997b, *Rocznik Statystyczny 1997*, Warsaw.
– 1997c, *Bezrobocie rejestrowane w Polsce. I–IV kwartał 1997*, Warsaw.
– 1996, *Rocznik Statystyczny Demografii 1996*, Warsaw.

Gorzelak G., 1996, *The Regional Dimension of Transformation in Central Europe*, Jessica Kingsley Publishers and Regional Studies Association, London.

Ingham M., Grime K., 1994, *Regional Unemployment in Central and Eastern Europe*, Regional Studies, 28, 8, 811–818.

Ingham M., Ingham H., Karwinska A., Węcławowicz G., 1998, *Women in the Polish Labour Market: Is Transition a Threat*?, [in:] Bull M. J., Ingham M. (eds), *Reform of the Socialist System in Central and Eastern Europe*, Macmillan Press Ltd., Basingstoke.

Ingham H., Ingham M., 1997, *Regional Unemployment in Poland: A Spatial Autocorrelation Approach*, paper presented at European Urban and Regional Research Network Conference, Frankfurt an der Oder, September.

Kabaj M., 1996, *Programmes and Strategies for Counteracting Unemployment and the Pro-*

motion of Productive Employment in Poland, ILO-CEET Report No. 15, Budapest: International Labour Office.

Kalaska M., Witkowski J., 1997, *Labour Market in Poland in 1996: Continuation of Positive Trends*, Warsaw: GUS, Department of Labour Statistics.

Kowalska A., *School-Leavers on the Polish Labour Market in the Period of Transition*, Studia Demograficzne, 1–2/123–124, 173–183.

Krajewska A., 1995, *Education in Poland*, Eastern European Economics, 33, 4, 54.

Lindbeck A., 1991, *Microfoundations of Unemployment Theory*, Seminar Paper No. 503, Institute for International Economic Studies, Stockholm.

Lindbeck A., Snower D., 1988, *The Insider-Outsider Theory of Employment and Unemployment*, Mass: The MIT Press, Cambridge.

Lindley R., 1996, *The School-to-Work Transition in the United Kingdom*, International Labour Review, 135, 2, 159–180.

OECD, 1977, *Selection and Certification in Education and Employment*, Paris.

– 1996, *Reviews of National Education Policies for Education: Poland*, Paris.

Quah D., 1996, *Regional Economic Convergence Clusters Across Europe*, European Economic Review, 40, 5, 951–958.

Reise M., 1993, *Improving the Use of Administrative Data Sources in Poland,* [in:] *OECD, Employment and Unemployment in Economies in Transition: Conceptual and Measurement Issues,* Paris: Eurostat.

Roberts K., 1995, *Youth and Employment in Modern Britain*, Oxford University Press, Oxford.

Ryan P., Garonna P., Edwards R. (eds), 1991, *The Poblem of Youth: The Regulation of Youth Employment and Training in Advanced Economies*, Macmillan, Basingstoke.

Saganowska E., Śmilowska T., 1996, *Estimates of GDP by Voivodships in the Years 1992–1994*, Research Bulletin of RECESS, 5, 3, 45–69.

Scarpetta S., Huber P., 1995, *Regional Economic Structures and Unemployment in Central and Eastern Europe: An Attempt to Identify Common Patterns*, [in:] *OECD, The Regional Dimension of Unemployment in Transition Countries*, Centre for Co-Operation with the Economies in Transition, Paris.

Sinn H. W., 1995, *Staggering Along: Wages Policy and Investment Support in East Germany*, The Economics of Transition, 3, 4, 403–426.

Węcławowicz G., 1996, *Contemporary Poland: Space and Society*, UCL Press, London.

GEOGRAPHIA POLONICA, 73, 1, Spring 2000
PL ISSN 0016-7282

UNEMPLOYMENT DYNAMICS IN THE PROCESS OF SYSTEM TRANSFORMATION IN POLAND

TERESA CZYŻ, PAWEŁ CHURSKI, JAN HAUKE
Institute of Socio-Economic Geography and Space Economy,
Adam Mickiewicz University, Fredry 10, 61-701 Poznań, Poland

ABSTRACT: The article presents an analysis of changes in the level and character of unemployment in Poland in the years 1990–1998 on the basis of a national and a regional approach. It embraces: (1) relations between national unemployment and changes in the economy, (2) a classification of regions on the basis of relative unemployment, and (3) the testing of the process of petrifaction of the regional structure of unemployment using the T-technique of principal components analysis.

KEY WORDS: Poland, national unemployment and stages of transformation, regional variation in unemployment over time, T-technique of principal components analysis.

INTRODUCTION

The process of systemic transformation initiated in Poland in 1989 and involving a gradual introduction of mechanisms of a free-market economy has changed the state and structure of the country's labour market. The one characterised by excessive employment and a permanent labour deficit (254,000 vacancies in 1989) has changed into one of a great demand for work, one in which job seekers far outnumber job offers. The economic crisis in the initial phase of the transformation as well as the restructuring of the economy and its concomitant redundancies led to unemployment which reached a national peak in 1993 when it embraced 16% of the active population. Despite the downward trend in unemployment over the years 1995–1997, the 1998 unemployment rate was 10,4%.

Unemployment has become a serious social problem in Poland. Its persistence contributes to the lowering of standards of living and a negative attitude towards the process of economic transformation.

The aim of the present article is to analyse the level and character of unemployment in Poland over the years 1990–1998 on the national and regional scales.

The analysis of changes in unemployment refers to the country as a whole and to the country as a regional system. In the first approach, unemployment is studied as an aggregate quantity on the national scale, and in the other, it is the regional structure of unemployment that is considered, i. e. its structure in the

pattern of 49 administrative regions (voivodeships) that existed until 31 December 1998. A comparison of the results of these two approaches was to help define the relations holding between changes in national unemployment and those in the regional structure of unemployment.

The work made use of statistical materials published by the Central Statistical Office and the Government Centre for Strategic Studies.

NATIONAL UNEMPLOYMENT IN THE YEARS 1990–1998

1990 was the first year in which unemployment was registered as a significant property of the labour market in Poland. At the close of 1990 the unemployment figure reached 1,126,000, which meant a 3.5-fold increase over the level registered in April of that year. In December the unemployed constituted 6% of the active population, as against 1.7% in April. Unemployment kept climbing sharply in 1991: by the close of that year it had reached 2,157,000, which meant an unemployment rate of 11%. In the following years unemployment grew less dynamically, but even so it rose to 2,509,000 people in 1992 (a rate equal to 13.6%) and to 2,890,000 people in 1993 (16.4%). The peak was recorded in the first quarter of 1994, when the unemployment figure reached 2,953,000 (a rate of 16.7%). Since that time unemployment has began to fall steadily, to 2,838,000 in December 1994; 2,629,000 in 1995; 2,360,000 in 1996; 1,826,000 in 1997; and finally 1,831,000 in 1998, with the respective unemployment rates decreasing from 15.9% to 10.4% (Tab. 1).

An analysis of the dynamic curve of national unemployment in Poland allows for the distinguishing of 4 periods characterised by (Fig. 1):

(1) a steep increase in unemployment (1990–1991),

Table 1. The labour market in Poland in the years 1990–1998.

Years	Employment in thous. (as on 31 Dec.)	Employment in the private sector in thous.	Unemployment	
			in thous.	as % of working population
1990	16485	7902	1126	6.0
1991	15772	8390	2157	11.0
1992	15356	8404	2509	13.6
1993	15117	8700	2890	16.4
1994	15282	9045	2838	15.9
1995	15324	9344	2629	14.8
1996	15842	10074	2360	13.5
1997	15927	10990	1826	10.5
1998	16232	11250	1831	10.4

Source: Little 1997,1998, 1999 Yearbook. Warszawa: GUS.
Centre for Government Strategic Studies, 1998.

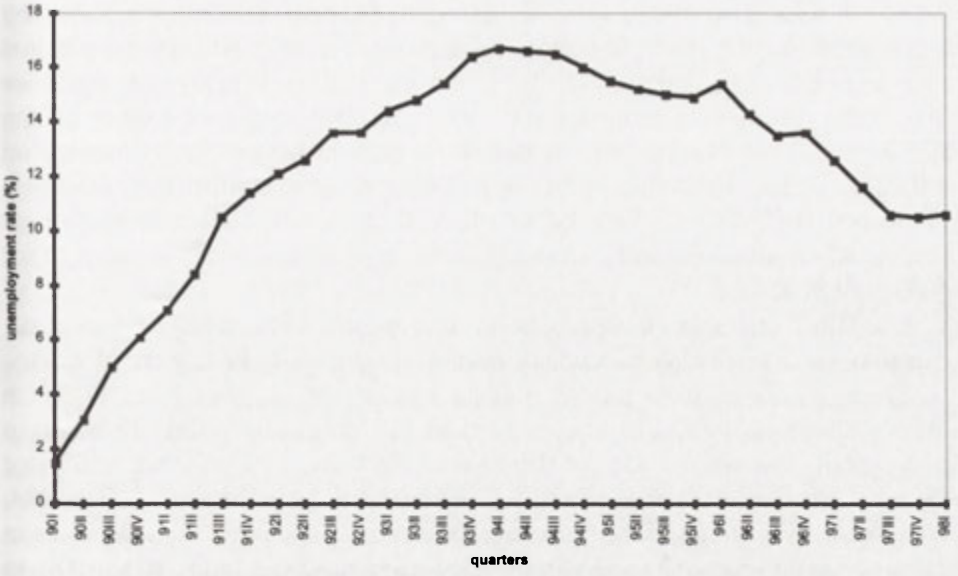


Figure 1. Unemployment rate in Poland in the years 1990–1998.

(2) a further rise in unemployment and sustained high level (1992–1994, with a peak in the first quarter of 1994),

(3) a fall in unemployment (1995–1996),

(4) a further drop and stabilisation (1997–1998).

In the years 1990–1997 unemployment in Poland was connected with a decrease in employment in the economy. 1990 ended with the working population one million lower than in 1989 (December-to-December basis), while the years 1990–1993 witnessed a fall of 1.4 million (Tab. 1). Between 1990 and 1997 there was a drop in the number of people in jobs in the public sector. As of 1997 there were 4.9 million people working in this sector, or only 57% of the number in 1990. In turn, the number of people employed in the private sector grew slowly but steadily in the period; its share in total employment increasing from 48% to 69%. The rise in employment in the private sector slowed the rate of decline in the total number employed in the years 1990–1993, and has even caused a slight increase since 1994.

The level of registered unemployment in Poland also depends to some extent on unemployment regulations and how they change.

In the years 1990–1992, when the number of working people dropped from 16.5 million to 15.4 million, unemployment increased to 2.5 million. Of this figure, an estimated 0.5 million were people unemployed hitherto, but registering with employment agencies at that point in order to claim unemployment benefits. It was in 1997, when the number of people in jobs increased by 85,000 and the size of the population of working age rose by 185,000 with respect to the previous year, that the highest fall in unemployment was recorded – of 533,000

people. A significant role in diminishing registered unemployment was played by amendments to the unemployment regulations.[1] In 1997, 800,000 people lost their unemployed status because they had not notified employment agencies about their readiness to undertake jobs, while 200,000 people were taken off the files because they had become entitled to pre-retirement benefits or enrolled on a training course. The requirement for the unemployed to confirm their readiness for work frequently was intended to eliminate those who drew unemployment benefit while simultaneously working in the grey sector of the economy (not registered officially).

The Polish unemployed population displays specific demographic and socio-occupational features. Also characteristic of it is a long period of being out of work.

There is a large proportion of jobless women, which grew from 47.8% in 1990 to 58.5% in 1998. The unemployed are predominantly young. Those up to 24 years old constituted 35% of the total in 1991 and 31% in 1998, and those from the 25–34 age bracket, 30% and 27% respectively. In the years 1991–1998, with a downward trend in the proportion of young unemployed and a decrease in the proportion of school leavers in the total unemployment figure (from 10% to 6%), the share of those over 45 increased from 10% to 16%. As to the level of education, people with basic vocational and elementary education predominate (67% in 1991 and 72% in 1998). Those with higher education form the smallest proportion of the unemployment figure: 3% in 1991 and 1.6% in 1998.

As to the duration of unemployment, there is a persistently high, or even growing, proportion of people out of work for a long time. The rise in long-term unemployment accompanied by a downward trend for unemployment in general is indicative of structural unemployment. In 1998 23% of those of work had been so for more than 2 years (19% in 1996).

The turnover among the unemployed increased in the years 1993–1995: there were rises in the numbers newly registered and dropped from the files. It was only in 1996 that the turnover slowed owing to a drop in the number of new unemployed.

A significant feature of unemployment has been a steady decrease in the proportion of those newly registered unemployed thanks to mass lay-offs, and an increase in the proportion of those dropped from the files after they have found jobs. In the years 1991–1998, the share of unemployed people not entitled to benefit grew from 20% to 77%. In September 1991 there were 74 unemployed per job offered, while in 1993 this figure fell to 68, and in 1997 to 58. It is estimated, however, that enterprises notify employment offices of only 30–50% of their vacancies.

[1] The amendments consisted in: (1) extending the period of work entitling a person to a benefit from 12 to 18 months, which limited its accessibility, (2) defining different periods of drawing the benefit depending on the unemployment rate on the local labour market, (3) introducing new forms of occupational mobilisation of the unemployed, (4) making it obligatory for an unemployed person to confirm a his/her permanent readiness to take up a job, and (5) introducing pre-retirement benefits.

UNEMPLOYMENT DYNAMICS AND CHANGES IN THE ECONOMY

The years 1990–1998 saw unemployment in Poland first triggered by economic crisis and then by structural changes taking place in the economy.

The years 1990–1992, i. e. the period of recession, were marked by a dramatic decline in GDP and high inflation (Tab. 2). In 1990 GDP decreased by 8% with respect to the previous year, while the annual inflation amounted to 686%. The decline in GDP was mainly a result of a breakdown in industrial production. Many state-owned industrial enterprises went bankrupt, or started to be liquidated or privatised. The fall in industrial employment was the principal cause of unemployment.

In the years 1992–1994 a rapid rise in unemployment coincided with what were already conditions of a slow increase in GDP and a considerable decline in inflation. Growing industrial production was accompanied by a further fall in industrial employment, also an indicator of the rationalisation of employment. The level of unemployment was affected greatly by the liquidation of unproductive state farms. What exerted a decisive influence on the economy was the development of the private sector. The number of state enterprises under privatisation grew substantially. The setting-up of new businesses, mostly small and medium-sized, represented by commercial companies, led to an increase in the number of job offers, mainly in services. However, the rise in employment in the private sector only managed to slow down the rate at which unemployment rose, being too small to improve the situation on the labour market.

The years 1994–1998 were a period of a marked economic revival that was favourable to beneficial changes in employment. Starting with 1995, unemploy-

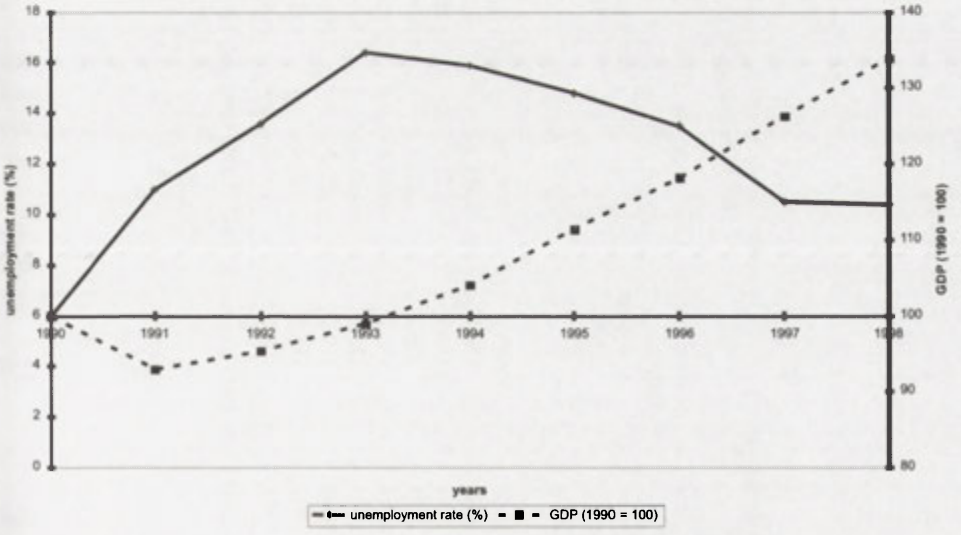


Figure 2. Unemployment and GDP in Poland, 1990–1998.

Table 2. Indices of Poland's economic development in the years 1990–1998.

Year	Annual rate of inflation, previous year = 100	GDP, previous year = 100	Industrial production sold, previous year = 100	Industrial employment in thous.	Number of state enterprises being privatised	Number of state enterprises	Number of commercial companies	Global agricultural production (constant prices), previous year = 100
1990	685.8	92.0	76.8	4620	130	8453	36267	97.8
1991	170.3	93.0	92.0	4250	1128	8228	53774	98.4
1992	143.0	102.6	102.8	3778	1401	7245	69907	87.3
1993	135.3	103.8	106.4	3671	1271	5924	83283	106.8
1994	132.2	105.2	112.1	3640	791	4955	95017	90.7
1995	127.8	107.0	109.7	3729	501	4357	104922	110.7
1996	119.9	106.1	108.3	3754	362	3847	115739	100.7
1997	115.5	106.9	111.5	3761	290	3369	126465	99.9
1998	108.6	104.8	104.8	3650	264	3219	128778	106.6

Source: Little 1997 Yearbook. 1993, 1996, 1998 and 1999 Yearbooks. Centre for Government Strategic Studies, 1998.

ment began to fall steadily. The symptoms of economic development were annual increases in GDP of 5–7% and declining inflation (8.6% in 1998). Privatisation of the economy is an ongoing process. The number of state-owned enterprises combines to dwindle: in 1998 these constituted 38% of the 1990 figure. In 1997 the private sector employed 69% of the total workforce and generated 65% of GDP. Industrial output is showing a steady upward trend, of 9% a year on average. Employment in the private industrial sector is rising slowly. However, it is the development of services that has been decisive for the increase in the number of jobs. In 1997 those employed in the services constituted 43% of total employment (as against 37% in 1990). The dynamic of change is low in agriculture. Private farming, overpopulated, small-scale and inefficient, does not undergo modernisation. The situation in agriculture is also aggravated by the slow pace at which state agricultural property is restored to former owners as well as by its poor development. Lack of investment capital in rural areas makes new job opportunities outside agriculture appear only slowly. As a result, while national unemployment is on the decline, rural unemployment persists at a relatively high level.

Economic growth and unemployment are closely related in Poland. A comparison of the income curve reflecting economic growth in the years 1990–1998 with that of the unemployment rate allows two periods to be distinguished in the pattern of relations between these magnitudes (Fig. 2). In the years 1990–1993, i.e. the period of crisis and the slow breaking out of it (a drop and a weak increase in GDP), unemployment grew rapidly. In the years 1994–1998, after it had reached the 1990 level in 1994, the income showed a high level of annual increase (5–7%). The upward trend to income is accompanied by a fall in unemployment.

TEMPORAL VARIABILITY OF REGIONAL UNEMPLOYMENT

The national unemployment curve for the years 1990–1998 is a generalised and smoothed image of regional trends. Regional curves differ, because regions have various patterns of unemployment growth (decline) over time. The 49 curves of regional unemployment plotted against 33 quarters of the period in question reflect a variety of unemployment rates and differ in shape.

The classification of the dynamic curves of regional unemployment is made on the basis of an analysis of relations between the regional and the national unemployment rates in the years 1990–1998. This relation is measured by the b_i ratio which Gleave (1987) calls relative unemployment.

The b_i ratio defined for the 49 regions for the entire period 1990–1998 embracing 33 quarters is given by the equation:

$$\frac{u_{it}}{U_t} = b_i \quad i = 1,2,3,\ldots,49;\ t = 1,2,3,\ldots,33,$$

where: u_{it} = unemployment rate in region i at time (quarter) t; U_t = national unemployment rate at time (quarter) t.

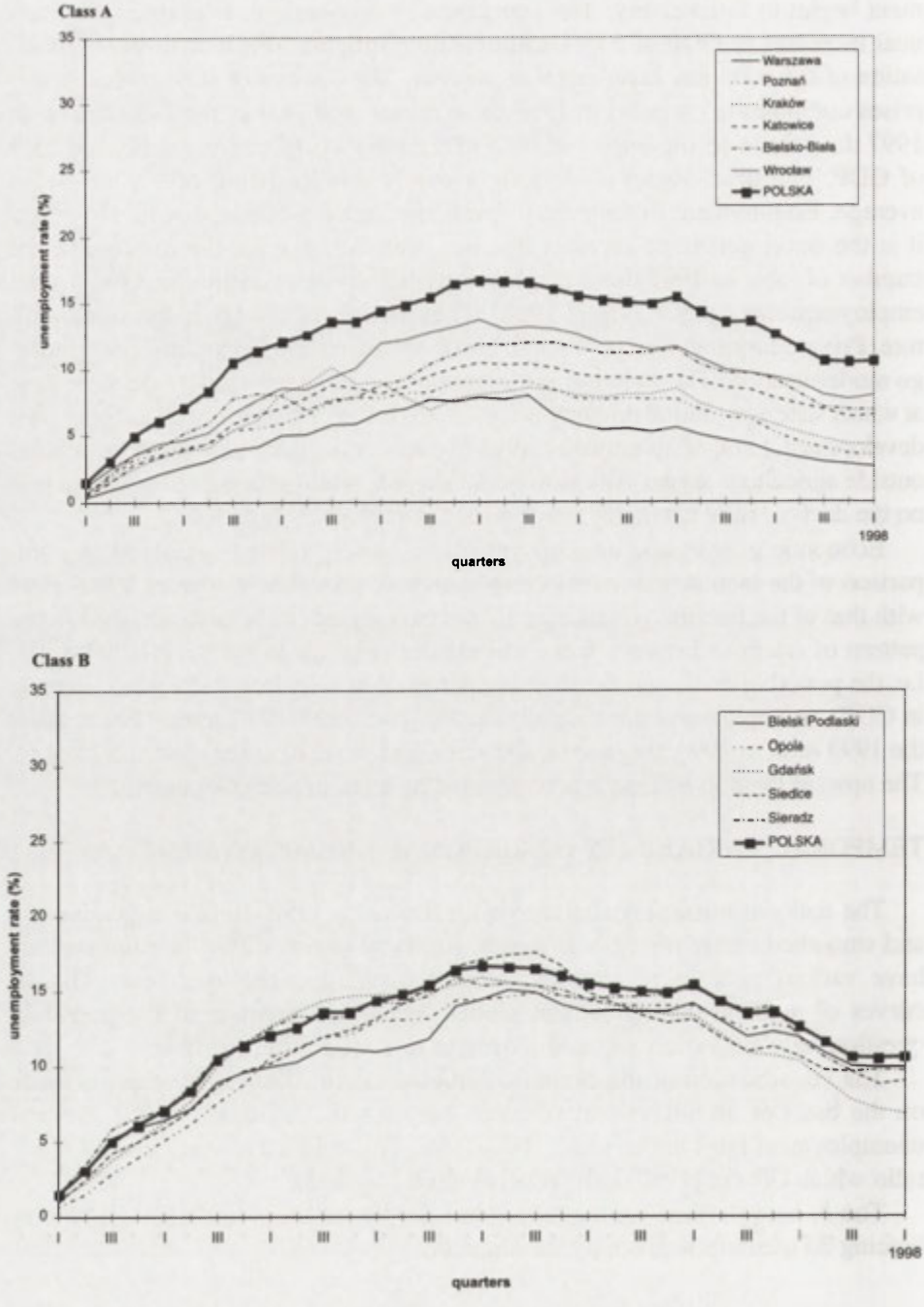
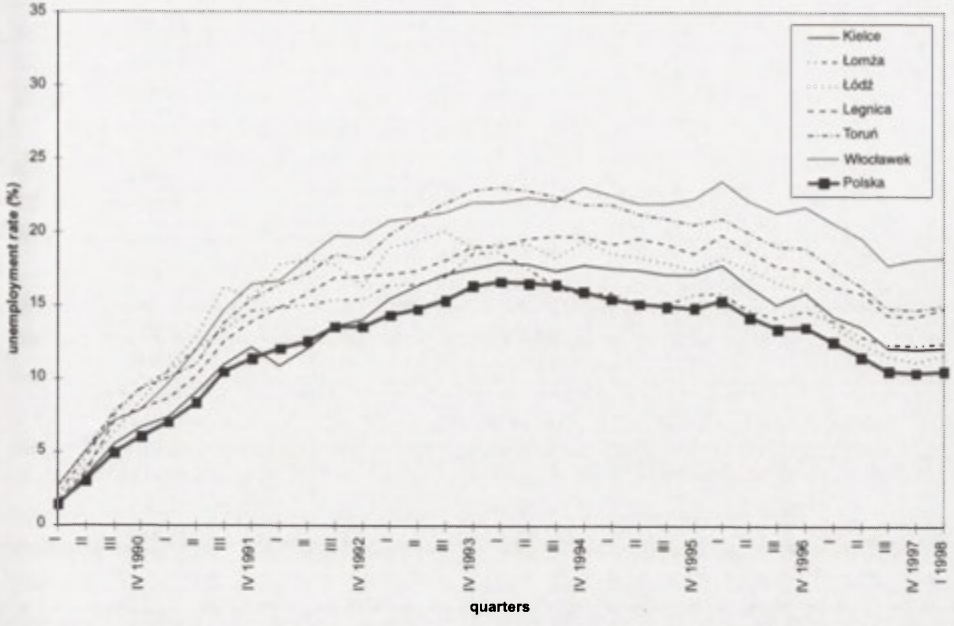
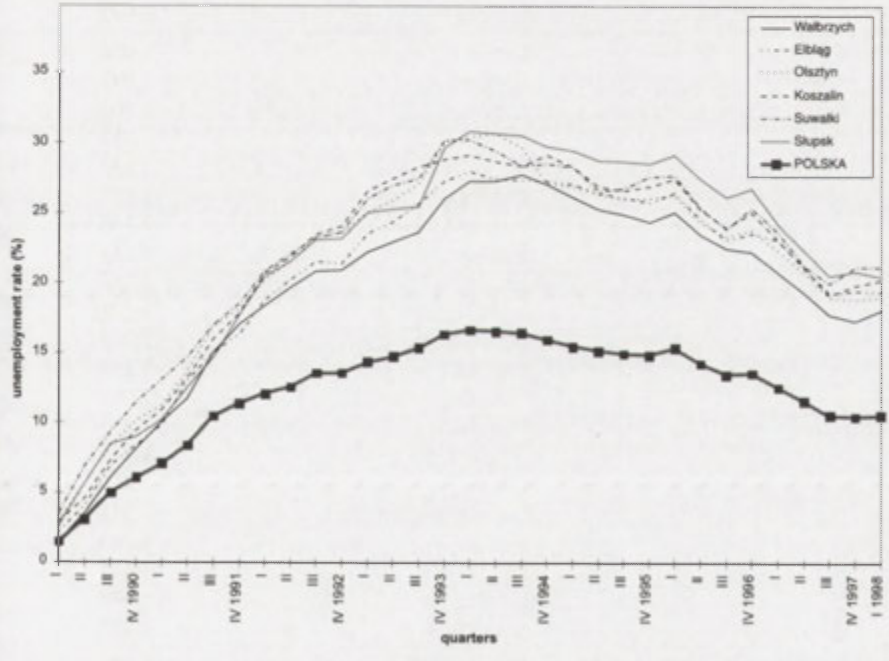


Figure 3. Regional unemployment curves by classes A, B, C and D

Class C



Class D



The $\hat{b}_i$ ratio was estimated using the least-squares method. Its values are significantly different from zero for $\alpha = 0.01$. The coefficient of determination is higher than 80%.

Table 3. Relative unemployment.

Class	Region (i)	b_i ratio
A	Warsaw	0.396
	Poznań	0.508
	Kraków	0.530
	Katowice	0.615
	Bielsko-Biała	0.694
	Wrocław	0.766
B	Biała Podlaska	0.874
	Lublin	0.883
	Opole	0.898
	Szczecin	0.902
	Skierniewice	0.903
	Leszno	0.911
	Tarnów	0.912
	Białystok	0.916
	Gdańsk	0.918
	Siedlce	0.921
	Zamość	0.933
	Częstochowa	0.935
	Nowy Sącz	0.941
	Chełm	0.952
	Tarnobrzeg	0.968
	Sieradz	0.972
C	Kalisz	1.081
	Kielce	1.099
	Łomża	1.107
	Krosno	1.117
	Rzeszów	1.130
	Przemyśl	1.135
	Zielona Góra	1.146
	Radom	1.188
	Konin	1.216
	Łódź	1.228
	Bydgoszcz	1.244
	Legnica	1.245
	Piotrków Trybunalski	1.261
	Ostrołęka	1.282
	Jelenia Góra	1.311
	Płock	1.318
	Gorzów Wielkopolski	1.345
	Toruń	1.389
	Piła	1.415
	Ciechanów	1.449
	Włocławek	1.467
D	Wałbrzych	1.609
	Elbląg	1.678
	Olsztyn	1.748
	Koszalin	1.771
	Suwałki	1.798
	Słupsk	1.817

The results of the b_i estimation for the 49 regions in the years 1990–1998 are contained in Table 3. The b_i ratio assumes values from the [0.396, 1.817] interval. On the basis of an analysis of the b_i distribution the regions are divided into four classes: A, B, C and D, varying in relative unemployment.

The unemployment curves of regions in the particular classes (A, B, C and D) display a high degree of convergence, or even overlap (Fig. 3). This is indicative of a similar pattern of temporal change in the unemployment rate in regions of the same class.

Classes A and D contain extreme values of b_i. Class A embraces regions with the lowest b_i values in the interval [0.396, 0.766]. It is made up of six regions: Warsaw, Poznań, Kraków, Katowice, Bielsko-Biała, and Wrocław, where the unemployment rate was much below the national average throughout the period 1990–1998. Class D is composed of regions with the highest b_i values contained in the interval [1.609, 1.817]. It includes six regions: Wałbrzych, Elbląg, Olsztyn, Koszalin, Suwałki, and Słupsk, in which the unemployment rate is 1.5 times the national one. Class B with 16 regions and b_i values ranging in the interval [0.874, 0.972] shows unemployment rates lower than, but close to, the national one. The most numerous is class C with its 21 regions, in which unemployment rates exceed the national one, but by not more than 1.5 times (b_i in the interval [1.081, 1.467]).

The regions of the particular classes differ in their level of development and their socio-economic structures.

In the low-unemployment class A, there are economically strong regions with well-developed industry and services and a high level of urbanisation. The high-unemployment class D is composed of average and weak regions, agricultural-industrial and agricultural in character (with predominantly socialised agriculture until 1992). An exception is the industrial voivodeship of Wałbrzych.

In classes B and C, in turn, there is a mixture of strong, average and weak regions, as well as industrial, industrial-service, agricultural and industrial-agricultural ones. The lack of correlation between the membership of classes B and C, and the level of development and industrialisation results from different responses of the particular industries to the crisis in the initial period of transformation. The high unemployment in some strong and average regions was mainly connected with the fall in employment in the textile industry (Łódź, Jelenia Góra), the power or extractive industries (Płock, Legnica, Piotrków), and the engineering and metallurgical industries (Bydgoszcz, Rzeszów, Kielce, Radom).

It should be noted that owing to the lack of a significant relation between the level of unemployment in a region as measured by the unemployment rate and the absolute number of unemployed in that region, the shares of the particular classes of voivodeships (A, B, C, D) and individual voivodeships in national unemployment are poorly correlated with the value of the b_i ratio. To give an example, over the long-term period of 1990–1998 Warsaw voivodeship with b_i = 0.396 contributed 2.5% to the national total of unemployed, Katowice voivode-

ship with b_i = 0.615 had a share of 5.7%, while Łomża voivodeship with b_i = 1.107, had only 1%. The proportion of the six-member class A with the lowest b_i values in the national total is 15%, and of class D with the highest b_i values, comprising also six voivodeships, 14%.

The regional unemployment curves have irregular shapes. In the generalised approach, they are positively skewed, and show symmetry only exceptionally. The asymmetry of the curves is connected with the high growth dynamic to the unemployment rate in the years 1990–1993 and, after peak unemployment, a slow decline. The curve pattern also reveals sharp, short-term fluctuations from quarter to quarter which reflect seasonal changes in the regional unemployment rates.

A characteristic feature of the regional unemployment curves is their flattening, which means the persistence of the highest unemployment over a longer period. Hence, the peak period of unemployment in a region cannot be determined by the occurrence of the maximum unemployment rate. Rather, it is determined by fitting a theoretical parabolic curve (of a quadratic polynomial) to the flattened empirical curve (Fig. 4). The least-squares method is used to determine the parameters of the parabola with the equation:

$$\hat{u}_i = a_0 + a_1 t + \sqrt{a_2 t^2} \qquad \begin{matrix} i = 1, 2, 3, \ldots, 49 \\ t = 1, 2, 3, \ldots, 33 \end{matrix}$$

where: $\hat{u}_i$ = unemployment rate in region i;
t = time periods (quarters);
a_0, a_1, a_2 = regression coefficients.

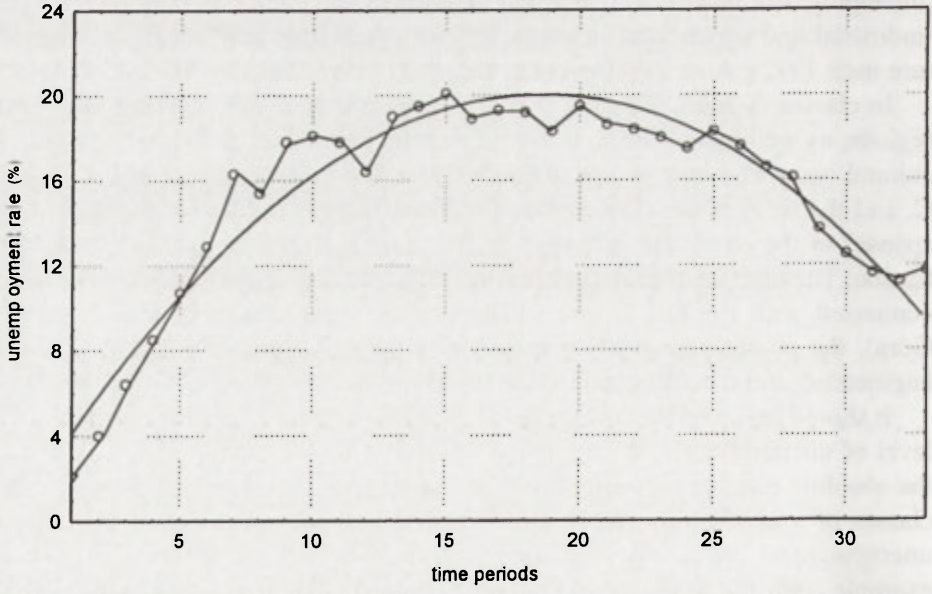


Figure 4. Goodness-of-fit of a parabolic curve given by the equation $\hat{u}_i = 2.238 + 1.913t - 0.051t^2$ (R^2= 92.59%) to the curve of the unemployment rate in the Łódź region.

Table 4. Co-ordinates (u_{max}, t_{max}) of regional unemployment curve peaks.

Regions	u_{max}	t_{max}
Warsaw	6.91	17.79
Poznań	8.60	18.07
Kraków	8.86	18.35
Katowice	10.10	20.02
Bielsko-Biała	11.34	20.10
Wrocław	12.66	19.68
Biała Podlaska	13.96	20.54
Lublin	14.14	19.90
Białystok	14.48	18.86
Szczecin	14.62	20.56
Opole	14.74	20.79
Tarnów	14.75	20.12
Leszno	14.77	19.56
Skierniewice	14.89	18.90
Zamość	14.96	20.68
Nowy Sącz	15.02	20.53
Częstochowa	15.31	18.78
Chełm	15.31	21.86
Tarnobrzeg	15.38	21.83
Siedlce	15.42	19.18
Gdańsk	15.61	18.26
Sieradz	15.72	19.39
Kalisz	17.36	19.71
Łomża	17.39	19.14
Krosno	17.73	20.57
Kielce	17.78	20.41
Przemyśl	18.21	19.79
Rzeszów	18.23	19.94
Zielona Góra	18.37	19.90
Konin	19.10	20.99
Radom	19.28	21.35
Legnica	19.91	20.59
Łódź	20.00	18.57
Bydgoszcz	20.05	19.78
Piotrków Trybunalski	20.07	19.78
Ostrołęka	20.38	20.37
Jelenia Góra	20.62	20.07
Płock	21.18	19.25
Gorzów Wielkopolski	21.99	19.05
Toruń	22.40	19.80
Ciechanów	23.07	19.96
Włocławek	23.36	21.26
Piła	23.62	18.94
Wałbrzych	26.43	20.29
Elbląg	27.33	20.67
Olsztyn	28.44	19.89
Suwałki	28.87	20.27
Koszalin	28.92	20.21
Słupsk	29.62	20.77
Poland	**16.25**	**19.71**

The goodness-of-fit of the curves is given by the coefficient of determination R^2, which assumes values above 85% (including 46 cases of more that 90%, and 35 cases of more than 95%). The co-ordinates of the vertex of the parabola: t_{max} (on the time axis) and $\hat{u}_{max}$ (on the unemployment rate axis), define the position of the estimated peak of the regional unemployment curve, which is shifted in time with relation to the moment of the maximum level of unemployment in the region. The peak sections of the 49 regional unemployment curves define unemployment rates ($\hat{u}_{max}$) in the interval [6.91, 29.62], and t_{max} values in the temporal interval [17.79, 21.86], that is, between the second quarter of 1994 and the second quarter of 1995 (Tab. 4). The estimated peak of the national unemployment curve is determined by the co-ordinates $\hat{U}_{max}$ = 16.25 and T_{max} = 19.71 (the 4th quarter of 1994).

The distribution of unemployment rates in the peak period looks as follows in the particular classes: class A, 7–13%; class B, 14–16%; class C, 17–24%; and class D, 26–30%. The time of occurrence of the peak varies from region to region. It appeared first in the 2nd quarter of 1994 in three regions of class A: Warsaw, Poznań and Kraków, and in the Gdańsk region from class B; it appeared latest in the 2nd quarter of 1995 in the regions of Chełm and Tarnobrzeg from class B. However, a decided majority of the regions (43) experienced peak unemployment in the third or fourth quarter of 1994, or the first quarter of 1995. There were 10 regions with peaks in the third quarter of 1994 (5 from class B and 5 from class C), 21 regions with peaks in the fourth quarter of 1994 (3 from class A, 3 from class B, 11 from class C, and 4 from class D), and 12 regions – in the first quarter of 1995 (5 from class B, 5 from class C, and 2 from class D). After the 1995 peak there began a process of decline in regional unemployment, which lasted until the end of 1997. In 1998 the level of regional unemployment stabilised noticeably.

CHANGES IN THE REGIONAL STRUCTURE OF UNEMPLOYMENT IN THE YEARS 1990–1998

Changes in the regional structure of unemployment over time were studied by way of: comparative and process approaches.

In the first, on the basis of a map series, the states of the regional structure of unemployment in specific time periods are described and compared. The chosen time is the December of the years 1991, 1992 and 1994, and the March of 1998, which are representative of the successive stages of unemployment in Poland. The comparative analysis allows differences and similarities to be traced between the consecutive states of the regional structure, which can then be associated with specific components of this structure.

In 1991, towards the end of the first stage of unemployment, the mean national unemployment rate was of 11% while its extreme values were of 4% in

Warsaw voivodeship and 19% in Suwałki voivodeship, so the regional structure of unemployment was already highly diversified (Fig. 5). There were three compact areas of the highest unemployment rates (in excess of 15%): (1) the north-eastern and central parts of Poland (9 voivodeships); (2) north-western Poland (4 voivodeships); and (3) south-western Poland (2 voivodeships). Low-unemployment rates (of 4–7%) were noted in the voivodeships of Warsaw, Poznań, Katowice and Kraków.

At the start of the second stage to the rise in unemployment in 1992, the mean national unemployment rate was of 14%, while the highest regional rates exceeded 19% and formed patterns of smaller spatial extent than those of 1991 (Fig. 6). The voivodeships with the lowest unemployment rates were joined by Bielsko-Biała. The extreme values for the rate were recorded in Warsaw (6%) and Koszalin (24%).

The years 1993–1994, which were the period of peak unemployment in Poland (mean national rates of 16%) saw the distribution of the highest unemployment rates (over 20%) form patterns similar to those of 1992 (Fig. 7). The voivodeships with the lowest unemployment (below 12%) were again those of Warsaw, Poznań, Katowice and Kraków. The extreme figures were 7% (in Warsaw voivodeship) and 30% (in Słupsk voivodeship).

In the stage of the steady fall in unemployment, as in the first quarter of 1998,

The regional (voivodeship) system, 1998.

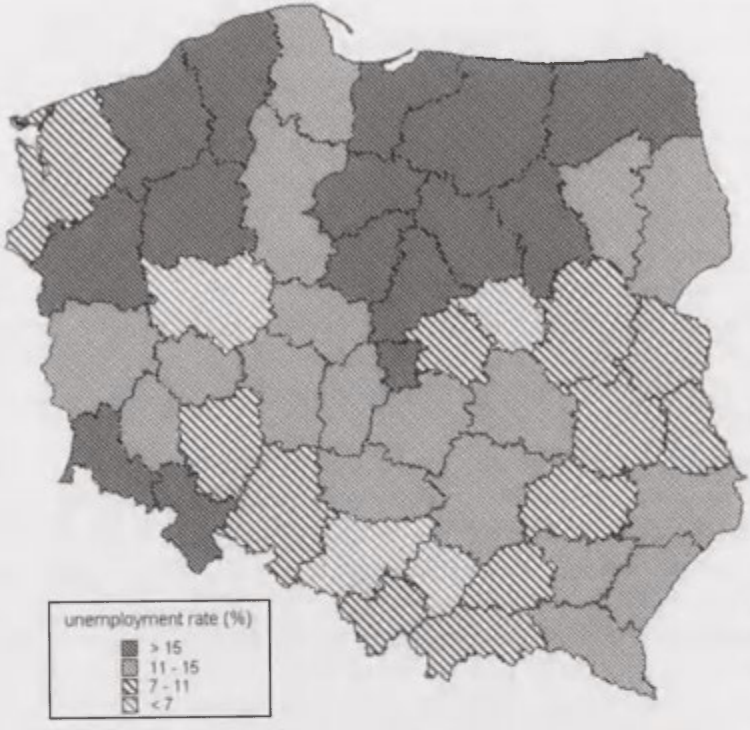


Figure 5. Regional structure of unemployment in 1991 (4th quarter).

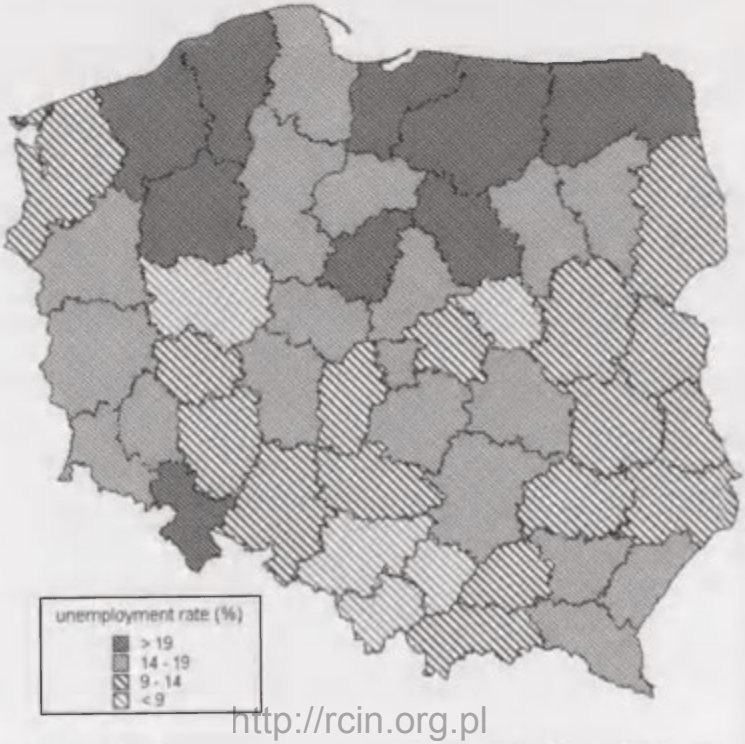


Figure 6. Regional structure of unemployment in 1992 (4th quarter).

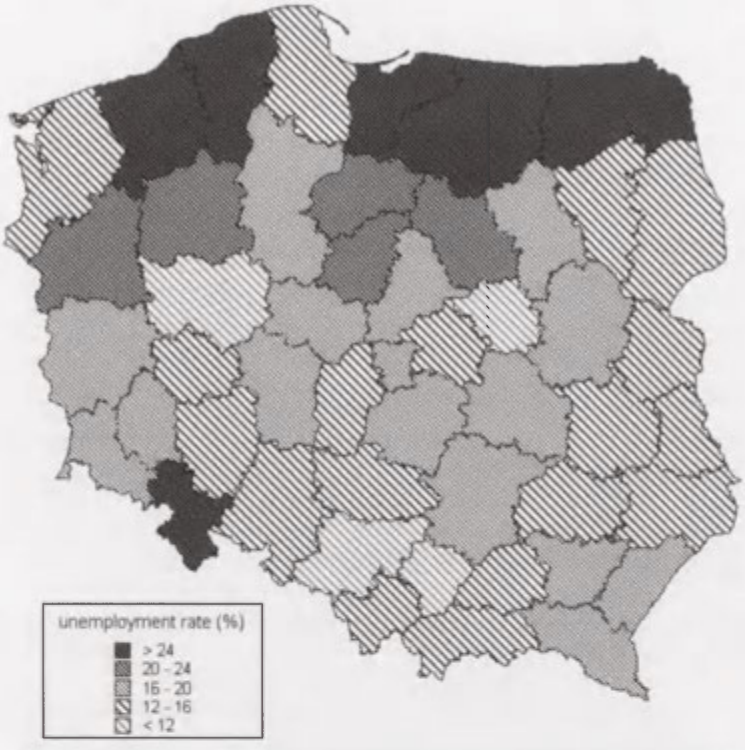


Figure 7. Regional structure of unemployment in 1994 (4th quarter).

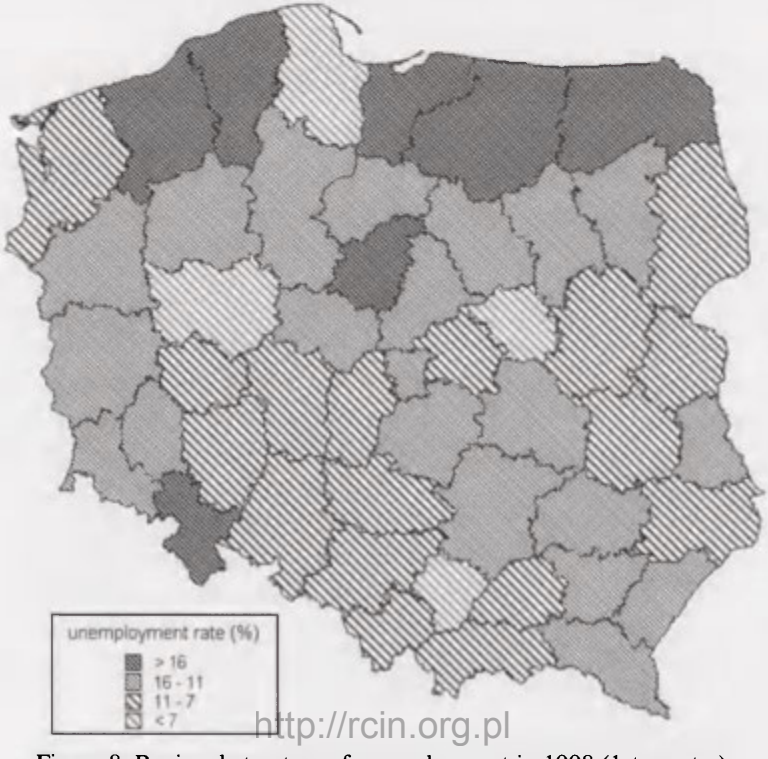


Figure 8. Regional structure of unemployment in 1998 (1st quarter).

the mean national unemployment rate was of 11%, while the patterns of highest unemployment (rates exceeding 16%) remained unchanged (Fig. 8). The regions with the lowest figures (below 7%) were joined by Bielsko-Biała and Gdańsk voivodeships. Regional contrasts were still very pronounced (3% in Warsaw voivodeship and 21% in Suwałki voivodeship).

The comparative analysis of the states of the regional structure of unemployment leads to the conclusion that in the time periods studied there were voivodeships that were permanent components of the highest regional unemployment patterns (eg. Słupsk, Olsztyn) and those that were areas of low unemployment, (e.g. Warsaw, Poznań).

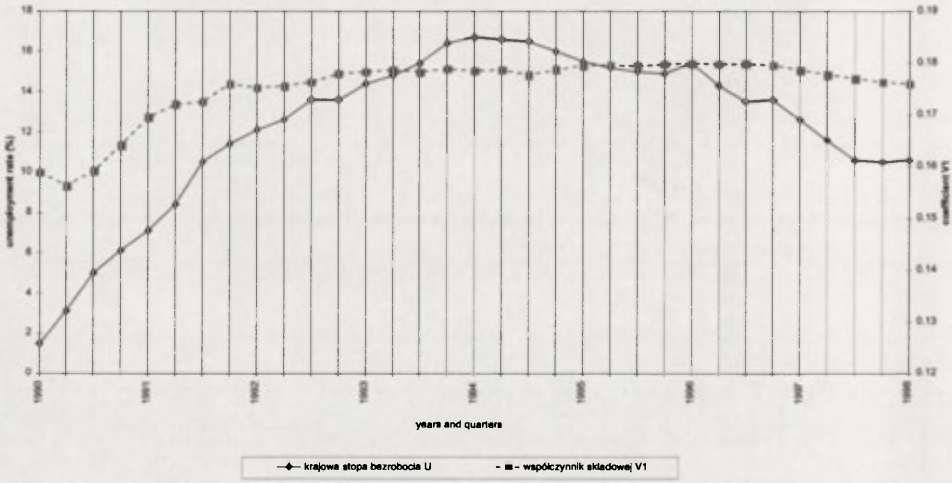


Figure 9. Coefficients of the first principal component V_1.

The other approach studies the process of temporal variations in the regional unemployment structure on the basis of an analysis of a quasi-continuous time series of the voivodeship unemployment rates over the 33 quarters of the years 1990–1998. The method employed is that of principal components analysis using the T-technique (cf. Chojnicki, Czyż 1976; Pedersen 1978). This method combines an analysis of the regional structure with an analysis of its temporal variability over the period in question, viz. 1990–1998, which allows for a generalized description of the process of change. Principal components analysis using the T-technique is carried out on an unemployment rate matrix of the order: 49 voivodeships x 33 quarters. For each quarter of the period between January 1990 and March 1998, a 49-element vector is defined representing the unemployment structure in the pattern of the 49 voivodeships. In the T-technique approach, the 33 vectors are treated as variables and the 49 regions as units. In principal components analysis correlations among the 33 regional unemployment structures only indicate their inter-temporal similarities. Two time periods (quarters)

Figure 10. Regional structure of unemployment in the years 1990–1998.

can be characterised by a similar regional unemployment structure and differ simultaneously in unemployment rates. The principal components identify time periods (quarters of specific years) significant for a description of changes in the regional unemployment structure.

The following components have been distinguished: the first principal component accounting for 91.1% of the variance, the second component (3.8%), the third (1.7%), and the remaining 30 components accounting among themselves for 3.4% of the original variance associated with the 33 vectors. Given such a distribution of the variance among the particular components, only the first one is interpreted. The curve of the coefficients of the first principal component at 33 time periods (quarters) is a curve of the similarity of the regional unemployment structure over time.

A comparison of the pattern and shape of the national unemployment curve (U) and the curve of the principal component (V_1) for the 49 regions makes it possible to distinguish three periods in the relation between the national unemployment level and the regional unemployment structure in the years 1990–1998 (Fig. 9).

The first period embraces the years 1990–1992 and is characterised by a rapid increase in the national unemployment rate (the first section of the steep curve

U) as well as by the formation of and constant changes in the regional unemployment pattern (curve V_1 shows irregular fluctuations).

The second period, from 1993 to the 1st quarter of 1994, is one of a further rise in national unemployment (the second section of the steep curve *U*) which is characterised by a stable regional unemployment structure (curve V_1 remains at the same level).

The third period covers the years 1994–1998 and displays a fall in the national unemployment rate (curve *U*) with the simultaneous persistence of differences in regional unemployment (the path of curve V_1 is equalised).

The map of values of the first principal component V_1 shows the regional unemployment structure in the years 1990–1998 (Fig. 10). Its comparison with maps of the unemployment rate distribution in 1992 (4th quarter) and 1998 (1st quarter) (Figs 6 and 8) reveals that the regional unemployment structure had formed by 1992 and has not undergone major changes since then. An analysis of the voivodeship values for component V_1 leads to the conclusion that the main role in preserving the stable regional unemployment structure in the years 1990–1998 has been played by the voivodeships with high unemployment, and by the voivodeships with low unemployment. Those voivodeships with permanently high or low unemployment rates follow their own specific pattern of transformation.

The regions of permanently high unemployment throughout the time studied were:

(1) Elbląg, Olsztyn, Suwałki, Koszalin and Słupsk voivodeships.

They are voivodeships in which, in the first period of transformation, a high incidence of closures of industrial plants coincided with a massive collapse of the state farms which were predominant there. The leaving-unsolved of the problem of the demise of state-owned farming and the underdevelopment of the tertiary sector caused the critical situation in their economies and high rural unemployment to persist. There have been cases, though, of a successful restructuring of individual large enterprises (in Elbląg and Olsztyn voivodeships).

(2) Ciechanów and Włocławek voivodeships.

These belong to less-developed regions of central Poland and are characterised by low indices of economic transformation. They are agricultural regions, poorly equipped with technical infrastructure and modern services, in which the poorly-developed industry has broken down.

(3) Wałbrzych voivodeship.

This is an old coal-mining region in which the economic transformation has been advancing despite a number of difficulties. In the years 1994–1996 the EU STRUDER Programme for regional development was implemented in the voivodeship.

The regions of permanently-low unemployment throughout the time studied have been:

(1) Warsaw, Poznań and Kraków voivodeships.

These embrace urban agglomerations with diversified, well-developed eco-

nomies and high adaptability to market conditions. They are leaders of the transformation. The relatively high qualifications of their labour force and the entrepreneurial spirit of their residents lay the foundations for the prosperity of these regions.

(2) Bielsko-Biała voivodeship.

This is a new winner of the transformation. A leading region in industrial production where substantial foreign investment contributes to the high share of the private sector in industry.

(3) Katowice voivodeship.

An old industrial, predominantly coal-mining region (the Upper Silesian Industrial District) with delayed economic, technological and ownership restructuring, and hence with a risk of high unemployment in the future.

CONCLUSIONS

The curve of national unemployment in the years 1990–1998 reflects the transformational stages in Poland.

The relations obtaining between unemployment and the stages of transformation associated with the recession and changes in Poland's economy can be used to test the validity of the hypothesis put forward by Blanchard (1997) concerning interactions occurring in the economics of post-communist transition (cf. also Bivand 1998). Blanchard (1997: 99) assumes the existence of "the interaction between growth of the new private sector, restructuring of firms in the state sector, and unemployment. Private sector growth decreases unemployment. Restructuring of state firms increases output, but increases unemployment. Unemployment, in turn, affects both private sector growth and the speed of restructuring. In the benchmark model, these effects combine to give two phases of transition". In the first, initial shocks connected with the reform lead to high unemployment. High unemployment prevents restructuring: the threat of a lay-off and becoming unemployed leave workers in state firms opposed to restructuring. During that phase, private employment creation decreases unemployment over time. As unemployment decreases, it eventually reaches a point when social resistance to enterprise restructuring becomes much weaker. There then begins a second phase of transition in which the economy proceeds along a balanced path where the private sector absorbs employment losses from restructuring. The output grows faster than in the first phase, and unemployment now remains constant.

In Poland, the development of the private sector and growth in its share of total employment have been accompanied by ongoing restructuring of the state sector, as well as an increase in investment outlays. As a result, Poland's labour market reached a turning point in 1997: the rate of unemployment declined to 10.5%. 1998 saw a combination of a fall initially – to 9.8% in June, but then

a renewed increase from September 1998 onwards, to reach 10.4% in December. This time the cause was a short-term slowdown in economic development and a slackening of the rate at which new jobs were generated. Thus, in spite of symptoms of stability, unemployment in Poland still tends to demonstrate seasonal fluctuations.

It should be noted that the change in the relations between unemployment and development in the period of the socio-economic transformation is only evident on the scale of the country. At the regional level, the change is less clear-cut and differs from region to region. In the regions with stable but high rates, unemployment is still a manifestation of a prolonged recession or poor socio-economic development.

REFERENCES

Bivand R., 1998, *Dynamic externalities and regional manufacturing development: an exploration of the Polish experience before and after 1989*, International conference on *The changing map of Europe*, The trajectory Berlin – Poznań – Warsaw, Poznań.

Blanchard O., 1997, *The economics of post-communist transition*, Clarendon Press, Oxford.

Chojnicki Z., Czyż T., 1976, *Some problems in the application of factor analysis in geography*, Geographical Analysis, 8, 416–427.

Gleave D., 1987, *Dynamics in spatial variations in unemployment*, [in:] Fisher M. M., Nijkamp P. (eds), *Regional labour markets*, North-Holland, Amsterdam, 269–288.

Pedersen P. O., 1978, *Interaction between short and long run development in regions. The case of Denmark*, Regional Studies, 12, 6, 683–700.

GEOGRAPHIA POLONICA, 73, 1, Spring 2000
PL ISSN 0016-7282

GEOGRAPHICAL STUDY ON CHANGES IN THE SPATIAL STRUCTURE OF THE DISTRIBUTION SYSTEM BY INFORMATIZATION

KENJI HASHIMOTO
Osaka Gakuin University, Kishibe Suita City, Osaka 564-8511, Japan
e-mail: kenjih@utc. osaka-gu.ac.jp

ABSTRACT: The purpose of this paper is to analyze the effect of informatization (influence of the development of information technologies) on the spatial structure of industrial organizations. The analysis was based on case studies of the distribution system in Japan. The following three viewpoints are identified for further studies: the first is a change in the spatial structure in a physical distribution system, the second is a change in the spatial structure in a sale system held by manufacturers and wholesalers, the third point is a change in information flows in the distribution system.

KEY WORDS: information technology (IT), distribution system, sales system, information network, Japan.

INTRODUCTION

There are two aspects to industrial organizations' informatization (large social transformation caused by information technology (IT)). These are the introduction of the computer and the networking of communications. Informatization in today's sense means interlinkage between these two changes, i.e. the obtainment of an environment in which computer-readable data are exchanged in real time through communication lines.

Where the effect of informatization on a business organization is concerned, many geographers are engaged in study. For example, Miles and Robins (1992) have summarized the developments in a study focused on economic aspects, according to which the theories of the socio-economic changes brought about by information technology are broadly classified into the post-industrial society theory of Bell (1973), as well as the information activities analysis of Porat (1977) and others, the theory of technological innovation cycles of the neo--Schumpetrians, and the flexible specialization theory of Piore and Sabel (1984), as well as others. Miles and Robins further point out that the latter two are closely related to the regulation theory represented by post-Fordism.

This article analyzes the effect which informatization has had on the spatial structure of industrial organizations using as a basis case studies of the distribution industry in Japan.

SPATIAL CHANGES IN CORPORATE ORGANIZATION INDUCED BY INFORMATIZATION

As pointed out by Pred (1973), a corporation consists of an in-house network of headquarters, regional offices and sales offices. Simultaneously, a number of networks are formed between corporations, typically in connection with the purchase of materials, research and development, distribution, settlement of accounts, etc. These various spatial linkages in corporate activities are necessarily accompanied by the flow of information; and with time constraints factors, such as lead time, added to this situation. Thus, the effect of the sophistication of information transmission and processing on the spatial form and base arrangement of an organization becomes quite substantial. It was from the latter half of the 1980s onwards that study of the effect of the sophistication of information technology on corporate organization was started in earnest in geography; there are four broad viewpoints in such study.

The first of these is the formation of new intra-organizational linkage through information networking. Hepworth published a series of studies on informatization and corporate organization in the latter half of the 1980s, pointing out that the sophistication of information technology in such various forms as LANs, digital networks, satellite communication, etc., causes the concentration or decentralization of offices and factories. He also suggested that internationalization of interbusiness transactions and the growth of transnational corporations gives rise to new interurban linkages. Miles and Matthews (1992) spoke of the necessity to differentiate between the argument of labor including skill, and that of hardware, stressing that there are limits to such simple categorization as the information sector when it comes to the assessing of corporate organizations into whose various divisions information instruments have penetrated.

Second, there is the promotion of interbusiness collaboration through the medium of information. For instance, the results published successively in the middle of the 1990s by Li (1995), Cecil et al. (1996), and so forth, all make it clear that interbusiness collaboration aimed at the mutual utilization of highly strategic factors such as new technology, economies of scale in production and access to markets, is structured through the medium of sophisticated information systems. Amirahmadi and Wallace (1995) have also examined the actual circumstances of interbusiness collaboration based on the sharing of information, and have classified such collaboration into: 1) the integration of production processes between different enterprises, such as the joint development of new products, 2) the formation of continuous business links between different enterprises, such as just-in-time distribution of products, and 3) complete organizational integration through the medium of a network.

Third comes organizational restructuring. There are two directions concerning the restructuring of a corporate organization based on progress in information technology. One is spatial rearrangement of the organization itself: separation of

logistics centers in distribution is a typical example. The other is rearrangement of individual posts/functions within the organization; the spinout or outsourcing of posts/functions in which exchange of information through face-to-face contact can be replaced by cyber communication has attracted special attention.

Fourth is the relationship between the sophistication of information technology and the labor market. Dunning and Norman (1987), for example, have pointed out that a highly-skilled labor force that can cope with progress in information technology is concentrated in major urban areas, and such spatially--uneven distribution of human resources may affect the structure of corporate organization and give rise to regional disparity in employment and wages. On the other hand, Goddard (1975), Richardson and Marshal (1996) etc., suggest the possiblity that new information industries (or information processing divisions of major corporations) may be located in developing areas by acquiring a certain level of information-processing capability and by training inexpensive young labor.

In corporate activities the base for the processing and exchange of information (in the broad sense of the term) is the office, and examination has also been made of the spatial effects related to changes in the form, function and locational conditions of the office that result from informatization. Typical examples are: 1) Goddard and Pye (1977); Randy and Selwood (1983); Graham and Marvin (1996); Kutay (1986) etc., who examined changes in office location, especially diffusion of offices to the suburbs, 2) Longcore and Rees (1996); Daniels and Bobe (1992) etc., who examined the functions and locational characteristics of intelligent buildings, and 3) Tauchen and Witte (1983); Kumar (1989) etc., who conducted simulation experiments concerning the optimal location of offices and advanced arguments concerning the possibility of new forms of work such as working from home or satellite offices.

DIFFUSION OF INFORMATION SYSTEMS IN THE DISTRIBUTION INDUSTRY

Generally speaking, the merits of informatization in a business organization can be classified broadly into hard merits produced directly and mainly from the development of hardware, and soft merits, whose attainment depends on software such as data analysis and manpower. In the case of the distribution industry, the former is represented by the introduction of the equipment necessary for the collection, accumulation and analysis of data, such as POS (point of sales) register handy scanners, computers, etc., the development of a communications environment that provides computer readable data, such as POS- and EOS (electric ordering system) -related data, on-line, and the automatization of distribution operations, such as automated warehouses, automated selection systems, etc. The latter is represented by the development/standardization of microdata, such as

POS- or EOS-related data, or the diffusion of source marking, and computerization, ranging from routine data processing to strategic decisionmaking in management or marketing.

The introduction of computers in the distribution industry began in earnest in the United States in the 1970s. The United States, with its vast national land area, has been leading in the area of communications networking with, in particular, major wholesale businesses starting with data exchange via satellite in the mid--1980s. With the intensification of competition in these areas, the spatial effect of informatization is becoming more and more important. For example, progress in informatization at a distribution center, such as the placing/receiving of orders on-line, the preparation of vouchers using a computer, automated pickup inside the warehouse, etc., provides for the establishment of a wider delivery area through a substantial reduction of work time required at the center and the assigning of remaining lead time to delivery. Moreover, wholesalers, who have succeeded in introducing an advanced information system, can build a strategic alliance with discount chains and promote the oligopolization of regional markets. In the retail industry also, CVSs (convenience stores) based on large-item--small-scale distribution have flourished, thereby attesting to the remarkable effect of informatization in the distribution industry (Arai and Yamada 1994).

Informatization in the distribution system has an effect on manufacturers situated "upstream" in the distribution channel as well. For example, among manufacturers who see that a smooth transfer of information and goods among factories becomes important as the structure to the spatial division of labor in production evolves, there is a more and more pronounced move to connect the headquarters and each factory with a communications network and to adopt factory location and production plans giving priority to the efficiency of the distribution of products to markets. Moreover, with a view to copying with market uncertainties, the construction of a strategic alliance relationship is promoted, with the information system supporting decisionmaking in management and promoting the introduction of networks connecting distribution channels.

In Japan also, the introduction of computers made headway in the first half of the 1980s while networking of strategic points through ISDN lines began in the latter half of that decade. Needless to say, lead time from receipt of an order to delivery and the accuracy of deliveries themselves constitute important competitive studies dealing with Japan's distribution industry and are meaningful as instances by which to examine the relationship between informatization and the spatial structure of industrial organizations.

PRESENT SITUATION OF THE DISTRIBUTION INDUSTRY IN JAPAN

The distribution industry in Japan is characterized by the fact that the share of medium and small enterprises is substantial in both retailing and wholesaling,

Table 1. Number of establishments and total sales during the year by number of employees (Retail Trade).

Number of employees	Number of establishments				Total sales during the year (bilion yen)			
	1994	1997	1997–1994	rate of change 1997/1994 (%)	1994	1997	1997–1994	rate of change 1997/1994 (%)
1–2	764.771	709.000	–55771	–7.3	13.332	12.485	–847	–6.4
3–4	370.942	350.304	–20638	–5.6	20.054	19.573	–481	–2.4
5–9	222.548	212.440	–10108	–4.5	28.997	28.556	–441	–1.5
10–19	89.618	93.455	3837	4.3	23.819	26.050	2.231	9.4
20–29	26.337	27.512	1175	4.5	12.163	13.013	850	7.0
30–49	15.655	15.801	146	0.9	11.719	12.182	463	4.0
50–99	7.191	7.922	731	10.2	9.998	11.081	1.083	10.8
100–	2.861	3.251	390	13.6	23.213	24.814	1.601	6.9
Total	1.499.923	1.419.685	–80238	–5.3	143.295	147.754	4.459	3.1

Census of Commerce (1994, 1997).

Table 2. Number of establishments and total sales during the year by number of employees (Wholesale Trade).

Number of employees (person)	Number of establishments				Total sales during the year (bilion yen)			
	1994	1997	1997–1994	rate of change 1997/1994 (%)	1994	1997	1997–1994	rate of change 1997/1994 (%)
1–2	90.382	83.085	–7297	–8.1	6.595	6.524	–71	–1.1
3–4	103.004	94.065	–8939	–8.7	19.413	18.760	–653	–3.4
5–9	120.148	109.032	–11116	–9.3	54.115	53.728	–387	–0.7
10–19	67.776	61.811	–5965	–8.8	72.162	71.725	–437	–0.6
20–29	21.296	19.266	–2030	–9.5	44.863	44.335	–528	–1.2
30–49	14.714	13.468	–1246	–8.5	54.129	51.676	–2453	–4.5
50–99	8.394	7.632	–762	–9.1	64.689	61.086	–3603	–5.6
100–	3.588	3.209	–379	–10.6	198.351	171.855	–29496	–13.4
Total	429.302	391.568	–37734	–8.8	514.317	479.689	–34628	–6.7

Census of Commerce (1994, 1997).

while the distribution route for supplying products to retailers is long. These characteristics have often given rise to criticism of inefficiencies in the distribution structure and have constituted issues both at home and abroad, along with a commercial policy that tends towards the protection of small businesses, and extreme price differentials between domestic and overseas markets. In this chapter, the present situation of the distribution industry in Japan, a subject of case studies, will be surveyed based on the Census of Commerce 1997 (Tabs. 1–3).

According to the Census, the distribution industry in Japan can be summarized as follows:

1) The absolute number of small businesses is rapidly decreasing in both

Table 3. Number of establishments and total sales during the year by type of management (Retail Trade).

Type of management	Number of establishments (rate of change 1997/1994 in %)	Total sales during the year (rate of change 1997/1994 in %)
Large-sized Department store	2.3	0.3
Department store (others)	12.3	5.3
GMS	13.5	11.2
Meduim-sized Supermarket	–22.7	–23.4
Specialty Supermarket (apparel)	46.3	29.4
Specialty Supermarket (food)	9.5	11.9
Specialty Supermarket (hausehold)	68.2	48.3
Convenience store	27.9	29.9
Convenience store (24 hours open)	52.9	52.7
Supermarket (others)	42.7	19.7
Specialty store (apparel)	–14.3	–16.3
Specialty store (food)	–12.7	15.7
Specialty store (hausehold)	–6.9	3.5
Miscelaneous retail stores	–5.3	–3.1

Census of Commerce (1994, 1997).

retailing and wholesaling. In the wholesale industry, in particular, declines in numbers have been registered for businesses of all employment levels and the downward trend is more serious than that for retailing;

2) self-service stores have generally recorded growth while the decline rate is high for small retailers whose merchandise is in competition with theirs;

3) regeneration is taking place within the self-service industry. This is best illustrated by the sluggishness in medium-sized supermarkets. That is to say, industries which have advantages directly related to the interests of consumers continue to grow; for example, large supermarkets are at an advantage in terms of selection of goods and the convenience of one-stop shopping, category killers or discount stores in terms of price, and convenience stores in terms of the merits of night operation. Many of the industries which continue to grow have commonly adopted POS and EOS data at an early stage and promoted informatization in their operation as a whole.

CHANGES IN THE DISTRIBUTION SYSTEM THROUGH INFORMATIZATION

In order to discuss the effect informatization has had on the spatial structure of the distribution industry, it is necessary to clarify the change which informatization has brought about in the distribution system itself. These changes can be summarized in the following five points:

The first point is standardization of the data used in the distribution system and the accompanying power shift to chain stores. Standardization and diffusion of information used frequently at the sites of distribution, such as sales information including POS data and ordering information including EOS data, has made the trading operation dependent on data real, diffused it as an industry-wide standard and enhanced the strength of the chain stores which are at the starting point of information. A series of changes in trade terms, such as sophistication of merchandising in chain stores, reduction of lead time, curtailment of delivery lots and reduction of delivery prices, reflect such a power shift. This is expected to have a wide-ranging effect on the distribution system as a whole, from store--opening strategies to production/delivery.

The second point is integration of distribution centers and restructuring of wholesale trade. Routine operations in wholesale trade consist of the receipt of orders, pickup, delivery and settlement of accounts; this is a cyclical operation based on orders placed by retailers where the time required for a cycle is defined by the lead time specified in the contract. Integration of distribution centers is effected on the assumption of a realization of economies of scale that enlarges the delivery area and increases the delivery volume of each distribution center. However, demand for a reduction in lead time, notably from chain stores, acts in the direction of contraction of the delivery area. In contrast, progress in informatization is expected to reconcile two contradicting propositions, a reduction of lead time and integration of distribution centers, by reducing work time at distribution centers through automatization of order receipt and pickup. On the other hand, in wholesale trade, sophistication of the information system is always necessary in step with the system of the clients, that is, chain stores or manufacturers, and firms which are unable to cope with such advancement in terms of capital or technology are likely to be excluded from the trade system, thereby accelerating restructuring in wholesale trade.

The third point is a change in the production/delivery structure among manufacturers of consumer goods. With regard to factory location by manufactures of consumer goods that have largely homogeneous markets around the country, it has been deemed rational to divide the country into several delivery regions and, as a rule, to complete production within each region. However, as multi-type production becomes more common in response to the needs of the market, a more rational system in terms of total cost is often achieved through wide-area delivery of products manufactured intensively at specific factories rather than at various factories which each have production lines for a variety of products. If such thinking also becomes common in respect to products typically produced in consumption areas, such as confectionery, beverages, beer etc., then relocation of production centers will likely become common with a view to rationalizing delivery.

The fourth point is a change in the location or functions of the operation centers (sales offices) of manufacturers, something that reflects the informatiza-

tion of trade. Japan's national manufacturers with nationwide branch networks generally adopt a hierarchical structure under which the headquarters is at the top with several regional offices placed under it and sales offices under them. As a rule, such manufacturers adopt the territory system under which a sales office is responsible for sales operations vis-à-vis all retailers/wholesalers in the area under its charge. However, there are limits for an individual salesman of a sales office to conduct all dealings with chain stores at an advanced level of business computerization. As a measure to cope with such a situation, an idea has emerged of limiting the number of clients with whom the salesmen of sales offices have dealings to those with a low level of informatization, while dealings with retailers, that require high-level information, such as chain stores or the headquarters of CVSs, are delegated to regional offices or specialized divisions set up at headquarters (Fig. 1). Such a change in sales strategies means a shift from a simplistic vertical territory system to a multi-stratified system under which different departments conduct dealings related to different types of business, and relocation of sales offices and a substantial transfer of authority are expected accordingly.

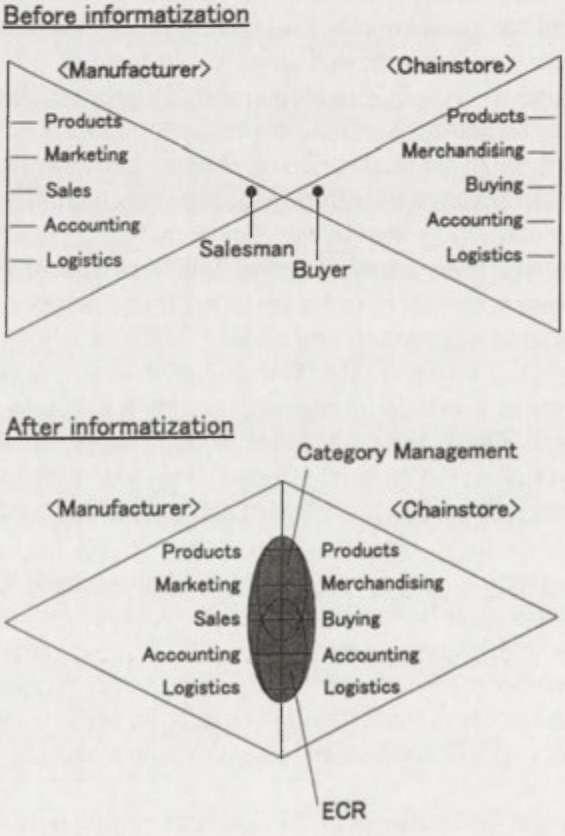


Figure 1. Change in sales strategies between manufacturers and chainstores.

The fifth point is progress in collaboration between businesses based on the networking of businesses. Today, a high-level communications infrastructure has been developed and the development of a computer system and data standardization have progressed in many companies. This will facilitate business collaboration such as division of operations and joint use of databases. Moreover, a reduction of communications costs will reduce distance constraints in information exchange and promote outsourcing and integration of information-processing operation as well as the decentralization of business to regions where personnel expenses are low. Such progress in the networking of businesses will not only enhance economies of scale, reduce the inventory risk and reduce personnel expenses in distribution, but also serve as a trigger for the structuring of a strategic alliance through the sharing of an information system. In turn, this is highly likely to introduce a change in the structure of the conventional open distribution system.

CHANGES IN THE SPATIAL STRUCTURE OF THE DISTRIBUTION SYSTEM BROUGHT ON BY INFORMATIZATION

The issues involved here are classified into three groups: the distribution system with delivery centers at its core; the sales system with regional offices/sales offices at its core; and the spatial pattern of information flows. I will now attempt to clarify the effect of the progress in informatization on each spatial structure.

CHANGES IN THE SPATIAL STRUCTURE IN A DISTRIBUTION SYSTEM

Changes in the function and location of the distribution system accompanying progress in informatization can be broadly assigned to three stages (Fig. 2). First, with respect to the location of distribution centers before the progress in informatization, small centers, where various commercial distribution functions coexisted with the sales functions at the core, were densely located, notably in urban centers, since the need to concentrate stock was small in reflection of the small number of brands handled and long lead times.

Then, as informatization of the distribution system made headway and the number of brands increased and lead times became shorter, it became necessary to integrate centers to reduce stock and relocate distribution centers to the suburbs, which have fewer congestion problems, so as to enhance delivery efficiency. Moreover, progress in informatization facilitated automatization of operation at centers and raised the ratio of delivery time to lead time, so distribution centers came to be located along main highways in the suburbs as large distribu-

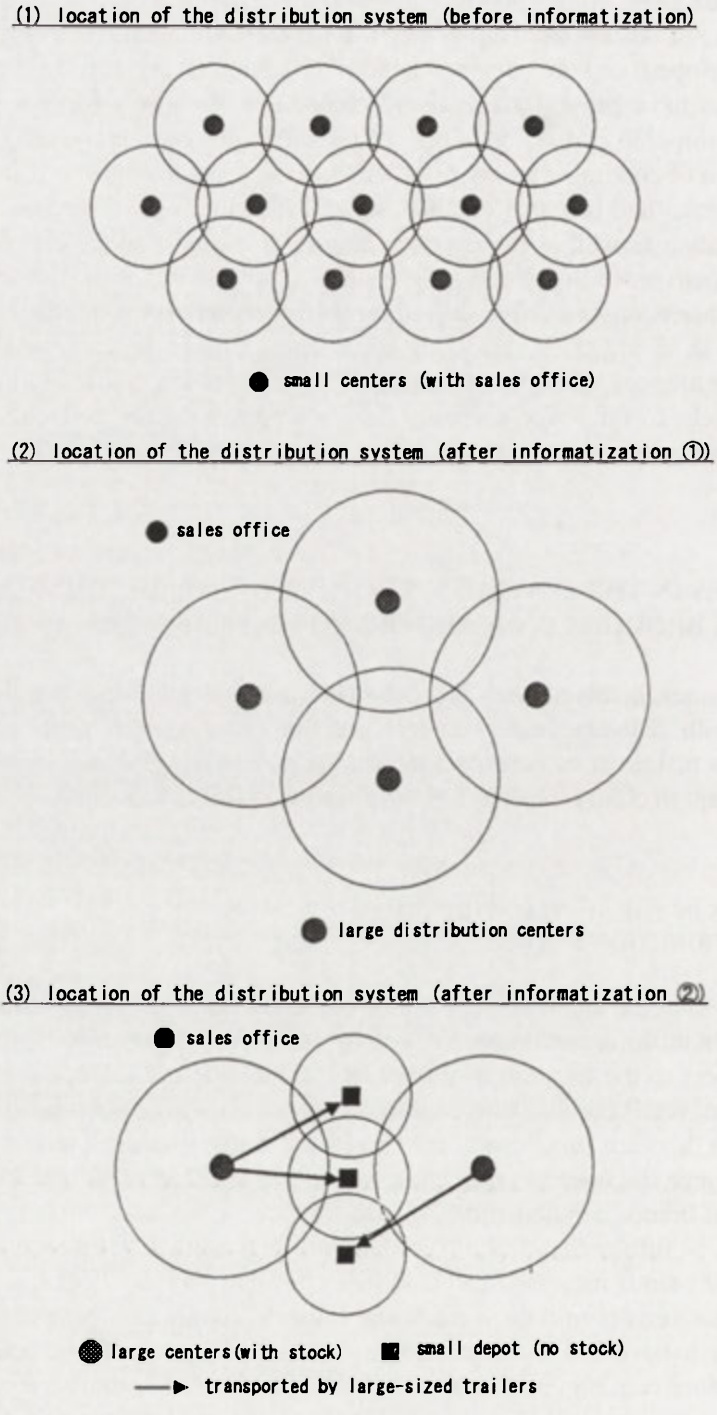


Figure 2. Changes in the spatial structure in a distribution system.

tion centers that perform wide-area delivery of goods from integrated stock, in separation from commercial distribution centers oriented toward urban centers.

However, this brings about a decline in efficiency for the large distribution centers to be charged with the delivery of products to all parts of the delivery area expanded through informatization. The delivery time for supermarkets is generally limited to the two to three hours between the time when clerks report to work and the time when the store opens. So, it is necessary to complete delivery within a short time using a large number of medium-sized trucks (with a load capacity of approx. 2 tons). As a result, stores located on the periphery of the distribution area have small transport volumes relative to the transport distance and it becomes difficult for them to absorb delivery costs. A system that combines a large distribution center and depots has been devised as a solution to this problem. Under this system, depots which do not have stock are placed on the periphery of the delivery area, and several medium-sized trucks are kept at each of them. In the middle of the night, goods are transported from the large distribution center to each depot in large-sized trailers at a stretch, and on the following morning, terminal delivery is made from the depots to stores in medium-sized trucks. This method manages both the merit of concentration of stock in the form of large distribution centers and efficient terminal delivery. At present, this system is adopted by Kao Corp., a major manufacturer of toiletry products, and Ryoshoku Ltd., a major food wholesaler, and can be considered the most efficient form of distribution system under which high-frequency-small-lot distribution is effected over a wide delivery area.

CHANGES IN THE SPATIAL STRUCTURE IN A SALES SYSTEM

The most significant effect that informatization of distribution has had on the spatial structure of a sales system is seen in the restructuring effected in nationwide sales systems held by national manufacturers and wholesalers. To take national manufacturers as an example, the conventional sales system had a hierarchical structure consisting of 6 to 10 regional offices set up under headquarters and 30 to 60 sales offices under a regional office, particularly in prefectural capitals. Under this system, sales offices, performing the main sales activities, were each responsible for a different territory which did not overlap with any other in terms of space and was charged with the sales activities for all the clients in the territory. In contrast to this vertical regional sales system, the sales system, after the progress in informatization, has been transformed to a multi-stratified system that selects a base for each business dealing from among sales offices, regional offices and headquarters according to the degree of strategic importance reflected in such factors as the type and size of the client's business (Fig. 3). Such a shift brought about the following changes in the spatial structure of a sales system.

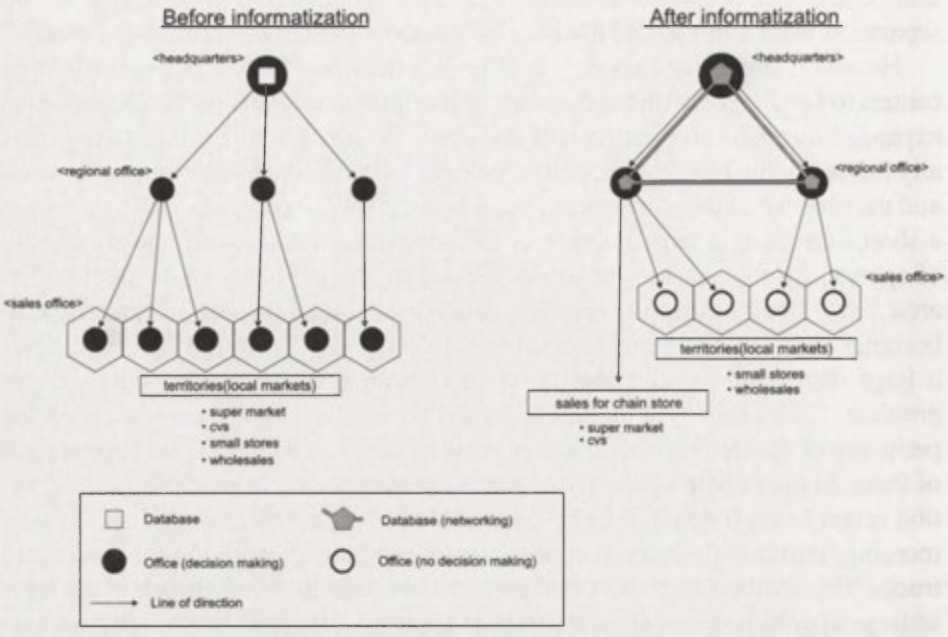


Figure 3. Changes in the spatial structure in a sales system.

First, with the regional office coming to assume responsibility for conducting dealings with chain stores where the importance of data is emphasized, the communications network between headquarters and each regional office has been upgraded and a new information system has been built under which the sharing of databases is possible between the two. Moreover, qualitative information of sales data exists, such as daily sales bulletins, records of success cases of sales activities and negotiation records, that is exchanged mainly via e-mail, in addition to quantitative information (such as POS and EOS data), that is updated periodically by headquarters. Each regional office stores information transmitted from the sales offices under its jurisdiction through a mail server installed at each and such information is in turn transmitted to the sales offices under the jurisdiction of other regional offices through on-line networks connecting the regional offices. Thus, with the diffusion of new methods of communication, such as e-mail, regional offices play an increasingly important role as nodes of in-house information.

Secondly, rapid reorganization is in progress at terminal organizations such as sales offices. To be specific, the operations of a sales office located within the area of a one-day trip from a regional office are integrated in the operations of the regional office and, even in a case where more than one sales office is located in the one-day-trip area, many of their operations are subject to a reduction in functions. Various factors, such as the expansion of spheres of activities due to

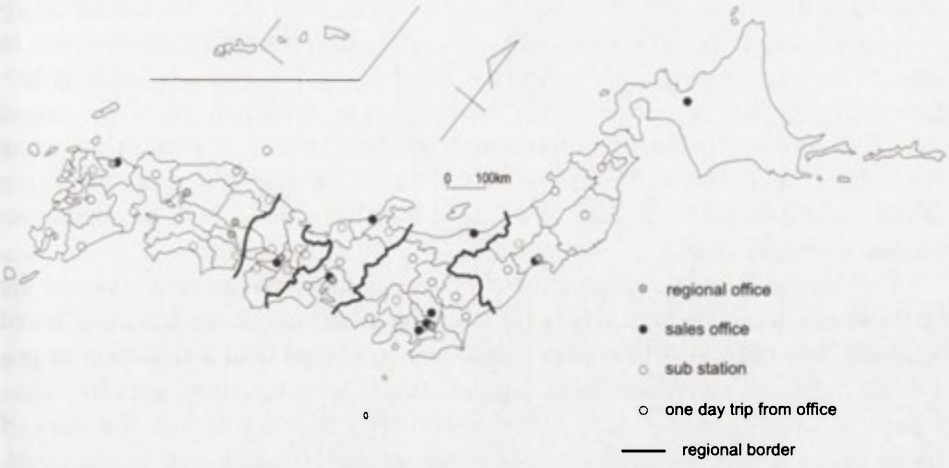


Figure 4. Reorganization of regional and sales offices (SUNSTAR Corp.).

the development of expressway networks and cost reductions through reorganization in indirect sectors including general and personnel affairs divisions are involved in such progress in rationalization. Essentially, however, such progress should probably be regarded as reflecting a relative decline in the sales functions assigned to sales offices (Fig. 4).

Finally, one can regard the tendency for high-level decisionmaking in sales activities to be integrated at a regional office as a transitional form of system innovation in large enterprises. That is, when an enterprise intends to effect innovation of a technological nature, it tends to integrate such a function which has been dispersed in terminal organizations in the key organization, such as the headquarters, thereby restructuring the system, and then to redeploy the function in subordinate organizations in respect of those operations for which a manual or software has been introduced successfully. Informatization in sales activities is also contained within such a cycle and is likely to be delegated to sales offices eventually. However, both the function and spatial location of sales offices will at that time have gone through a process of substantial integration and are expected to differ greatly from their present function and spatial location.

CHANGES IN INFORMATION FLOWS

Information flows in the distribution system are mainly those concerned with ordering or billing. As analog information such as vouchers, they have so far been instrumental in controlling the transfer of goods and the flow of payments. However, progress in informatization has introduced three changes in information flows in the distribution system.

The first is standardization and digitization of data. Their realization has

enabled real-time on-line exchange of information, with the convenience of information storage in databases and simultaneity in transmitting information to different centers being greatly enhanced. This has resulted in a substantial reduction in information exchange time in the trade cycle and expansion of the area of jurisdiction of each center as represented by the distribution area; at the same time, there has been a decline in relative terms in the importance of such operations as information exchange or protocol transformation which has been performed by wholesalers.

The second change concerns speed-up in communications networks and the introduction of private lines. These two technological innovations have facilitated long-distance transfers of raw data, and have combined with a reduction in private-line charges, to reduce the constraints from communication costs for information processing substantially, thereby increasing in relative terms the share of personnel expenses and hardware cost in the entire information processing cost. As a result, integration has made headway in such routine operations as the arrangement of databases and update of product masters with a view to reducing the cost of hardware. Moreover, since variation in communication costs depending on distance has disappeared due to the use of private lines, information processing operations can to expected to relocate to peripheral areas in pursuit of lower personnel expenses.

The third change is an increase in non-fixed, ad hoc information exchange, as represented by e-mail. E-mail has become common as a business tool not only because its degree of simultaneity in communication is high and secondary processing of information is easy with it, but also because e-mail can be used as a communication means that overcomes the organizational barrier. Such cross-organizational exchange of information is essential in proceeding with business negotiations with major chain stores which deploy offices, such as regional offices and sales offices, beyond the geographical territories of manufacturers. This can be said to show the limits of regionally vertical sales organizations in terms of information flows.

Such a series of changes not only transforms the spatial pattern of information flows but also has a considerable effect on the functions and location of centers serving as nodes. At chain stores which are at a high level of informatization, measures are taken to reduce lead time by concentrating ordering information from each store at the information center at headquarters, classifying it by manufacturer and then transmitting it simultaneously to the delivery centers of wholesalers and manufacturers. Manufacturers, on the other hand, are accelerating the trend towards the direct conducting of operations vis-à-vis chain stores without going through wholesalers, and the function of wholesalers as nodes in information flows is compelled to decline as far as dealings with chain stores are concerned. As a result, the raison d'ètre is gradually declining for wholesalers which are inferior in their ability to cope with high-frequency-small-lot delivery and those which cannot add intellectual values in the form of operational proposals based on data, for example.

CONCLUSION

The operation centers or distribution centers dealt with in this study have all contributed to the development or concentration of tertiary industries in major local cities, particularly prefectural capitals. In particular, the mass production/mass distribution system that presupposes mass consumption has contributed to the strengthening of the urban economy by causing the bases of manufacturers and wholesalers to be located in local cities of a size above a certain level and bringing in employment and consumption there.

However, though the merits which informatization has introduced in various aspects have all accelerated the concentration of functions to a small number of core cities, they have caused the functions of relatively small local cities to decline substantially. Such a tendency toward bipolarization is irreversible and expected to accelerate in coming years. In this sense, the progress in informatization in the distribution system may be considered a change that puts an end, for the time being, to the growth which has been brought about with local cities becoming distribution centers for national brands.

Moreover, although this study does not deal with it directly, a similar upward integration appears to be occurring within prefectures and municipalities. For example, reduction of wholesalers or integration of distribution centers means a loss of sales in the wholesale trade in the smaller cities in a prefecture. Among local cities, prefectural capitals may be capable of maintaining their central roles dependent on prefectural governments through, inter alia, concentration of government offices and financial institutions including regional banks. On the other hand, if cities other than prefectural capitals are to continue to grow, it will be necessary to cast around for new growth strategies.

First, it will be necessary to aggregate industries which are essentially incapable of mass production and require a certain level of labor force in each local market, that is, industries such as services, advertisement, software development, events, etc. Second comes the attracting of on-line information processing business, that is, telebanking by banks, management of ordering data and product masters for chain stores, direct marketing, etc. In this case, measures must be taken to develop the infrastructure necessary for attracting businesses, such as high-grade communications networks, including a network of ISDN lines, with the administration taking the lead. Third comes intensive location of distribution centers taking advantage of changes in traffic conditions. The reorganization of the distribution system as a whole in step with progress in informatization is having an effect on the growth strategy itself in local cities, and it may be necessary for the direction of a new growth strategy to be re-examined in the light of overall regional policy.

REFERENCES

Amirahmadi H., Wallace C., 1995, *Information technology, the organization of production and regional development*, Environment and Planning, A, 27, 1745–1775.

Arai Y., Yamada H., 1994, *Development of Convenience Store Systems in Japan, 1970s–1980s*, [in:] *Comparative Study on Retail Trade Tradition and Innovation*, Ryutsu Keizai University, 117–126.

Bell D., 1973, *The coming of post-industrial society: a venture in social forecasting*, Basic Books.

Cecil B., Green M., Mcnaughton R., 1996, *Patterns of strategic alliances among information technology firms in the United States*, Geografiska Annaler, B, 78, 129–146.

Daniels P. W., Bobe J. M., 1992, *Office building in the City of London: a decade of change*, Area, 24, 253–258.

Dunning J. H., Norman G., 1987, *The location choice of offices of international companies*, Environment and Planning A, 19, 613–631.

Goddard J. B., 1975, *Office location in urban and regional development*, Oxford University Press.

Goddard J. B., Pye R., 1977, *Telecommunications and office location*, Regional Studies, 11, 19–30.

Graham S., Marvin S., 1996, *Telecommunications and the city: electronic spaces, urban places*, Routledge.

Kumar A., 1989, *Impact of technological developments on urban form and travel behavior*, Regional Studies, 24, 137–148.

Kutay A., 1986, *Effects of telecommunications technology on office location*, Urban Geography, 7, 243–257.

Li F., 1995, *The geography of business information*, John Wiley & Sons.

Longcore T. R., Rees P. W, 1996, *Information technology and downtown restructuring: the case of New York City's financial district*, Urban Geography, 17, 354–372.

Miles I., Matthews M., 1992, *Information technology and the information economy*, [in:] Robins K. (ed.), *Understanding information: business, technology and geography*, Belhaven Press, 91–112.

Miles I., Robins K., 1992, Making sense of information, [in:] Robins K. [ed.], *Understanding information: business, technology and geography*, Belhaven Press, 1–26.

Piore M. J., Sabel C. F., 1984, *The second industrial divide: possibilities for prosperity*, Basic Books.

Porat M. U., 1977, *The information economy: definition and measurement*, U. S. Department of Commerce.

Pred A., 1973, *The Growth and Development of Systems of Cities in Advanced Economies*, Lund Studies in Geography, B, 38, 9–82.

Randy W., Selwood D., 1983, *Office location and the density-distance*, Urban Geography, 4, 302–316.

Richardson R., Marshall J. N., 1996, *The growth of telephone call centres in peripheral areas of Britain: evidence from Tyne and Wear*, Area, 28, 308–317.

Tauchen H., Witte A. D., 1983, *An equilibrium model of office location and contact patterns*, Environment and Planning A, 15, 1311–1326.

GEOGRAPHIA POLONICA, 73, 1, Spring 2000
PL ISSN 0016-7282

THE DEMOGRAPHIC STATUS OF AND PERSPECTIVES FOR THE RUSSIAN FEDERATION

PIOTR EBERHARDT
Institute of Geography and Spatial Organization, Twarda 51/55, 00-818 Warszawa, Poland
e-mail: mar.kow@twarda.pan.pl

ABSTRACT: The paper presents the demographic problems of the Russian Federation. The demographic developments in Russia to date are first shown. In spite of two World Wars and the enormous losses incurred by the Russian population, this population continued to feature strong demographic dynamics until the disintegration of the Soviet Union began. The turning point came during the 1990s, as difficult politico-economic conditions brought an abrupt decline in the birthrate along with an increase in mortality. The population then started to decrease. The subsequent part of the paper presents the demographic forecasts for Russia. They are pessimistic indeed, showing that the 50 years to come will see the population in Russia decreasing steadily. The author outlines the socio-economic consequences of this already persistent phenomenon, and shows that the demographic decline in Russia may have quite fundamental geopolitical repercussions. Indeed, this question has become the focus of a very broad scientific discussion in Russia, so the paper presents the views of numerous Russian demographers and politologists. Many of these opinions are of an alarmist character and sere to inflame the political atmosphere in Russia.

KEY WORDS: population of Russian Federation, demographic forecasts, socio-economic consequences, demography and geopolitics.

THE DEMOGRAPHIC DEVELOPMENT OF RUSSIA IN THE 20TH CENTURY

In accordance with statistics as of July 1st, 1998, the territory of the Russian Federation is inhabited by 146.5 million persons. Thus, with respect to population, Russia takes sixth place in the world, after the People's Republic of China (1208 million), India (916 million), the United States (260 million), Indonesia (193 million), and Brazil (159.1 million). Before its disintegration in 1990 the Soviet Union was inhabited by 285 million people and had occupied third place in the world, being ahead of the United States, Indonesia and Brazil. With respect to surface area the Russian Federation is still by far the largest country in the world (at more than 17 million sq. km), ahead of Canada (10 million sq. km), People's Republic of China (9.6 million sq. km) and the United States (9.4 million sq. km). Simultaneously, Russian Federation is one of the least densely-populated countries of the world (with an average of 9 persons per sq. km).

In spite of the two World Wars, the Bolshevik revolution, and the losses

caused by the repressive nature of the political system, Russian lands featured a high demographic dynamic during the 20th century. The first population census carried out there in 1897 showed population of 67.5 million for the present territory of the Russian Federation. Just prior to the outbreak of World War I, by 1913, the population had increased to 89.9 million. Immediately after the cessation of military operations, in 1922, the population was at 87.7 million. Russian territories were characterized by a very high natural increase in the inter-War period and, notwithstanding the losses incurred by forced collectivisation, the then territory of the Russian Federation was inhabited by 110.1 million persons on the eve of Hitler's invasion.[1] Enormous losses caused by World War II caused the population in Russia to decrease. By 1950 there were still only 101.4 million inhabitants, but the following decades saw significant demographic growth take place (Tab. 1).

In the forty post-War years 1950–1990, the population of Russia increased by almost 50 million. Deep structural transformations were taking place simultaneously, due to urbanization and the growth of towns, especially of large agglomerations. Owing to the mass inflow of rural population to towns, the urban population increased by approximately 66 million, while that in rural areas decreased by close to 20 million. Consequently, the percentage share taken by the rural population, still at 56.9% in 1950, had dropped to 26.1% by 1991, i.e. by 30.8 percentage points.

The decisive role in population dynamics has been played by the natural increase. The role of external migrations in increasing the population of Russia has been marginal. Some well-defined trends appeared in the domain of natural increase. Thus, the natural increase was high in 1950s and 1960s, but then started to decrease gradually. Nevertheless, until 1990 there was a clear surplus of birth rates over rates of mortality (Tab. 2).

When analysing the problems of the demographic development of Russia to date one should not overlook the essential differentiation in conditions set by the natural environment. The enormous area of the country, stretching from the Baltic Sea to the Pacific, as well as its specific features, related to the historical process by which human settlement develops, ensure that the spatial distribution of population is very uneven. More than three quarters of the whole population are concentrated in the European part of Russia, and more than half on 13% of the whole area of the country (this part bounded by the river Volga to the East and North as well as by the Moscow – St. Petersburg railway line). In turn, the processes of industrialization caused the concentration of population in the industrial centers and regions, and consequently entailed depopulation processes in rural areas. All the conditions indicated have influenced the status of population distribution, characterized by a high degree of polarization and spatial unevenness. This will have an important bearing on future developments, because the

[1] The paper had been written before the outbreak of the new Chechenyan war of 1999–2000.

Table 1. Changes in population on the territory of the Russian Federation, 1950–1995.

Years	Population in millions	Urban population		Rural population	
		in millions	in %	in millions	in %
1950	101.4	43.7	43.1	57.7	56.9
1959	117.5	61.6	52.4	55.9	47.6
1970	130.1	81.0	62.2	49.1	37.8
1979	137.6	95.4	69.3	42.2	30.7
1989	147.4	108.4	73.5	39.0	26.5
1991	148.5	109.8	73.9	38.7	26.1
1995	147.6	107.7	73.0	39.9	27.0

Source: *Naseleniye SSSR 1987. Statisticheskii Sbornik* (*Population of the USSR 1987. Statistical Collection*), Moskva 1988, p. 8; *Rossiyskii Statisticheskii Ezhegodnik* (*Russian Statistical Yearbook*), Moskva 1996.

Table 2. Natural demographic processes on the territory of the Russian Federation, 1960–1990.

	Births		Deaths		Natural increase	
	in thousands	in ‰	in thousands	in ‰	in thousands	in ‰
1960	2782.4	23.2	886.1	7.4	1896.3	15.8
1970	1903.7	14.6	1131.2	8.7	772.5	5.9
1980	2202.8	15.9	1525.8	11.0	677.0	4.9
1990	1988.9	13.4	1656.0	11.2	332.9	2.2

Source: *Naseleniye SSSR 1987. Statisticheskii Sbornik* (*Population of the USSR 1987. Statistical Collection*), Moskva 1988, pp. 112, 129; *Rossiiskaya Federatsiya v 1992 godu*, [in:] *Statisticheskii Ezhegodnik* (*Statistical Yearbook*), Moskva 1993.

historically-shaped settlement concentrations are characterized by a high degree of persistence.

The disintegration of the Soviet Union, which caused a serious socio-economic crisis, also influenced the course of the demographic processes in an essential manner. Specifically it had an impact on the breakdown in the demographic development of the country. This appeared in a spectacular way in 1992, yet in 1991 there had still been a small but positive natural increase. However the subsequent year saw the death rate start to exceed of the birth rate and the situation worsened as years went by. A natural decrease which amounted to 1.5‰ in 1992 had grown by the year 1995 to 5.7‰. In peace-time conditions such a natural decrease is a singular phenomenon in the modern world. This statement is especially true in view of the speed of demographic transformations, and of the nature of the evolution on both the birth side (decrease) and the death side (increase). Analogous phenomena were observed in urban as well as rural areas (Tab. 3).

Presentation of the scale of natural loss in absolute quantities shows the magnitude of the phenomenon considered in a more tangible manner. The number of births decreased between 1990 and 1995 by 625.1 thousand. Simultaneously, the number of deaths increased by 547.8 thousand. Consequently, in place of

Table 3. Natural demographic processes on the territory of the Russian Federation, 1990–1995 (per thousand inhabitants).

Year	Births			Deaths			Natural increase & decrease		
	total	in towns	in countryside	total	in towns	in countryside	total	in towns	in countryside
1990	13.4	12.7	15.5	11.2	10.4	13.3	2.2	2.3	2.2
1991	12.1	11.2	14.5	11.4	10.6	13.4	0.7	0.6	1.1
1992	10.7	9.8	13.2	12.2	11.5	14.1	–1.5	–1.7	–0.9
1993	9.4	8.6	11.5	14.5	13.8	16.4	–5.1	–5.2	–4.9
1994	9.6	8.9	11.4	15.7	15.0	17.5	–6.1	–6.1	–6.1
1995	9.3	8.6	10.9	15.0	14.4	16.5	–5.7	–5.8	–5.6
1996	8.9	8.3	10.4	14.2	13.4	16.2	–5.3	–5.1	–5.8

Source: *Rossiiskii Statisticheskii Ezhegodnik* (*Russian Statistical Yearbook*), Moskva 1996, pp. 50–51.

a natural increase which had been at the level of 332.9 thousand until 1990, there was already a natural loss of 840.9 thousand by 1995. These disadvantageous demographic processes have been growing stronger and have encompassed both towns and villages. (In 1997 the decrease amounted to 759.9 thousand.)

Due to the abrupt decrease in the birth rate and jump-like increase in the death rate, the natural loss of population in the Russian Federation amounted to 2,703.2 thousand persons in just four years (1992–1995). The depopulation processes took place over virtually the whole territory of the Federation, although there was definite spatial differentiation. The drop in the birthrate appeared simultaneously with the decrease in life expectancy. We are thus dealing with a full-fledged demographic breakdown, which will certainly have very important future consequences in all domains of life. In spite of the thus significant difference between death and birth rates there was no automatic reflection in the form of a decrease in the absolute population size in the Russian Federation. During the whole first half of the 1990s, that population remained relatively stable at of 147–148 million. This is the result of positive net inmigration. The positive balance of external migrations for Russia was 164 thousand in 1990, 51.6 thousand in 1991, 176.1 thousand in 1992, 430.1 thousand in 1993, 810 thousand in 1994, and 502.2 thousand in 1995. Thus, over six years (1990–1995) 2134 thousand more people came to Russia than left the country. In 1996 the positive net inmigration balance was 343.6 thousand, while in 1997 it was 359 thousand persons (Tab. 4).

This large positive migration balance result from the mass return to their home country of the Russians residing previously in the various Soviet republics which have recently gained independence. After the republics gained independence the Russians living there became a largely unnecessary or undesired ethnic minority. This was compounded by ethnic conflicts or misunderstandings. The effect was the *exodus* of Russians from the Caucasian and Asian republics. Thus, the positive migration balance for Russia has so far been compensating for the natural loss. As the role of migration decreases and the tendencies to date in

Table 4. Natural demographic processes on the territory of the Russian Federation, 1990–1995 (in '000).

Year	Births			Deaths			Natural increase & decrease		
	total	in towns	in countryside	total	in towns	in countryside	total	in towns	in countryside
1990	1988.9	1386.3	602.6	1656.0	1140.6	514.4	332.9	245.7	87.2
1991	1794.6	1230.5	546.4	1690.7	1168.9	521.8	103.9	61.6	42.3
1992	1587.6	1068.3	519.3	1807.4	1254.8	552.6	–219.8	–186.5	–33.3
1993	1379.0	930.5	448.5	2129.3	1488.3	641.0	–750.3	–557.8	–192.5
1994	1408.2	960.4	447.8	2301.4	1615.0	686.4	–893.2	–654.6	–238.6
1995	1363.8	933.3	430.3	2203.8	1554.2	649.6	–840.0	–620.7	–219.3
1996	1304.6	897.9	406.7	2082.2	1446.0	636.2	–777.6	–548.1	–229.5

Source: *Rossiiskii Statisticheskii Ezhegodnik* (*Russian Statistical Yearbook*), Moskva 1996, pp. 50–51.

natural processes are maintained, an unavoidable demographic regression will occur. The data as of 1995 bear witness to this. During that year 840 thousand more persons died in Russia than were born. The positive migration balance was at 502.2 thousand. Thus, in spite of a mass inflow of people the population in Russia still decreased by 337.8 thousand. The persistence of the recent, demographically-disadvantageous tendencies constitutes an essential threat to the future of Russia.

DEMOGRAPHIC FORECASTS FOR RUSSIA

The State Statistical Committee of the Russian Federation recently published a forecast for population development in the country until 2010. In view of the importance of this document, the presentation of its main statements is certainly worthwhile. The forecasts are based upon three different scenarios, namely:

– the intermediate scenario – constructed on the assumption of a slow overcoming of the economic crisis; this will induce a decline in the death rate and simultaneously make it possible for families to fully realize their reproductive intentions, in accordance with the microcensus conducted in Russia in 1994;

– the pessimistic scenario – based upon the assumption of continued crisis; it refers to an extension of the upward trend to death rates initiated in the mid 1960s and existing currently; while birthrates are assumed to remain at a very low level;

– the optimistic scenario – which assumes that it will be possible to overcome the crisis relatively quickly; in the nearest years there will therefore be a decrease in death rates; while birth numbers will initially return to the levels from just before the crisis, and thereafter show a long-term trend of increase.

The demographic forecast thus prepared accounts for the migrations pertaining to the Russian Federation. A positive net migration balance with the former Soviet republics is envisaged. Inmigration will be exceeding outmigration until

Table 5. Forecast of natural demographic processes in the Russian Federation until 2010 (per thousand).

Year	Pessimistic scenario			Intermediate scenario			Optimistic scenario		
	births	deaths	natural loss	births	deaths	natural loss	births	deaths	natural loss or increase
1995*	9.3	15.0	–5.7	9.3	15.0	–5.7	9.3	15.0	–5.7
2000	8.8	16.0	–7.2	11.2	15.6	–4.4	12.7	14.8	–2.1
2005	9.3	17.5	–8.2	11.8	16.2	–4.4	14.2	14.0	0.2
2010	8.8	18.8	–10.0	11.4	15.9	–4.5	14.2	14.2	0.5

* Preliminary data.

Source: Y. Yurkov, 1997, *Prognoz chislennosti naseleniya Rossiiskoi Federatsii do 2010 goda* (*Forecast of population for the Russian Federation up to 2010*), Voprosy Ekonomiki, no. 7, pp. 103–106.

the year 2000 by 400 thousand persons a year. Then, over five years, the positive migration balance will amount to 200 thousand, before decreasing further to 150,000 a year in the period between 2005 and 2010. It is predicted that the positive net balance of migrations with the former Soviet republics will amount to 3.6 million persons up to 2010. On the other hand, the balance of migrations with respect to other countries will be negative at the level of some 1 million persons. Hence, the total migration balance will be positive in the period of 1996–2010, and will entail a migration-related population increase of approximately 2.6 million people.

As noted, the forecasts concerning the natural processes take the form of three scenarios. In the first, pessimistic one, births will remain at a very low level of 8–9‰, while deaths will display an upward trend, to attain the rate of 18.8‰ in the last year of the forecast (2010) Consequently, the natural increase will be negative, falling from –5.7‰ (in 1995) to –10.0‰ (2010).

The intermediate scenario envisages a slight increase in the birthrate and a continuation of death rates at virtually unchanged levels. There will be a marginal improvement in natural increase (decrease), from –5.7‰ to –4.5‰.

In turn, the optimistic scenario assumes, on the one hand, an increase in childbearing, with the birthrate growing to 14‰, and, on the other, the appearance of a minor tendency towards decreased mortality. Consequently, a small positive natural increase will appear after 2000. These three variants are presented in terms of rates in Table 5.

These three variants are also defined in absolute numbers. In the pessimistic scenario the natural loss will increase steadily. After the year 2000 it will exceed 1 million persons a year. The so-called intermediate variant also predicts a significant natural loss of population, but with a slight tendency towards improvement. Yet, even in this variant the population of the Russian Federation will be decreasing by more than 600 thousand persons per annum after 2000. Only in the so-called optimistic variant will the end of the forecasting period see the popula-

Table 6. Forecast of natural demographic processes in the Russian Federation until 2010 (in '000).

Year	Pessimistic scenario			Intermediate scenario			Optimistic scenario		
	births	beaths	natural loss	births	deaths	natural loss	births	deaths	natural loss or increase
1995*	1364.0	2204.0	–840.0	1364.0	2204.0	–840.0	1364.0	2204.0	–840.0
2000	1270.0	2314.0	–1044.0	1627.0	2279.0	–652.0	1658.0	2167.0	–309.0
2005	1295.0	2447.0	–1152.0	1686.0	2321.0	–635.0	2088.0	2047.0	+42.0
2010	1188.0	2521.0	–1333.0	1596.0	2239.0	–643.0	2097.0	2013.0	+84.0

* Actual (initial) data.

Source: Y. Yurkov, 1997, *Prognoz chislennosti naseleniya Rossiiskoi Federatsii do 2010 goda*" (*Forecast of population for the Russian Federation up to 2010*), Voprosy Ekonomiki, no. 7, pp. 103–106.

Table 7. Forecast of the population numbers in the Russian Federation until 2010 (in million).

Year	Pessimistic scenario			Intermediate scenario			Optimistic scenario		
	population	of which		population	of which		population	of which	
		in towns	in countryside		in towns	in countryside		in towns	in countryside
1995*	147.6	107.7	39.9	147.6	107.7	39.9	147.6	107.7	39.9
2000	144.3	105.4	38.9	145.5	106.3	39.2	146.5	106.9	39.9
2005	134.4	102.3	37.1	143.0	104.9	38.1	146.7	107.3	39.4
2010	133.6	98.9	34.7	140.3	103.9	36.4	147.6	108.8	39.0

* Actual (initial) data.

Source: as to Table 5.

tion of the Russian Federation start to increase somewhat. The subsequent table (Tab. 6) shows the results of the forecast in absolute numbers.

After external migrations, births and deaths had been accounted for, population forecasts for the Russian Federation at three time points (years 2000, 2005 and 2010) were put together, with a distinction made between urban and rural populations (Tab. 7).

In accordance with the so-called pessimistic variant the population of Russia will have decreased by 14 million by the year 2010 (by 8.8 million in towns and 5.2 million in the countryside). It goes without saying that such a significant decrease in population coupled with simultaneous worsening of the age structure (a significant increase in the share of the elderly and a major drop in the share of young people) can be referred to as a demographic catastrophe for Russia in the intermediate scenario. The size of the urban population will decrease by 3.8 million up to the year 2010, that is by 3.5%. This will be the consequence of an actual natural decrease at a level of 8 million offset by a net migrational increase of approximately 4.2 million persons. The decrease in the rural population

Table 8. Forecast of population changes in the Russian Federation in percentage points in the period 1996–2010.

Year	Pessimistic scenario			Intermediate scenario			Optimistic scenario		
	population	of which		population	of which		population	of which	
		in towns	in countryside		in towns	in countryside		in towns	in countryside
1996–2000	–3.3	–2.3	–1.0	–2.1	–1.4	–0.7	–1.1	–0.8	–0.3
2001–2005	–4.9	–3.1	–1.8	–2.5	–1.4	–1.1	+0.2	+0.4	–0.2
2006–2010	–5.8	–3.4	–2.4	–2.4	–1.0	–1.7	+0.9	+1.3	–0.4

Source: S. Vykolova, G. Vasina, 1997, *Prognoz chislennosti naseleniya Rossiiskoi Federatsii do 2010 goda* (*Forecast of population for the Russian Federation up to 2010*), Voprosy Ekonomiki no. 1, p. 56.

amounting to 3.5 million will be the result of a natural decrease of 1.9 million and a migrational outflow of 1.6 million.

If the pessimistic variant comes into being, the population in Russia will have decreased by 3.3% in absolute terms between 1996 and 2000, by 4.9% during the subsequent five years and then by 5.8% in the period 2006–2010. In this particular scenario the depopulation processes will be more intensive in towns than in the countryside. Likewise, the intermediate scenario also envisages a significant decrease in the population expressed in percentage terms. There will be only smaller differences between urban and rural areas. It is only the so called optimistic scenario that forecasts an improvement in the situation after the year 2000 (Tab. 8).

The demographic processes taking place in the case of the optimistic scenario are the necessary condition for stabilisation of the population. Yet, the realization of this scenario is not very probable as achievement of a situation with simultaneously lowering mortality and increased fertility is not realistic in the coming years.

The indicators of the natural demographic processes within the area of Russia will be quite differentiated spatially. Without going into details concerning this problem area, since this would require a separate report, we may mention that the lowest birthrate will be noted over the whole period analysed in Moscow, in St. Petersburg, and in the districts of Moscow, Tula, Penza, Jaroslav, Murmansk, and Vladimir. Relatively higher fertility will be observed in the areas inhabited by the non-Russian islamic population (the Ingush Republic or Dagestan). A similarly high fertility characterizes the Chechen Republic, now belonging in the Russian Federation only formally.

There is no certain method of saying which of the three scenarios will most closely approximate the actual future developments of the coming dozen or so years. Demographic forecasts very often date quickly when unexpected circumstances occur to change the nature of trends to date.

Nevertheless, taking this reservation into account we can try to depict the most probable scenario of future demographic processes. When approaching the questions considered here from a more intuitive and somewhat less hard-fact angle we can, it seems – as already indicated – reject the so-called optimistic scenario, and perhaps also the one situated between the optimistic and pessimistic scenarios. This intermediate scenario can only be fulfilled if there is a significant improvement in the political, economic and social conditions in Russia, something which does not appear very probable at present. The most probable scenario would thus seem to be the pessimistic one. Under the conditions of this scenario the population in Russia will decrease by 10 to 13 million people between the years 2000 and 2010, due to natural processes alone.

When persisting over a longer time period low birthrates bring an impairment in the age structure of the population. First, the number of children decreases, followed by the number of children and the young, and then by the number of people entering the professionally-active age. All this takes place along with a constant increase in the population of post-productive age. The upkeep of the ever growing population group of retirement age is an economic problem weighing ever more heavily on society. This is a difficult issue to solve even in the rich countries. For Russia, a country with a relatively low level of national income, this might have very deep social repercussions.

It is necessary to take into consideration one further possibility which cannot be excluded *a priori*. This is that mortality is largely determined by natural aspects, related to life expectancy, and a sudden outside intervention can hardly change anything here. On the other hand, though, there is at least a theoretical possibility of exerting a radical influence on birthrate. A rigorous ban on abortion in conditions of adequate pro-natalistic propaganda might lead to an abrupt increase in the number of births. However, this would require the formation of an effective repressive administrative and legal system, which cannot be thought likely to exist in a state other than one having totalitarian features.

The probability of the appearance of this kind of social conditioning may be considered as nothing more than a hypothesis, though it certainly cannot be overlooked. Russia always had ambitions of empire and an expansionist attitude. The awareness of its demographic degradation has a significant psychological meaning for Russian society. It evokes in Russia a longing for a strong authority which will not only put the economic situation in order, but also create the conditions allowing for increase of population. On the other hand, it is known that a democratic lawful state is not capable of taking effective steps in this domain. Only a strong dictatorial power is capable of taking the respective decisions which will be carried out effectively in practice. Will this type of social atmosphere not be conducive to the strengthening of authoritarian tendencies, allowing for violations of human rights in the name of attaining national goals? This problem is extremely important, insofar as it concerns the future of Russia.

The different variants of future demographic development in the Russian

Table 9. Population of the fourteen most populous countries in 1995, 2010 and 2050.

1995		2010		2050	
country	population in '000	country	population in '000	country	population in '000
China	1208.0	China	1388.5	India	1529.0
India	916.0	India	1189.1	China	1478.0
United States	260.0	United States	297.5	United States	349.0
Indonesia	193.0	Indonesia	239.6	Pakistan	346.0
Brazil	159.1	Pakistan	210.1	Indonesia	312.0
Russia	**147.3**	Brazil	199.3	Nigeria	244.0
Pakistan	126.3	Nigeria	168.4	Brazil	244.0
Japan	124.8	Bangladesh	162.5	Bangladesh	213.0
Bangladesh	117.8	**Russia***	**133.6**	Ethiopia	170.0
Nigeria	107.9	Japan	127.1	Congo	160.0
Mexico	91.9	Mexico	117.6	Mexico	147.0
Germany	81.1	Vietnam	98.5	Philippines	131.0
Vietnam	72.5	Iran	95.2	Vietnam	127.0
Philippines	66.2	Philippines	88.1	**Russia**	**122.0**

* The result shown for Russia was for the pessimistic scenario.

Source: *The Economist*, 1997; *Świat w liczbach*, Warszawa 1997, p. 14. United Nations Population Division. World Population Prospects 1998.

Federation presented here, can be compared to the expected demographic futures of the world's other largest countries. Thus, on the basis of the forecast from UN reports, as well as those published in the annual edition of *The Economist*, we can set the future demographic situation of Russia against the respective forecasts for the ten most populous countries in the world for the years 2010 and 2050 (Tab. 9).

The forecast presented implies that the coming dozen or so years will see the demographic ranking of Russia fall from 6th to 9th place. By 2050 there will be 13 countries whose populations are greater than Russia's. This is connected with the fact that Russia will be the only one of the countries here listed featuring a demographic decline. This is a cause of quite common concern among Russian politicians and demographers.

AN ASSESSMENT OF THE DEMOGRAPHIC SITUATION IN THE RUSSIAN LITERATURE OF THE SUBJECT

Russian scientific journals and other media publish much on the subject. Some publications tend to present the situation in terms of a catastrophic vision, indicating the negative consequences of the lowering demographic potential of Russia. One can mention here the paper by A. Antonov, who analyses the causes behind the drop in birthrate and puts forward a hypothesis concerning the crisis

of the Russian family and the disappearance of the bonds linking people, leading altogether to the self-destruction of society. In the opinion of that author a further continuation of the already deeply-rooted tendency towards limited childbearing will ultimately result in the deformation of the age structure of the population and an excess share of the elderly, deprived of adequate care, with all the related negative consequences.

This question is presented in an even more dramatic manner by V. Kozlov. This author maintains that a persistence of the high mortality rates, especially when coupled with low fertility, will bring consequences leading to national catastrophe. Such opinions have also been presented by A. Koreshkin, N. Titov, as well as by A. P. Sercova and G. Karpov. Their articles were published in Moscow in 1997 in a book devoted to the demographic future of Russia. Similarly, a number of political columnists try to demonstrate that the existing demographic trends will result in territorial losses and a return of a Russia with the boundaries of the Grand Duchy of Muscovy.

A more balanced, though also pessimistic, opinion was presented by I. Bestuzhev-Lada, an author who considers the contemporary social situation of Russia in a very deep manner. He maintains that the demographic breakdown in Russia has multiple causes of an objective nature. His diagnosis starts with the following statements: "For several years already deaths in Russia are more numerous than births. Population started to decrease first by tens of thousands, then by hundreds, and now [the decrease] nears [one] million [a year], with even bigger losses in the future being a realistic threat. Many demographers sound the alarm in view of this situation. They accept the assumption that continuation of these tendencies during the coming decades will result in the decrease of the population by 1/4 and that the decrease will yet go on at an important rate. It is not difficult to imagine what this can mean for us given that in the neighbouring Asian countries a further rapid population growth will continue."

Despite this catastrophic assessment the author approaches the question in a more balanced or even philosophical manner. He nevertheless treats it as an enormously important problem, requiring comprehensive countermeasures to be provided by creating an alternative civilisation in Russia, in which every family will have assured economic security, psychological stabilisation, adequate ecological conditions, etc. According to this author, Russia is capable of creating such a modern civilisation, based upon new moral principles. Only in such a case will "extinction of Russia cease to be a reality, become just a myth, one of the nightmares of the third 'smuta' ['turmoil'] period in the history of Russia (the first one having taken place at the turn of the 17th century and the second in the years 1917–1920)".

The approach here presented, quite typical for the Russian mentality, has more of a mythological-ideological character than a pragmatic one. Yet, it indicates the degree of understanding by Russians of the situation in which their country finds itself today.

GEOPOLITICAL CONSEQUENCES OF THE DEMOGRAPHIC DECLINE IN RUSSIA

The depopulation processes in Russia will also have essential geopolitical consequences. The economic crisis, linked with the demographic decline, will cause weakening of Russia's position power on the Asian continent. Simultaneously, the processes of demographic, and presumably also political, strengthening of the islamic countries located to the south of Russia, will go on. The former Soviet republics of Asia, as well as Turkey, Iran, or Pakistan, will dispose of an ever growing demographic potential. Then, in the Far East, depopulating Siberia borders with China, featuring an enormous pool of the young labour force. It is hard to forecast whether these negative geographical disproportions will ultimately bring a threat to the southern borders of Russia, but usually this type of disproportion sooner or later brings about mass migrations of population seeking free areas for colonization and settlement. The geopolitical consequences resulting from changes in demographic potentials are hard to predict, but considerations of such consequences are not quite irrelevant. This is well illustrated by the fact that they are the subject of detailed politological studies conducted by Russian experts.

The conclusions from such studies are usually quite alarming for Russia, since they indicate the future difficulties with maintaining complete domination over the vast border territories of southern and eastern Russia. In the middle of the 21st century a depopulating Russia will become an attractive place of permanent residence for the numerous inhabitants of the demographically dynamic Asia. The political conditions in Russia and the attitude of Russians with respect to the incomers will decide whether Russia becomes a state wide open to millions of migrants, or whether it will try to close its boundaries tightly in the face of the inflow from outside of people representing various cultures and languages. Resolution of this dilemma will be very difficult, as is demonstrated by the experience of western Europe. A solution to this question should of necessity be expected in conditions of future globalization. It is hard to say whether the large-scale migration processes will take place in an evolutionary and conflict-free manner, or whether they will bring about tensions as well as military conflicts.

It must be noted clearly that a similar demographic situation has been formed in other post-Soviet republics (except for the Muslim countries). The demographic decline has affected Ukraine, Belarus, Latvia and Estonia. The consequences of this state of affairs will only be uncovered 20 to 30 years from now, yet the evident indications are that the whole of eastern Europe, and above all Russia, face essential changes resulting from the strengthening of the disadvantageous demographic processes. They may bring truly far-reaching consequences in the future.

Studies of problems connected with the demographic perspectives for Russia are of political and cognitive significance. Hence, analyses of this question may

form the basis for considerations referring to the future role of Russia as a superpower and the future course of the internal socio-economic situation in this country. For it is known that demographic conditions are quite often the ground for definite political moves. Knowledge of the future demographic image of a country usually constitutes a starting point for all kinds of scientific analyses related to both society and the economy.

REFERENCES

Antonov A., 1995, *Depopulyatsiya i krizis sem'i v postsovetskoi Rossii: Kto vinovat i chto delat' (Depopulation and family crisis in post-Soviet Russia: Who is guilty and what to do)*, Vestnik MGU, Sotsiologiya i politologiya, seriya, 18, 2, 74–86.

Antonova O., Brui B., Gorbunova T., Nikitina S., 1997, *O demograficheskoi situatsii v Rossiiskoi Federatsii na period strukturnoi perestroiki i stabilizatsii ekonomiki, (On demographic situation of Russian Federation during the years of structural transformation and economic stabilization)*, Voprosy Statistiki, 12, 53–56.

Aralovets N., 1995, *Poteri naseleniya sovetskogo obshchestva v 1930-e gody: probemy, istochniki, metody izuchenuya v otechestvennoi istoriografii, (Population losses of soviet society in the 1930-ties. Problems, reasons, methods of investigations in national historiography)*, Otechestvennaya Istoriya, 1, 135–146.

Arutyunyan Yu., Drobizheva L., 1992, *Russkie v raspodayushchemsya Soyuze, (The Russians in disintegrating Soviet Union)*, Otechestvennaya Istoriya, 3, 3–15.

Bestuzhev-Lada I., 1996, *Vymorochnost' naseleniya Rossii: mif ili realnost', (Dying out of the population of Russia: myth or reality)*, Otechestvennaya Istoriya, 4, 113–128.

Brade I., Schulze M., 1997, *Russland – aktuell*, Institut für Länderkunde, Leipzig.

Brui B., 1997, *Vliyanie sotsial'nykh i mediko-demograficheskikh faktorov na uroven' rozhdaemosti v Rossii, (An influence of social and medical factors on the rate of birth in Russia)*, Voprosy Statistiki, 12, 72–76.

Eberhardt P., 1998, *Mniejszość rosyjska w republikach postsowieckich, (Russian ethnic minority in the post-soviet republics)*, Przegląd Wschodni, 3, 19, 489–507.

Khorev V., 1995, *Demograficheskaya katastrofa v Rossii, (Demographic catastrophe in Russia)*, Izvest'iya Russkogo Geograficheskogo Obshchest'va, 127, 49–56.

Koreshkin A., 1997, *Problemy priostanovki depopulyatsii russkogo etnosa, (Problems of permanence of depopulation of the Russian nation)*, [in:] *Natsional'nye interesy russkogo naroda i demograficheskaya situatsiya v Rossii (National interests of Russian nation and demographical situation of Russia)*, Shchtrikhton, Moskva, 119–126.

Kozlov V., 1989, *O ljudskikh poteryakh Sovetskogo Sojuza v velikoi otechestvennoi voine 1941–1945 godov, (On demographic losses of the Soviet Union during the war for fatherland 1941–1945)*, Istoriya SSSR, 8, 132–139.

Kozlov V., 1995, *Vymiranie russkikh: istoriko-demograficheskii krizis ili katastrofa, (Dying out of the Russians: a historico-demographical crisis or a catastrophe*, Vestnik RAN, 9, 771–781.

Maryański A., Mróz G., Szot Z., Troc M., 1996, *Geografia gospodarcza Rosji, (Economic geography of Russia)*, Państwowe Wydawnictwo Ekonomiczne, Warszawa (in Polish).

Naselenie SSSR 1987, Statisticheskii Sbornik, (Population of USSR in 1987, Statistical Collection), 1988, Gosudarstvennyi Komitet SSSR po Statistike informatsionno-izdateljskii tsentr, Finansy i Statistika, Moskva, 8 (in Russian).

Naselenie Rossii 1995, (Population of Russia in 1995), 1996, Gosudarstvennyi Komitet SSSR po Statistike informatsionno-izdateljskii tsentr, Moskva.

Rossiiskii Statisticheskii Ezhegodnik 1991–1999, (*Russian Statistical Yearbook 1991–1999*), 1999, Gosudarstvennyi Komitet SSSR po Statistike informatsionno-izdateljskii tsentr, Moskva.

Sercova A. P., Karpov G., 1997, *Tragicheskaya demografiya smutnogo vremeni*, (*Tragic demography of the years of depression*), [in:] *Natsional'nye interesy russkogo naroda i demograficheskaya situatsiya v Rossii* (*National interests of Russian nation and demographical situation of Russia*), Shchtrikhton, Moskva, 133–138.

Świat w liczbach 1997, (Pocket World in Figures 1997 edition), The Economist, 1997, Wydawnictwo Studio EMKA, Warszawa.

Titov N., 1997, *Blagosostoyanie i demograficheskue protsessy v sovremennoi Rossii*, (*Living conditions and demographic processes in contemporary Russia*), [in:] *Natsional'nye interesy russkogo naroda i demograficheskaya situatsiya v Rossii* (*National interests of Russian nation and demographical situation of Russia*), Shchtrikhton, Moskva, 127–132.

Vykolova S., Vasina G., 1997, *Prognoz chislennosti naseleniya Rossiiskoi Federatsii do 2010 goda, (Prognosis of the population size of the Russian Federation until 2010)*, Voprosy Ekonomiki, 1, 53–58.

World Population Prospects, 1998, United Nations Population Division, New York.

Yurkov Yu., 1997, *Prognoz chislennosti naseleniya Rossiiskoi Federatsii do 2010 goda, (Forecast of the population for the Russian Federation up to 2010)*, Voprosy Ekonomiki, 7, 103–106.

GEOGRAPHIA POLONICA, 73, 1, Spring 2000
PL ISSN 0016-7282

MAJOR DISPUTES ON THE METHODS OF CALCULATING LOCAL EXPENDITURE NEEDS. A CASE STUDY FOR LARGE GERMAN CITIES

CHANG WOON NAM
Ifo Institute for Economic Research; Poschinger Straße 5, 81679 Munich, Germany
e-mail: nam@ifo.de

ABSTRACT: Due mainly to rapidly diversifying and expanding urban activities in an enlarged Europe and also partly to the continued decrease in local tax revenue in recent years, large German cities are currently suffering from the lack of financial means to cover increasing expenditure needs. In this context, they argue that the German-style method of intergovernmental fiscal transfer and resource allocation applied in practice do not adequately consider the particular socio-economic conditions in large cities. This study compares the German way of measuring local expenditure needs with other empirical methods implemented in countries like Norway and Japan and introduces its reform alternatives. Additionally the study examines some significant socio-economic determinants from the point of view of the large German cities.

KEY WORDS: local expenditure needs, Germany, large cities, socio-economic determinants for urban development.

INTRODUCTION

Large cities in Germany and other European countries have recently been confronted with a number of similar problems which are caused by the spatial concentration of industries and services and the disproportionate share of the social, economic, health and environmental ills as well as by the provision of additional new city-specific infrastructure that is often better adapted to newly--emerging economic activities (European Institute of Urban Affairs 1992). Above all, the rapid changes in the economic framework and the increased competition which are led by German unification, the completion and expansion of the EU Single Market, the introduction of the free market system in eastern Europe and the globalisation of markets have had more significant impacts on the economic performance and development of large German cities than is the case for other areas, mainly due to the concentration of economic activities in these urban areas as well as their function as major exporters of goods and services and (national and international) gateways to the hinterlands. Apart from high-technology manufacturing and information-processing activities, large European cities have been competing against each other to capture international management and other

modern service functions, and, consequently, to secure sustainable growth in the future. Although large German cities located in the European core region are generally favoured in this competition, some improvement of 'hard' and 'soft' infrastructure, the ensurement of diversified economic structure as well as the provision of new commercial sites are in urgent need if location attractiveness and quality are to be enhanced.

At present, a number of large German cities (including the transformation cities like Berlin, Dresden and Leipzig) are still suffering from economic restructuring. In addition, the expansion of services has been unable to absorb the workers displaced from the industrial sector on a full-scale, partly due to the differences in required qualification and skills for the jobs. The internal-migration flow has made this type of structural unemployment more serious and persistent in the urban labour market. In the western part of Germany, employment growth has recently been distinctly higher in the surroundings of a large city than in the city itself (sub-urbanisation): Cities have not only had slower growth in the service sector but also suffered more seriously from the restructuring process in the manufacturing sector. Additionally, a number of industrial and service firms have already left their traditional city sites for the surroundings, mainly due to high rental costs and/or restricted possibilities of expansion. In the process of sub-urbanisation there has also been selective migration of higher-income households from city centres to suburbs.

In this context, large German cities complain that they are suffering from a serious lack of the necessary financial means to cover the increasing expenditure needs and to meet current challenges. At the same time, they argue that the German-style methods of intergovernmental fiscal transfer and resource-allocation as applied in practice do not adequately consider the particular socio-economic conditions in the large cities mentioned above. Such fiscal stress usually arises either when costs of providing local services increase faster than revenues needed to finance them, or when, given the costs of public-service provision, local government revenues are constrained by a declining economic base which reduces taxable resources (Pommerehne and Krebs 1991). In recent years, large German cities have continually lost part of their major income – the revenue from trade taxes (*Gewerbesteuer*) and income taxes – due particularly to the ongoing sub-urbanisation process and the recent economic recession. Besides, as a direct consequence of unification, grants from upper-level governments to the cities which are made in the context of the German intergovernmental resource-allocation system (*Finanzausgleich*) have been markedly reduced. On the other hand, a large city has to maintain intact its function as a central place, providing infrastructure, educational and cultural services also for the surroundings. This unfavourable financial condition has always been cited in political discussion as one of the major reasons and impediments why a number of urban development plans could not be (or are not being) carried out immediately and/or the construction of some infrastructure is hindered significantly (Nam 1997).

Following the so-called equalisation objectives, one easily tends to argue that those municipalities and cities with greater spending needs automatically require more financial support from central or upper-level government. Yet the sum of grants to municipalities should basically be deduced from a comparison of their (existing and/or anticipated) 'true' expenditure needs with local fiscal capacity from their own resources such as local tax revenue (Nam 1993). To be sure, the expenditure behaviour of municipalities is also, to a great extent, influenced by their present fiscal capacity as well as by the size of local debts. In providing infrastructure, local governments tend to consider an increase in local taxes, especially when intergovernmental grants to municipalities do not adequately compensate for the fiscal stress caused by large expenditure needs, and/or, when the total sum of local debts has already reached a level that should not be exceeded. In addition, a region well-equipped with infrastructure is obviously more attractive for investors and firms looking for a new location, while the increase in local taxes immediately means the loss of regional competitiveness (Nam, Nerb and Russ 1990). In a case in which the additional provision of infrastructure will mainly be financed by higher local taxes, local governments should also be well aware of such a trade-off relationship and their short- as well as long-term effects on the local economy.

It is a widely accepted fact that 'needs' are subjective and cannot, therefore, be quantified easily. Nevertheless, a large number of methods of resource transfer between different levels of government and of measurement methods of local expenditure needs have been developed in the past and also implemented in many industrialised countries, ranging from exclusively political to straightforward statistical ones. Furthermore, there have been serious and ambitious efforts to devise as well as improve these methodologies, so that the so-called true financial needs of municipalities could be measured in a more effective and systematic way. In particular, the dispute about the relationship between the per-capita level of local expenditure needs and the size of the municipality (in terms of population) has not yet been settled fully (Nam 1993).

This study compares the German way of measuring local-expenditure needs with other empirical methods implemented in Norway and Japan, and introduces their reform alternatives. Its main purpose is not to find a common (or best) measuring method for local government needs in the different cases, but mainly to signal an urgent need for further empirical research in the field of intergovernmental resource allocation. Additionally, the study investigates some large-city specific socio-economic determinants and major current challenges which should be better reflected in the measurement of their expenditure needs in Germany.

THE TRADITIONAL GERMAN WAY OF MEASURING LOCAL EXPENDITURE NEEDS

It is generally said that Germany has the best-developed system of urban (and regional) government among countries in Europe, which makes discretionary

resources available at a local level and favours local autonomy and identity (Cheshire 1992).[1] The lowest political level in the German federal system consists of municipalities and cities (*Gemeinden*) which, together with *Regierungsbezirke* (administrative regions), the hierarchic level next to the *Land* (state), have considerable administrative autonomy. Major areas of responsibility which belong to such local government levels are housing and construction, school education, water and energy supply, waste disposal services and the provision of transport systems (Albers et al. 1980).[2]

With regard to Germany's intergovernmental resource allocation system, the expenditure needs of a municipality are generally calculated on the basis of the so-called '*Haupt- und Nebenansätze*' (major and additional weighting systems). In the major weighting system (*Hauptansatz*) it has traditionally been assumed that the local expenditure needs per inhabitant increase with the size of the municipality measured in terms of number of inhabitants – the so-called Popitz rule (Hansmeyer and Kops 1985). Apart from the extra costs caused by a higher density of population in cases where infrastructure and social welfare activities are provided, this rule emphasises the fact that a central-place ('*zentraler Ort*'), or a large city, usually provides a wide range of public services also for the surrounding regions and smaller municipalities. In other words, there is a need to include all the beneficiaries of the local infrastructure provided by a municipality, which are both the number of inhabitants in the municipality and those in surrounding areas which also have access to the same public utility. According to this weighting system, the number of inhabitants of a municipality is multiplied by a rate (larger than 1), which gradually increases with population size. For example, in Bavaria the major weight amounts for a municipality with less than 5,000 inhabitants, to 108%, for one with 10,000 to 115%, with 25,000 to 125%, with 50,000 to 135%, with 100,000 to 140%, with 250,000 to 145% and with 500,000 to 150% (Parsche and Steinherr 1995).[3]

In practice, the major weighting system is supplemented by additional ones (Nebenansätze). For instance, the additional weights for population growth and social welfare increase the major weights by a certain number of percentage points. Two other typical examples of the additional weighting systems are:

– spatial development weights (*Raumordnungsansätze*) for centres in Rhineland-Palatinate and Saarland, which are determined by their central position and are mainly characterised by the spill-over of their functions to the surroundings,[4]

[1] This type of federal system has hindered the increase in economic and infrastructural disparities among German regions, cities and municipalities, although there have always been some significant differences between cities and other regions in the country (Mottershead 1994).

[2] The *Regierungsbezirke* are mainly responsible for providing public services of a supra-municipal nature such as mental hospitals, etc.

[3] However, in the two states Schleswig-Holstein and Rhineland-Palatinate each inhabitant is equally weighted regardless of the size of municipality (Steinherr and Parsche 1998).

[4] In Schleswig-Holstein the centres are promoted by additional lump-sum grants, which are

– different weights classified according to the number of pupils (*Schüleransatze*) in North Rhine-Westphalia and Rhineland-Palatinate.

In order to measure the expenditure needs of a municipality, its weighted number of inhabitants, calculated on the basis of major and additional weighting systems, is multiplied by the amount of basic needs per inhabitant which is the same state-wide. This basic sum is fixed annually so that the total amount of financial grants provided by the state government can be fully distributed to the municipalities, whose spending needs are greater than the sum of their local tax revenue and other local income (including charges). Subsequently, the amount of grants going to a municipality is determined by the differences between calculated expenditure needs and the fiscal capacity of the municipality (Parsche and Steinherr 1995; Steinherr and Parsche 1998).[5]

The present measuring system is criticised by a large number of economists, primarily for being too simple (Kuhn 1995). In addition, the combination of major and additional weights could lead to a systematic inconsistency. In most German states, the expenditure needs per inhabitant, when calculated solely on the basis of the major weight system, increase very rapidly with the size of municipalities (in terms of population). If the additional weights are also included, the expenditure curve generally moves upward but its progression becomes evidently slow (Leibfritz and Thanner 1986; Parsche and Steinherr 1995; Steinherr and Parsche 1998).

Furthermore, the present German system does not adequately consider the so-called 'spatial policy oriented' socio-economic determinants and the current large-city specific challenges which also have significant effects on expenditure needs and the behaviour of individual cities and municipalities. For instance, the list of such determinants includes:

– differences in age and employment structure of inhabitants (for example, the share of children and aged persons, of persons unemployed and social welfare recipients),

– varied financial and other types of problems caused by the increasing

determined by the share of inhabitants who are living in the surrounding municipalities but benefiting from the public services or utilities offered by the centres.

[5] The fiscal capacity of a municipality is shown in terms of the so-called tax capacity measurement index number ('*Steuerkraftmeßzahl*') which consists of individual tax capacity index numbers of local taxes like real property tax (*Grundsteuer*) A and B as well as trade tax (Gewerbesteuer), in addition to that of the municipality's share of current income tax revenue. The individual tax capacity index numbers of real property tax and trade tax are calculated, when the basic municipal-specific tax amount ('*Steuergrundzahl*') of each category of such local taxes is multiplied by the state-wide raising weights (in %), which are legally determined in the context of German intergovernmental resource allocation system ('*Nivellierungshebesatze*'). On the other hand, the basic tax amount of real property tax A is, for example, estimated annually so that this (estimated) basic sum can be fully in accordance with the current fiscal year's actual sum of tax revenue (for real property tax A) divided by the corresponding (legally fixed) raising weights (Steinherr and Parsche 1998).

number of commuters and by the provision of central-place-specific utilities in the city, including cultural activities,

– differences in local economic structure (for example, the share of the service sector and of industries in trouble),

– differences in environmental damage and preservation, and,

– different rehabilitation requirements of old buildings in the municipality or in the city centre (Budde and Junkernheinrich 1986; Leibfritz and Thanner 1986; Kuhn 1995).

CURRENT ECONOMIC CHALLENGES FOR LARGE GERMAN CITIES AND SOCIO-ECONOMIC DETERMINANTS FOR THEIR EXPENDITURE NEEDS

Large German cities are presently facing a variety of challenges. The economic integration in Europe, the rapidly increasing reorganisation and internationalisation of markets, etc. have made different national urban hierarchies merge into one common European urban system with a resultant intensification in rivalry for status between the European 'world' cities (Rozenblat and Pumain 1993).[6] As a consequence, the future European urban system appears likely to be increasingly shaped by the dynamics of economic competition, especially for the investment of multinational enterprises, the location of public institutions and hallmark events such as major sporting occasions, cultural festivals and trade fairs that have significant multiplier effects (Lever 1993). Apart from high-technology manufacturing and information processing which, according to a widespread opinion (in theory and practice), guarantee continued growth, large German cities (and their counterparts in Europe) are nowadays specialising in functions that particularly serve European economic integration. In the future, their success will depend heavily on capturing international management and other key functions including banking and insurance centres, high-speed transport connection, etc. (ERECO 1994; Bosman and De Smidt 1993).

The competitiveness of a large city is generally shaped by the consequence of the complex inter-play between economic, political, social and cultural factors (Nam, Russ and Reuter 1990). In addition, a firm's location is determined by the relative advantages and disadvantages perceived to exist in alternative locations:

[6] A world city can be defined as "a large urban agglomeration specialising in international control capabilities, manifested spatially by three interrelated components: a management and financial centre of global reach, a very high concentration of advanced producer services and an extremely rich physical and social infrastructure. Almost all large metropolitan areas do contain some elements of these... characteristics, but only those which have all of these components developed to the highest level can be defined as world cities" (Shachar 1994, p. 389).

For example, in the case of the selecting of a large city as a location in Germany firms will accept additional costs – usually reflected in higher rents, labour costs and taxes – to take advantage of the special agglomeration benefits that this particular location affords (see below). In order to enhance competitiveness and location quality, large cities in west Germany currently apply a number of development strategies in common, which include:

– improvement of business-related infrastructure, the housing situation and inner-city traffic systems, as well as active engagement in transport policy co-ordination between the city and neighbouring municipalities,

– ensurement of continuous economic development while maintaining diversified economic and business structure, and enhancement of reputation as, for example, centres of high-tech, R&D, culture, international trade-fairs and congresses, or other modern services,

– efficient spatial policy in providing new sites for industrial firms and office blocks (also in the context of close co-operation between the city and its surroundings) and development of attractive inner-city commercial areas,

– improvement of the availability and quality of general and professional education facilities to ensure favourable development of the urban labour market,

– maintenance and improvement of quality of life and cultural and leisure infrastructure,

– active city marketing to attract tourists and investment of multinational firms and international institutions from abroad, and

– active participation in EU integration, especially in the context of various city network programmes in economic, political and cultural fields.

Traditionally, cities have been the motor force of regional economic development and centres of key (regional) political and business functions. For example, in the process of transformation from a planned to a market economy in the new German states, direct investment from the west and economic modernisation has heavily concentrated on Berlin, Dresden and Leipzig, and these cities have played a significant role as growth poles for the (economic) development and recovery of the entire east German territory. According to the popular export-base theory, large cities export a wide variety of goods and services to the surrounding region, the nation and the world, and the export of services in various forms appears to be growing. Furthermore, it is generally argued that most innovations are made in cities – the incubator hypothesis (Glaeser et al. 1992). In line with the standard spatial economic theory growth is triggered by the agglomeration economies which include localisation economies resulting from the expansion of a particular industry and/or service in a certain place, and urbanisation economies arising from the greater array of business, social, educational, cultural and other public services and opportunities available in larger central places (Glaeser et al. 1992; Mills and McDonald 1992; Moulaert and Djellal 1995; Nam 1999). In other words, agglomeration can stimulate efficient production and generate productivity growth leading to higher per-capita income.

On the other hand, there are popular criticisms of demographic and economic growth in large cities which can also be well applied in Germany:

– Since large agglomeration areas provide an unparalleled business environment to economic sectors, rural regions are at loss to compete, with the result that the regional disparity in a country increases. Lagging regions have suffered from a net reduction in population size, decreasing corporate and income tax revenue, fewer job opportunities, etc.

– If population and economic activities exceed a certain level in the limited area of a large city, residents can no longer enjoy its prosperity but will suffer from poor living conditions which could be well characterised by restrictions on the utility of urban facilities and services like commuter trains and roads, waste disposal, the prevention of air pollution and the supply of affordable housing, etc. In other words, there is a point at which the (demographic and economic) growth of a city reaches its limit.

In the recent structural changes, a large number of jobs in the industrial sector have disappeared from large German cities. In particular, the cities whose economy is traditionally based on older heavy industrial and port sectors have generally fared worse. Furthermore, the transformation cities like Berlin, Dresden and Leipzig and cities in the Ruhr area such as Cologne, Düsseldorf, Essen, Duisburg and Dortmund are still carrying out the modernisation of industrial structure and are in the active de-industrialisation process from a manufacturing to a service-based economy. As a consequence of these urban economic changes, it is unskilled blue-collar workers that have suffered the most. However, the recent expansion of services in the city regions may be unable to absorb them on a full-scale, partly due to different requirements of skills and qualifications for new jobs. In addition, the anticipated increase in migration flow from rural areas and also from abroad into German city areas could make this type of structural unemployment more serious and persistent. In this context, the aspect of (re-)-training has generally appeared as an important issue in (urban) labour market policy, especially in order to reduce the long-term unemployment rate.

In Germany the economic relationship between cities and their surrounding regions seems to have been complementary more often than competitive (Nam 1997). The economic development has been shared between the city and its surroundings or the economic problem of the city at its centre has restricted the development of the entire agglomeration area to a large extent. This relationship is particularly well demonstrated in the experiences of the labour, as well as the housing and land markets. For example, Munich has played a motor role in the remarkable economic development of the entire city region, providing job opportunities for commuters from the surroundings (and also migrants from other German regions to this city). On the other hand, land prices in the surroundings have been pushed up rapidly by the economic attraction of proximity to the city of Munich.

In recent years, however, the relative employment growth was distinctly

higher outside the large agglomerated areas in the western part of Germany. The winners of this so-called 'deglomeration' process were the peripheral regions, while surrounding municipalities of large cities have also achieved a remarkable growth rate in general. Moreover, the movement of firms to the suburbs or the establishment of branches or back offices have gradually reduced the demand for the city's labour. The city cores have not only had the lowest (relative) growth in the service sector, partly due to the existing high level of activity of this sector, but have also suffered most seriously from the restructuring process in the manufacturing sector (Bade and Kunzmann 1991; Bade 1994). In other words (apart from the explicit movement of firms and plants) these types of spatial decentralisation have to a large extent taken place, because of the more serious decline, death and lack of firms' growth in cities and the birth and rapid growth of firms elsewhere (Crampton and Evans 1992). With regard to the major reason for firms' movement to the suburbs around large German cities, firms generally complain that they can no longer find sites to a new location or additional space for further expansion in the city. Some firms have already left their traditional sites for the surroundings, mainly because of high rental costs and/or restricted possibilities of expansion (IHK Munchen und Oberbayern 1992).

Conceding that in recent years surrounding municipalities of large German cities have generally achieved faster employment and GVA (gross value added) growth in services (Koll et al. 1997), many urban and regional economists argue, that the concentration of (more productive) high-order services (including real estate, legal, financial and business services) has taken place in large cities at the same time (see, for example, the cases of New York, London and Frankfurt). Numerous empirical studies carried out in the US which examined the appropriateness of the traditional central place theory with particular regard to development of business and professional services in urban areas, find a significant correlation between the size of city and the development of these types of services. For instance (the centres of), large cities house a number of specialised services including investment banking, commodities trade and information-based services, and the incidence of these modern services decreases with city size (Hall 1993; Preston and McLafferty 1993; Esparza and Kremenec 1994). The concentration of high-order services has become more intensified in recent years in the large world cities, since the less-specialised, low-value services have been continually replaced in the city centre in a more dynamic way, compared to the case in peripheral regions (Brake 1990).

In the process of sub-urbanisation there has also been the selective migration from central city to suburbs of higher-income households who could afford long-distance commuting and expensive housing in the city fringes. One additional reason for this type of movement could be "the lack of equivalence between... taxation and the perceived consumption benefits of central city services... the provision of education, of public health facilities and cultural and many other amenities requires high tax payments, in particular by the non-poor...

People move to nearby towns... [and]... can still use the central city's facilities, although they are not taxed for them [additionally]..." (Pommerehne and Krebs 1991, p. 783).

In west Germany, however, not all the higher-income households have opted for a home in a suburban setting. They are also found in pockets of attractive, expensive housing in the cities themselves. The successful enhancement of inner-city attractiveness has also hindered the process of sub-urbanisation in many large German cities: Such activities have brought back to the city centres the high-income 'yuppie' households who prefer the excitement of living close to the central business district. "But it has been shown... that many high-income households compete with [rather] low-income ones for the same housing... "(van Weesep and van Kempen 1992, p. 989). Triggered by the serious housing shortage, this competition has led to an increase in housing prices and rents in large German cities, something which has, in turn, reduced disposable income of households. In this context, a term 'new poverty class' has recently emerged in Germany indicating the economic problems of low- and middle-class households living in large cities. Furthermore, there are also examples of suburban governments which try to preserve the existing low densities by zoning, to attract higher-income groups and exclude lower-income groups.[7] For example, on several occasions, suburban governments have prevented the construction of low-income (publicly-supported) housing for inner-city poor (van der Veer 1994).

ALTERNATIVE METHODS OF CALCULATING LOCAL EXPENDITURE NEEDS DEMONSTATED IN INTERNATIONAL REFERENCES

As a possible alternative to the existing German model, Hanusch and Kuhn (1985) developed a concept for expressing local expenditure needs in terms of the 'social need index'. Unfortunately, this approach has not yet become popular in Germany. In order to calculate such an index, one first identifies, for relevant public activities in individual municipalities, corresponding statistical data that are available and expresses these as indicators per inhabitant (Table 1).

These statistical values are then compared among the investigated municipalities and standardised on a common scale between 1 (the lowest value) and 100 (the highest). In order to reflect the different preferences among local activities, a weighting system can additionally be introduced: It is simply assumed by Hanusch and Kuhn (1985) that the distribution of total amount of expenditure among local activities is an indicator of differences in municipal preferences. The

[7] The municipal zoning regulations can generally be differentiated into two types. "The traditional justification for zoning is that it minimises the impact of externalities by separating incompatible land uses. Fiscal zoning refers to attempts to use zoning to attract property owners with high tax-to-service ratios. Fiscal zoning may thus act to exclude certain classes of potential residents such as the poor..." (Pogodzinski and Sass 1991, p. 597).

Table 1. Local activities and statistical indicators considered by Hanusch and Kuhn in the calculation of the social need index in Bavaria.

Municipal activities	Statistical indicators (examples)
Public security (fire protection)	Number of fires
Education	Number of pupils for primary and secondary school, number of students for technical college and other types of professional schools, etc.
Science, research and culture	Size of exhibition areas in museums as well as number of visitors to museum and theatre, etc.
Social welfare	Number of social welfare recipients, the number of older persons (over 65), number of inhabitants aged under 5 for kindergarten etc.
Health, sport and leisure	Number of in-patients in hospital, size of sport and leisure areas etc.
Construction, housing and transport	Size of living space, length of streets, etc.

Source: Hanusch and Kuhn, 1985, *Messung des kommunalen Finanzausgleichs – Ein alternativer Ansatz für die Schlüsselzuweisungen*, [in:] Akademie für Raumforschung und Landesplanung (ed.), *Räumliche Aspekte des kommunalen Finanzausgleichs*, Hannover.

aggregated need-index, the sum of standardised (and also weighted) indicators of all the local activities considered, determines the rankings of municipalities, according to which the total size of expenditure needs of a municipality should differ correspondingly from the others.

For 23 large municipalities (central places and cities) selected in Bavaria, Hanusch and Kuhn (1985) developed, on the basis of the calculated index, a ranking of local expenditure needs for 1982, by which the large municipalities are in general badly scored. In particular, the two largest cities in Bavaria, i.e. Munich and Nuremberg, achieved the lowest values.[8] This fact suggests that in contrast with the case of the present German system shown above the local expenditure needs per inhabitant could decrease with the size of the municipality.

The resource transfer from central to local government takes place in Norway in the form of general and (certain public service-oriented) special grants. In general, the former are given to municipalities on the basis of their economic strength, measured in terms of income per-capita in previous years, while the latter are determined by the actual expenditures incurred by a municipality, which were necessary for providing local services in the fields of education, social welfare, churches and culture, transport infrastructure/water supply/waste disposal, parks and sports as well as healthcare.[9] For instance, local authorities receive from the central government a certain percentage share of their expendi-

[8] The highest index was given to the municipalities such as Schweinfurt (61), Aschaffenburg (59), Ansbach (58), while Munich was scored at 24 and Nuremberg at 29.

[9] These are six major areas of local government activity in Norway.

tures for every hospital-bed in use (50% in 1980) or teacher-hour in schools (Hansen and Gerhardsen 1981).

Since such special grants are made available on the basis of the actual expenditure level of each municipality, the total amount of grants depends to a certain extent upon the municipality's own priorities in satisfying its social and infrastructure needs as well as on the efficiency of local administration in providing the services. Some economists critically argue that under this system grants are solely distributed according to what local authorities actually do, and not according to what they might or should do, with the money in order to satisfy existing needs (Zimmermann 1981). Nevertheless, this Norwegian special grant system emphasises the importance of local political priorities, local service standards and the efficiency of local administration when determining the local expenditure needs.

In the 1980s, in order to promote the efficiency of such a block grant system and achieve greater equality in more relevant – and not only in per-capita – terms in the local service provision across the municipalities, the Norwegian central government gave serious consideration to possibilities for adopting a regression analysis to measure local expenditure needs (Hansen and Gerhardsen 1981). In the context of this calculating method, the total expenditure per inhabitant is estimated on the assumption that each local authority has the same set of effective choices with respect to the standard of public services. Using a number of selected 'expenditure' and 'need' indicators,[10] which seem to have the most significant effects on expenditure level and behaviour of municipalities, Hansen and Gerhardsen (1981) carried out a multiple regression analysis to identify a set of more objective criteria for the distribution of grants to municipalities.[11] The major concern of this investigation was to examine how individual need indicators interact with the size of municipality in their effects on expenditure levels.

The regression analysis carried out by Hansen and Gerhardsen postulates that the chosen need indicators generally have stronger impacts on the level of expenditure in larger municipalities (with a population of more than 15,000 inhabi-

[10] The following ten expenditure variables (measured in per-capita terms) are used in the investigation: total education expenditure, expenditure on elementary education, health care expenditure, total social welfare expenditure, services for children and youth, services for the elderly, expenditure on churches and cultural activities, expenditure on transport infrastructure/water supply/waste disposal, maintenance expenditure on parks and sport facilities, and expenditure on public administration. On the other hands, the thirteen need variables are: number of inhabitants, density, percentage share of children of school age, of children and of youth aged under 15 years, and of elderly people aged over 65 years, percentage share of old housing units, percentage share of workforce employed in industry, number of commuters as percentage share of total number of inhabitants, single parents as well as one-person families as a percentage of all families, number of employed woman as a percentage of all women, number of private cars per family, and length of municipal roads as a percentage share of all roads within the municipality.

[11] This analysis is based on 1974 data on expenditure and need indicators which encompass altogether 421 out of 445 Norwegian municipalities.

tants) than in smaller ones. This finding in turn suggests that it would be appropriate to have separate regression models for measuring expenditure needs corresponding to the size of a municipality, rather than a common one. Furthermore, in most activities (especially, education, healthcare, culture as well as transport infrastructure, water supply and waste disposal) the calculated coefficients, which show the relationship between per-capita expenditure and population size, are negative for smaller municipalities but positive for larger ones, indicating the possibility that different types of scale effects could operate in the two groups: For smaller municipalities per-capita expenditure decrease as the population size increases, while increases in population immediately imply higher per-capita expenditure in larger ones. Subsequently, the relationship between population size and per-capita expenditure can be illustrated in those activities as U-shaped (Nam 1993). In the rest of the cases (especially social welfare as well as parks and sports), per-capita expenditure increases continuously with population size.

According to the so-called normative approach, the level of future expenditure needs of a municipality or a region (required, for example, to provide different types of local infrastructure) is induced from the existing bottleneck of their endowment in this area, but only to the extent that the availability and quality of such infrastructure reaches the standard fixed nationally (Leibfritz and Thanner 1986). This type of measuring method has been applied in Japan to identify the general volume of financial support which should be made available by the central government to improve the general condition of local infrastructure.[12]

In the Japanese model, the financial needs of a municipality are calculated in a separate way, according to individual categories of local government activities (and also expenditures) which include, for example, fire brigades, construction (street, bridge, harbour, city planning, parks etc.), education (primary, middle and high schools), social welfare and health care, regional economic (agricultural and industrial) promotion and public administration. Several indicators are applied in Japan as the calculation basis for the different types of local expenditure needs. Examples are:

– Number of inhabitants covered by fire brigades, parks, social welfare and health care,
– Surface area and length of streets,
– Number of pupils in education, and
– Number of employees in agricultural and industrial sectors (Hirota 1985).

For the individual categories of local activities, a total amount of standard-expenditure is estimated by the central government on the basis of given demo-

[12] Politically Japan consists of a number of prefectures and municipalities (within a prefecture) with their own representative bodies. Both of them also enjoy substantial financial autonomy. The construction of local infrastructure (including streets) and provision of education are generally carried out in Japan by these two types of under-level governments in co-operation. On the other hand, responsibilities including for the police are among the typical activities of prefectural governments.

Table 2. Selected indicators of the 'fictitious' standard region and municipality in Japan.

	General indicators
Standard prefecture	
Number of inhabitants	1,700,000
Total area (km^2)	6,500
Number of cities located	10
Number of inhabitants in cities	900,000
Number of municipalities (except cities)	75
Number of inhabitants in municipalities (except cities)	800,000
Standard municipality (including cities)	
Number of inhabitants	100,000
Total area (km^2)	160
Number of households	29,000

Source: Hirota, 1985, *Die originäre Ermittlung des Finanzbedarfs am Beispiel des japanischen Finanzausgleichs*, [in:] Akademie für Raumforschung und Landesplanung (ed.), *Räumliche Aspekte des kommunalen Finanzausgleichs*, Hannover.

graphic and spatial conditions of the chosen 'fictitious' standard region or municipality (Table 2).[13] From this total sum, the amount of unit expenditure for each calculation basis (for example, costs per inhabitant or pupil, costs per km of street etc.) is derived and applied nationwide in practice when calculating the real expenditure needs of a municipality. There are two types of basic unit costs which are taken into account in Japan: maintenance costs for the existing infrastructure, and new investment costs. The Japanese model also contains a number of additional 'correction factors' (several weighting measures including location/size/infrastructure factor, population density factor, etc.), which have been introduced to discriminate between the measurement of local expenditure needs, the demographic concentration, climatic differences and the geographical and other region-specific disadvantages of municipalities in a more systematic way (Hirota 1985; Parsche and Steinherr 1995).

CONCLUSIONS

In practice, the expenditure needs of a city and a municipality are calculated by adopting different types of models. For example, the weighting system has been applied in Germany, the implementation of regression analysis has been considered in Norway and a normative approach has been implemented in Japan. However, these varied approaches unfortunately provide incompatible and disputable explanations regarding the relationship between per-capita local expen-

[13] In a number of activities, the derived fictitious expenditure needs of a municipality are also partly covered by the local income (charge, usage fees, etc.) attained from the provision of infrastructure.

diture needs and the size of a municipality (measured in terms of population). The German method is traditionally developed on the basis of Popitz rule which simply assumes the positive relationship between these variables. This rule emphasises the fact that a central place, like a large city, usually provides a wide range of public services (e.g. infrastructure and social welfare activities) also for the surroundings. By contrast, the social need index suggested by Hanusch and Kuhn (1985) as an alternative to the present one, apparently shows the possibilities for the negative correlation between the two. In Norway, the relation appears to be U-shaped, as observed in the regression analysis which treats expenditure structure and the behaviour of small and large municipalities separately. In the normative approach applied in Japan, the amount of expenditure per inhabitant is generally fixed for individual activities of local governments, regardless of the size of the municipality. These non-uniform outcomes immediately signal an urgent need for more systematic empirical research in the field not only of measuring local expenditure needs but also of intergovernmental resource allocation in Germany.

Furthermore, it would appear to be urgently necessary – particularly from the large-city point of view – to consider more adequately some significant socio-economic factors and current challenges in the German way of measuring local expenditure needs, which include, for example:

– the ongoing fierce competition among large European cities for capturing high-tech manufacturing and modern international service functions which guarantee sustainable growth,

– the provision of new commercial sites and city-specific soft and hard infrastructure better adapted to newly-emerging industries and services,

– the diminishing economic importance of large central cities as the growth motor caused by the recent sub-urbanisation of firms and rich households,

– the large-city specific economic restructuring and de-industrialisation and related urban labour market problems,

– the increasing social, environmental and housing problems caused by the growing demographic concentration, traffic congestion as well as the number of low-income households and social-welfare recipients (e.g. elderly people and unemployed persons) in the city, etc.

REFERENCES

Akademie für Raumforschung und Landesplanung (ed.), 1985, *Räumliche Aspekte des kommunalen Finanzausgleichs*, Hannover.

Albers W. et al., 1980, *Handwörterbuch der Wirtschaftswissenschaft (HdWW)*, Vol. 8, Stuttgart et al.

Bade F.-J., 1994, *Regionale Beschäftigungsprognose 2000*, University of Dortmund, Dortmund.

Bade F.-J., Kunzmann K. R., 1991, *Deindustrialization and Regional Development in the*

Federal Republic of Germany, [in:] Rodwin L., Sazanami H. (eds.), *Industrial Change and Regional Economic Transformation. The Experience of Western Europe*, London.

Bosman J., De Smidt M., 1993, *The Geographical Formation of International Management Centres in Europe*, Urban Studies, 30, 6, 967–980.

Brake K., 1990, *Der Einfluß der Tertiärisierung auf die Städte und das Städtesystem*, [in:] Kemming H. et al., *Tertiärisierung und Stadtstruktur. Zur Notwendigkeit der Neuorientierung städtischen Handelns*, ILS Schriften, 44, Dortmund.

Budde R., Junkernheinrich M., 1986, *Kommunale Inzidenz eines Arbeitslosenansatzes im nordrhein-westfälischen Finanzausgleich*, Ruhr-Forschungsinstitut für Innovations- und Strukturpolitik, Bochum.

Cheshire P. C., 1992, *European Integration and Regional Response*, CeSAER Working Paper, 1, Faculty of Urban & Regional Studies, University of Reading, Reading.

Crampton G., Evans A., 1992, *The Economy of an Agglomeration: The Case of London*, Urban Studies, 29, 2, 259–272.

ERECO, 1994, *European Regional Prospects and Forecasts to the Year 1998 for European Cities and Regions*, Cambridge.

Esparza A., Kremenec A. J., 1993, *Producer Services Trade in City Systems: Evidence from Chicago*, Urban Studies, 31, 1, 29–46.

European Institute of Urban Affairs (Liverpool John Moores University), 1992, *Urbanisation and the Functions of Cities in the European Community*, Regional Development Studies of the European Commission, 4, Brussels.

Glaeser E. L., Kallal H. D., Scheinkman A., Schleifer A., 1992, *Growth in Cities*, Journal of Political Economy, 100, 6, 1226–1252.

Hall P., 1993, *Forces Shaping Urban Europe*, Urban Studies, 30, 6, 883–898.

Hansen T., Gerhardsen R., 1981, *Expenditures, Social Needs and Size of Municipality*, [in:] OECD, *Measuring Local Expenditure Needs: The Copenhagen Workshop*, Urban Management Studies, 4, Paris.

Hansmeyer K. H., Kops M., 1985, *Die Popitz'sche Hauptansatzstaffel*, [in:]: Akademie für Raumforschung und Landesplanung (ed.), *Räumliche Aspekte des kommunalen Finanzausgleichs*, Hannover.

Hanusch H., Kuhn M., 1985, *Messung des kommunalen Finanzbedarfs – Ein alternativer Ansatz fur die Schlüsselzuweisungen*, [in:] Akademie für Raumforschung und Landesplanung (ed.), *Räumliche Aspekte des kommunalen Finanzausgleichs*, Hannover.

Hirota S., 1985, *Die originäre Ermittlung des Finanzbedarfs am Beispiel des japanischen Finanzausgleichs*, [in:] Akademie für Raumforschung und Landesplanung (ed.), *Räumliche Aspekte des kommunalen Finanzausgleichs*, Hannover.

Industrie- und Handelskammer (IHK) für München und Oberbayern, 1992, *Betriebsverlagerungen in das Münchner Umland. Ergebnisse einer IHK-Umfrage*, Munich.

Koll R., Nam Ch. W., Kiemer K., 1997, *Dienstleistungsstandort Bayern*, ifo Studien zur Regional- und Stadtökonomie, 10, Munich.

Kuhn Th., 1995, *Theorie des kommunalen Finanzausgleichs*, Heidelberg.

Leibfritz W., Thanner B., 1986, *Berücksichtigung der zentralen Orte im kommunalen Finanzausgleich Bayern*, Ifo Studien zur Finanzpolitik, 38, Munich.

Lever W. F., 1993, *Competition within the European Urban System*, Urban Studies, 30, 6, 935–948.

Mills E. S., McDonald J. F., 1992, *Sources of Metropolitan Growth*, Center for Urban Policy Research, Rutgers University, New Jersey.

Mottershead P., 1994, *Economic Problem of Central City Surrounds: Evidence from Three European Regions*, Regional Studies, 28, 3, 326–330.

Moulaert F., Djellal F., 1995, *Information Technology Consultancy Firms: Economies of Agglomeration from a Wide-area Perspective*, Urban Studies, 32, 1, 105–122.

Nam Ch. W., 1993, *Can the True Expenditure Needs of a Local Government Be Measured?*, Ifo Diskussionsbeiträge, 11, Munich.

Nam Ch. W., 1997, *Major Economic Challenges for Large German Cities in an Enlarged Europe*, Zeitschrift für Wirtschaftswissenschaften, 16, 323–346.

Nam Ch. W., 1999, *The Significance of Regional Technology Policy for Local Innovation Networks in Germany – Case Studies from Bavaria and North Rhine-Westphalia*, Geographia Polonica, 72, 1, 63–88.

Nam Ch. W., Nerb G., Russ H., 1990, *An Empirical Assessment of Factors Shaping Regional Competitiveness in Problem Regions*, Document Series of the European Commission, Brussels.

Nam Ch. W., Russ H., Reuter J., 1990, *The Competitiveness of EC Regions in the Nineties*, [in:] Kremmel P. (ed.), *The Single European Market in the Nineties*, Ifo Studien zur europäischen Wirtschaft, 4, Munich.

OECD, 1981, *Measuring Local Government Expenditure Needs: The Copenhagen Workshop*, Urban Management Studies, Paris.

Parsche R., Steinherr M., 1995, *Der kommunale Finanzausgleich des Landes Nordrhein-Westfalen*, Ifo Studien zur Finanzpolitik, 59, Munich.

Pogodzinski J. M., Sass T. R., 1991, *Measuring the Effects of Municipal Zoning Regulations: A Survey*, Urban Studies, 28, 4, 579–621.

Pommerehne W. W., Krebs S., 1991, *Fiscal Interaction of Central and Suburbs: The Case of Zurich*, Urban Studies, 28, 5, 783–801.

Preston V., McLafferty S., 1993, *Income Disparities and Employment and Occupational Changes in New York*, Regional Studies, 27, 3, 223–235.

Rozenblat C., Pumain D., 1993, *The Location of Multinational Firms in the European Urban System*, Urban Studies, 30, 10, 1691–1709.

Shachar A., 1994, *Randstad Holland: A 'World City'?*, Urban Studies, 31, 3, 381–400.

Steinherr M., Parsche R., 1998, *Der Ausgleich zwischen Finanzbedarf und Finanzkraft im kommunalen Finanzausgleich des Landes Rheiland-Pfalz*, Ifo Studien zur Finanzpolitik, 66, Munich.

Van der Veer J., 1994, *Metropolitan Government and City-urban Cleavage: Differences between Old and Young Metropolitan Areas*, Urban Studies, 31, 7, 1057–1079.

Van Weesep J., Van Kempen R., 1992, *Economic Change, Income Differentiation and Housing: Urban Response in the Netherlands*, Urban Studies, 29, 6, 979–990.

Zimmermann H., 1981, *Local Expenditure Needs under Alternative Policy Objectives*, [in:] OECD, *Measuring Local Government Expenditure Needs: The Copenhagen Workshop*, Urban Management Studies, 4, Paris.

GEOGRAPHIA POLONICA, 73, 1, Spring 2000
PL ISSN 0016-7282

CHALLENGES OF REGIONALISM: DEVELOPMENT AND SPATIAL STRUCTURE OF THE HUNGARIAN BANKING SYSTEM

ZOLTÁN GÁL
Centre for Regional Studies, Hungarian Academy of Sciences, Pecs, 7601 POBOX 199, Hungary,
e-mail: galz@dti.rkk.hu

ABSTRACT: This paper is concerned with the spatial characteristics of the Hungarian banking system. Financial services became the key sector in the processes of economic transformation and one differentiated by uneven regional development. The spatial structure of the banking sector is characterised by a large-scale concentration in Budapest, but the foundation boom of branch offices is also typical in the regions, as the necessity of a presence on local markets, as well as competition for the retail market, stimulate banks to expand their branch networks. Commercial banks, which have their headquarters exclusively in Budapest, largely confine themselves to the collection of deposits in their national network, resulting in capital drainage and net capital loss from most of the regions. The presence of the centralised capital market and the lack of a decentralised regional financial system can restrain and slow down regional development in the long run.

KEY WORDS: Hungarian banking system, transitional financial markets, banking network, global-local dichotomy, uneven regional development.

INTRODUCTION

As a newly established sub-discipline of economic geography, financial geography deals with the flows and transformation of money, and the spatial, institutional and regulatory structure of financial capital (Leyshon 1995; Leyshon and Thrift 1997). The recent growth of interest in the geography of money has been stimulated by an explosive growth in information technology and financial services and also by the profound changes and upheavals (crises) that have remapped, and are continuing to transform, the financial landscape of the world (Leyshon et al. 1988). These changes are usually associated with a single word, globalisation. Globalisation – giving a definition by Ron Martin – refers to the increasing integration of financial markets, the hybridisation, convergence and stretching of economic relationships across space, regardless of national borders and institutions, and to the growth of "stateless monies" that move electronically around the globe at a very high speed, ignoring national borders and economic territories (Martin 1999). Financial globalisation is inherently constituted geographically, a product of the organisational, technological, regulatory and corpo-

rate strategies of individual firms, financial institutions and authorities in a specific location. Divergent forces of deconcentration/decentralisation and concentration/centralisation are consistent with financial globalisation, and are shaping the evolving geographies of national, regional and global finance. Different monetary spaces – national, global and local, regional – coexist, as it is recognised that globalisation of finance is a global-local process. There is an inherent tendency for financial services to concentrate in particular locations and these services and their availability differ from one locality to another. National, and even more so regional, economies remain important, as investment portfolios remain highly skewed by different spaces, and while there has been a move towards equalisation of interest rates, this has not been matched by an equalisation of the profit rate (Leyshon 1995). Banking has experienced structural changes to a greater extent as it has evolved its early functions of intermediation to its present day system of securitization, so the geographical structure of the banking system has undergone successive phases of development (at first local and regional banking, then national branch banking, followed by the concentration of financial activity in selected regional and national centres, and more recently the growth of global centres). Not every country has followed the same sequence of spatial forms as they are at the different stages of banking development (Dow 1999).

These changes have several effects on the emerging single European market, where finance with the European banking licence lies at the heart of the policy. The emergence of the European Monetary Union (EMU) encourages merger and acquisition activity across the EU in order to strengthen the position of financial institutions to hold their own under increased competition. While cross-border acquisition has been limited, the emergence of new large national and universal banks, as the amalgamation of several national or regional institutions, is bound to have important spatial consequences, as these are located in the existing financial centres. These banks will have even more power to dominate the European market (Leyshon and Thrift 1997). Changes which impose a universal monetary space for Europe remove a significant element of national and regional autonomy concerning monetary control over economic territory. The consequences of financial integration will effect regional and local banks as well as different national banking systems. Small and local banks might suffer a competitive disadvantage initially, while eventually a two-tier banking system would emerge with one tier consisting of international banks and the second tier consisting of local banks (where local banks include local, regional and national banks devoted to their domestic markets). According to the European Commission's view, bank customers would enjoy a wider range of financial services and improved access to credit at a lower cost as a result of the competition. The growing literature on regional finance suggests that credit allocation in a regional banking system and the different national banking systems is different in line with the stage of development, and frictions also exist across regions within national economies, resulting in differing availability of capital (Porteous 1996; Mazucca 1993; Dow

1999; Alessandrini and Zazzaro 1999; Warf 1999). Less-developed banking systems – including also regional banks – have a lesser capacity to promote their economic development and might experience certain disadvantages as a result of the financial integration in Europe. Despite the fact that regional and local banks can serve local economic interests better than financial-centre banks whose priorities relate more to the single European and global markets, less-advanced banking systems can be controlled more easily by the large universal banks of the financial core areas.

This latter argument refers very much to the acceding countries of Central and Eastern Europe, such as Poland, the Czech Republic, Slovenia, Slovakia and Hungary, which followed their reintegration into the world financial market in the early 1990s. They not only have to adopt new technologies and the financial behaviour they accommodate, but also have to cope with a legacy of bad debts and a lack of experience in credit risk assessment. Central-Eastern European banking systems are accelerating through some features of the stages of development as a result of competition with more advanced systems and the state encouragement of banking development. As European Union membership approaches in Central Europe's more advanced economies, Western European banks are aggressively moving to expand into what will soon be a home market for them. The result is increasing pressure on margins, as more banks compete for relatively little business. This results in a reversal of the process of concentration in the EU, i.e. in a growing number of institutions. Making matters worse for the locals, the foreign banks often boast deeper pockets, greater expertise and a more solid reputation (Anderson and Kegels 1998). All these challenges which are to be faced are common in these countries, but what could be varied from country to country is the spatial and institutional structure of the national banking systems.

In this article I would like to give an overview of the institutional and spatial structure and development tendencies in Hungarian banking, which occurred during the first decade of two-tier banking. The allocation of branches of the banking institutions plays an important role, not only in the development of the national economy as a whole, but also in the development of the local and regional economy, in innovation and, last but not least, in the organisation of production and service sectors. The elements of the banking network connect to regional development as the supply section of business services, so the density and number of banks and their branches are important indicators of regional economic development. Unlike its Polish counterpart, the Hungarian banking system is based almost exclusively on the national branch-banking system, which means a lack of regionally-based independent universal banks. This very-centralised structure of the national banking system may combine with the growth of competition following further liberalisation and lowering of barriers to the national market (e.g. cross-border banking and branching which primarily targets the corporate, but not the retail, sector) to result in a decline in credit availability for SMEs situated in the Hungarian regions remote from the capital city and from European financial centres (Várhegyi and Gáspár 1997). In this respect, there

may well be a case for creating regional banks in the forming regions of Hungary in order to protect the local economy and facilitate the channeling and mediation of EU regional development funds in the future.

This analysis rests on data – usually published with a certain delay – in the Annual Report of the Bank and Capital Market Advisory Board, the Regional Statistical Yearbook and the Hungarian Almanac of Financial and Capital Markets. Most of these sources resemble publications on economics in still neglecting to consider geographical issues of the financial sector, thereby making research efforts more difficult. From the point of view of the geographer wishing to survey the rapidly-changing banking network in the expansion phase, it is made rather difficult to demonstrate accurate by the spatial structure of banking services.

THE FIRST DECADE OF DEVELOPMENT IN THE HUNGARIAN BANKING SYSTEM

The first important step forward in the modernisation of the Hungarian financial sector was the creation in 1987 of a two-tier banking system better adapted to a market environment. Following this Act, the National Bank of Hungary came to perform primarily central-bank functions, while institutionally-separated commercial banks were set up in January 1987. Commercial banks which originally had corporate clientele were admitted on to the retail market, while financial institutions were given commercial banking licenses. In denial of Hungarian traditions, a specialised rather than universal banking system was created in 1987, sorting different type of banks by functions (34 commercial banks, 8 specialised banks, mortgage banks and building societies and 236 co-operative savings banks; Tab. 1).

Since the reintroduction of two-tier banking (after 40 years of discontinuance), the banking system has opened up to the world as a competitive and rapidly-growing sector. The transformation into a market economy, radical dim-

Table 1. The structure of the Hungarian banking sector, 1997.

Banks	Number of banks			Number of branches		
	total	Budapest	provinces	total	Budapest	provinces
National Bank	1	1	–	19	1	18
Commercial bank	32	32	–	946	299	709
OTP (National Savings Bank)	1	1	–	420	96	324
Specialized bank	8	7	1	7	2	5
Investment bank	1	1	–	1	1	–
Mutual savings bank	246	1	245	1,740	31	1,710
Mortgage bank	1	1	–	1	1	–
Credit union	4	–	4	4	–	2
Total	292	43	250	2,722	265	2,446
Without savings b.	42	39	1	978	234	732

Source: Hungarian Almanac of Financial and Capital Markets, 1991, 1995–1996.

inution of the state's role in the business sector, privatisation, foreign capital inflow, more-intensive participation in the international division of labour and European integration have all provided new opportunities and challenges for banking. The new actors of a rapidly developing economy, the mushrooming of associations and corporations and the greater demand for corporate and retail markets will transform the present banking system. Money markets will undergo radical rearrangements in the future and the balance of power will change as a consequence of market competition. Banks are struggling for larger shares of the expanding market and for new customers which require both an expansion of the banking network and perpetual innovation in banking.

If we take ten years of development in the banking system into consideration, different periods can be indentified. The short period 1989–1992 was the peak time for the foundation of new banks. Competition was also increased by the entrance of new foreign-owned and joint-venture banks founding their own subsidiary banks in Budapest (Bácskai 1997).

After the period of rapid and extensive expansion, the banking system was characterised by the first bankruptcies and failures in the years 1992–1995. Over-geared expansion of balance-sheets and increasing risk-taking stood in contrast with the low level of financial standing and the huge sum of inherited debt that was accumulated in the central bank before 1987. This automatically led to the loss of market shares for the Hungarian-owned banks and to a strengthened position of foreign banks. The pecuniary difficulties of the mainly state-owned banks made the restructuring of the Hungarian banking sector, together with loan, bank and debtor consolidation inevitable. The main purpose of bank consolidation and privatisation was to reduce the percentage of state ownership in the banking sector to at least below 25%.

In the third period, commencing in 1995, a stabilised and more-competitive banking system emerged, characterised by successful privatisation of the banking system and resulting in a slower expansion in banking from 1996 onward. In this latter period of development, the branch network expansion was one of the major phenomena. This was due to business policies of banks shifting from the corporate to the retail market, intending to gain greater market shares through easier access to retail customers, and on the other hand strengthening competition which forced mainly foreign banks without branches to build networks in order to hold their ground. The growing retail market from the mid-1990s has urged banks to establish their extensive nation-wide networks of local branches. (This has occurred partly through the acquisition of offices of liquidated banks and partly through the opening of new branches). In parallel to the stabilisation processes, the growth of newly-established banks halted, while the founding of joint-ventures and subsidiaries of foreign banks was compensated by mergers and liquidations through a strengthening of concentration in banking. In the last few years, new types of banks have formed to serve the special interests of the money market (mortgage banks, building societies, land and mortgage banks), but the concentation in banking will continue (Tab. 2).

Table 2. Annual balance sheet footings and number of employees in Hungarian banking 1991–1996.

	Date of foundation	Owner-ship	Number of branches[c]	Balance sheet footings (million HUF)				Number of employees		
				1991	1993	1995	1996	1991	1993	1995
Big banks										
1. OTP.	1949	HT	423	669599	830958	1072777	1256085	12659	14367	–
2. Postabank	1988	HT	41	89863	168085	269632	378718	987	1472	2173
3. KH	1987	K	174	214234	231723	342294	360003	3891	4112	3837
4. MKB	1950	KT	20	238319	238293	330516	347901	1500	1512	1333
5. MHB[e]	1987	K	78	321824	357463	246091	250873	4180	4322	3095
6. Budapest Bank	1986	KT	74	146067	156605	212931	225380	2085	2952	2831
7. CIB	1979	KT	0	35000	76010	144113	175021	267	184	174
8. CIB Hungária	1988	KT	12	26341	45135	126424	134712	–	168	268
9. BA-Creditanstal	1990	K	12	28400	27202	77074	116119	111	193	343
10. Raiffeisen Unicbankt	1986	KT	9	28450	36083	77117	109630	88	152	309
Medium sized banks										
11. Mezőbank[d]	1986	K	81	29984	34544	80883	97152	401	657	720
12. Citibank	1985	K	0	35248	36474	59628	93196	97	150	216
13. MFB Rt.	1991	H	0	–	20468	57805	91921	–	117	184
14. IEB	1980	KV	14	35090	40540	62387	91117	202	323	417
15. ING Bank	1991	K	14	2564	20283	64558	84005	–	80	–
16. Commerzbank	1993	K	0	–	11027	57994	75270	–	40	74
17. ÁÉB	1922	K	8	19766	26148	37970	71657	270	294	279
18. Takarékbank	1989	H	22	36760	37323	40360	66789	281	509	613
19. BNP-Dresdner Bank	1990	K	0	9836	15014	35252	50639	–	90	115
20. Credit Lyonnais	1992	K	0	–	12577	28627	42160	–	–	–
21. HYPO-Bank	1993	K	5	–	4831	22349	40198	–	30	160
22. DAEWOO Bank	1989	K	0	11291	12854	25048	38774	44	54	76
23. ABN Amro	1993	K	0	–	5566	29826	38430	–	19	50
24. EKB	1991	K	4	6612	8700	23081	33631	34	46	110
25. WestLB Hungária	1985	KT	1	12890	9186	20372	31975	113	122	110
26. Pénzintézeti Kozpont	1983	H	3	–	15364	24001	29768	–	354	244
27. Polgári Bank	1993	H	15	–	2900	27069	29764	–	210	291
28. Konzumbank	1986	H	25	16016	11725	12894	27590	179	332	354

Small banks										
29. Realbank	1989	H	29	10200	13948	16622	20704	63	–	–
30. Volksbank	1992	K	5	–	1227	9356	15579	–	24	52
31. Merkantil Bank	1988	H	3	4508	5530	11772	13689	66	90	03
32. EXIMBANK	1994	H	0	–	1267	5196	12134	–	–	66
33. Corvinbank	1984	H	6	10449	18375	12558	11888	91	173	193
34. Porsche Bank	1994	K	0	–	500	7834	7072	–	–	52
35. Opel Bank							4630			
36. Hanwha Bank	1990	KT	1	4206	5613	4551	4371	65	74	59
37. Kvantum Bank	1991	H	0	1036	1050	3164	3581	16	39	68
38. Deutsche Bank							2994			
39. IC Bank	1993	K	0	–	979	1087	2330	–	–	22
40. Rákóczi Bank	1992	H	0	–	1056	1072	1689	–	18	22
41. Nomura	1991	K	0	1042	1037	780	774	8	11	10

K = foreign ownership, KT = majority of foreign shares, H = Hungarian ownership, HT = majority of Hungarian shares, KV = joint-venture with foreign ownership. [c]State of development at the end of 1996; [d]In 1995 together with balance sheet footings of Agrobank, purchased by Erste Bank; [e]Az ABN Amro Bank purchased 89% of MHB in 1996.

Source: Hungarian Almanac of Financial and Capital Markets, 1991, 1995–1996.

One of the most important alterations in the Hungarian banking system was that the role of foreign capital in ownership was determined. As a consequence of foreign capital inflow into Hungarian banking, the ownership structure was transformed entirely; in parallel with the process of significant decrease in state ownership (20% recently), shares of foreign capital in the banking system rose to 65%, leading to the gaining of a majority of market shares within a short time. This very high proportion of foreign capital is among the highest in the European context. In the UK 53% of ownership is foreign, but in Finland it is only 1% and in Germany 2%. According to the portion of invested capital, the main investors are still the leading German and Austrian entrepreneurs, followed by American, Dutch, French, Japanese and Korean investors. The activity of the Dutch banks is indicated by the fact that all the top-ranking Dutch banks (ABN Amro, ING, Rabobank) have opened subsidiaries in Hungary, though British banks are conspicuous by their absence (Várhegyi 1997).

To summarise the role of foreign capital in the Hungarian banking system it can be said that such a rapid process of privatisation of banking would have been impossible without foreign capital inflow. Foreign capital inflow into the banking system, together ownersip shares from privatisation, comprising a total of 220 billion HUF, was directly invested into banks based in Budapest but ran through the channels of a branch network. As early as in 1995, foreign banks, occupying one fourth of the total market, accounted for 70% of profit returns due to their high profitability, which was twice as great as in Hungarian-owned banks. Foreign capital investment has contributed significantly to the growth of the international competitiveness of Hungarian banking. (Per capita investment of foreign capital amounted to 1,950 USD in Hungary, compared to 1,040 USD in the Czech Republic and 832 USD in Poland up to 1999; Wachtel 1997).

As a consequence of successful privatisation, competitiveness of the banking system is strengthened: it has increased its balance sheet account and improved the quality of portfolios, though the proportion of the bad debts is still higher than the EU average, having been increased significantly by certain banks (Postabank, Realbank). By the end of the 1990s the Hungarian banking system was being affected by radical legislative, administrative and institutional changes, such as the privatisation of many banks, foreign capital inflow, the introduction of a more universal banking model, the liberalisation of the credit market and the consequent intensification of competition.

The Hungarian banking system is small as far as the size of banks and the ratio of balance sheet status to GDP are concerned (72% for Hungary and 110–240% for EU countries). The total assets in the Hungarian banking system are still only a fraction of those of a big European bank, and the largest Hungarian bank (OTP Bank) ranks only sixth among Central Europe's largest banks in terms of assets. As regards the number of employees in banking the number reached 45 000 in 1990, but then began to diminish and fell back to 38 260 employees by 1998 (Fig. 1). However, the percentage employment of banking

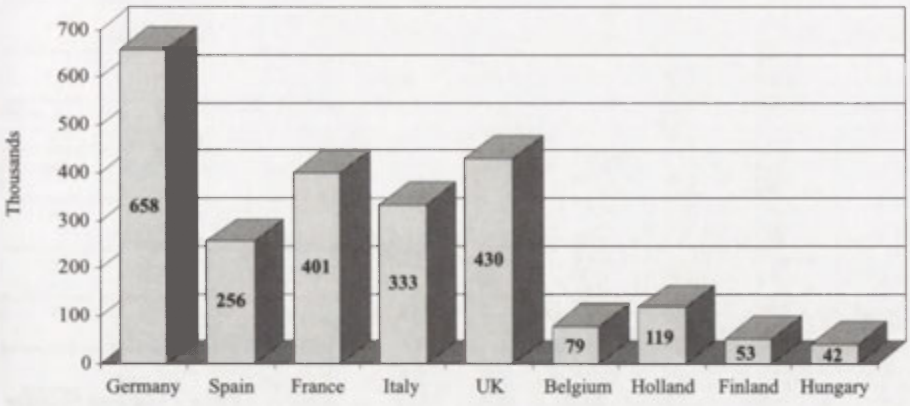


Figure 1. Employment in banking in some European countries (1994).

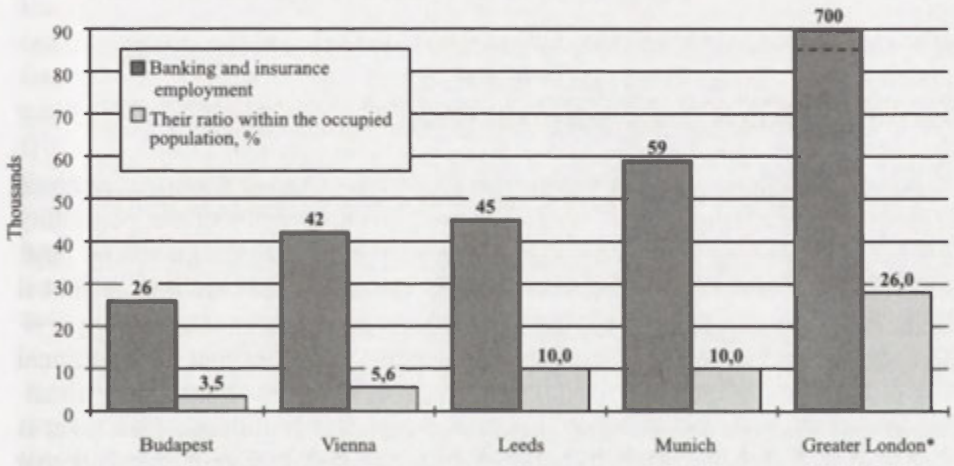


Figure 2. Employment in banking and insurance in some cities (1996).
* Financial sector employment.

among all employees rose from 1% in 1990 to 2.5% in 1997. (The rate in EU countries ranges from 2–4%). The banking and insurance sector of Budapest accounted for 47% of the total in 1996, or 3.65% of the total employment in the city. (In contrast, Vienna had 42,000 (5.6% of total employment), and Munich 59,000 (9.8%) in banking and insurance (Fig. 2).

The smaller size and extension of the banking network is highlighted by European comparisons. Smaller countries, such as Belgium and Holland have seven times as many branch offices, while the less-densely-populated Finland has twice as many. The number of branches in 1995 amounted to 1000, though the consequence of rapid expansion in the last few years raised this to 1400 (Gál 1998; Fig. 3).

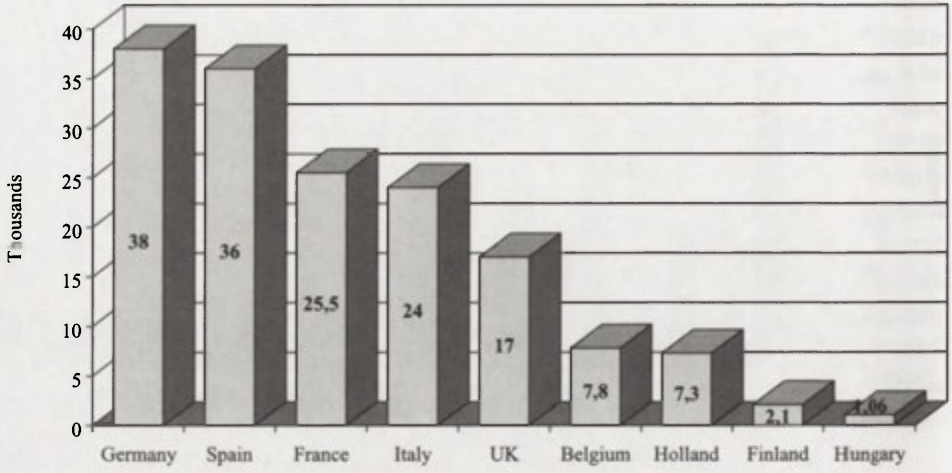


Figure 3. Number of branches of banking in some European countries (1995).

STRUCTURAL AND SPATIAL POLARITY OF THE BANKING SYSTEM IN THE 1990s

As regards the merger and acquisition activity in different European national banking systems, there is a recognition among member states of the possibility of increasing concentration in banking in Europe, leading to domination by large banks situated in a few financial centres of the single market. Recent global financial crises have highlighted the vulnerability and constrained state supervision of international financial markets, and one that questions that the traditional control functions of the nation-state. As a consequence of the diminishing financial role of the state the growing importance of *EMU* at supranational level is paralleled by strenuous effort to build up strong regional financial markets that will able to serve the interests of the regional economy better and represent a link between local economies and financial centres. The Hungarian banking system is characterised by the lack of strong local and regional banks – a fact that can arguably be explained in part by adjustment to more-concentrated international banking structures and in part by structural polarisation.

The spatial structure of the banking system is polarised compared to the network which existed at the turn of the century (when the number of independent banks scattered throughout the countryside were overshadowed within the banking network, and there were proportionally few branches in banking before World War I, with the result that only 5.7% of the network was concentrated in Budapest), the recent banking system is characterised by strong spatial concentration (Gál 1999a; Fig. 4).

The fact that all but one of the 41 banks are based and headquartered in Budapest results in a deformed structure in the banking system. Banking in

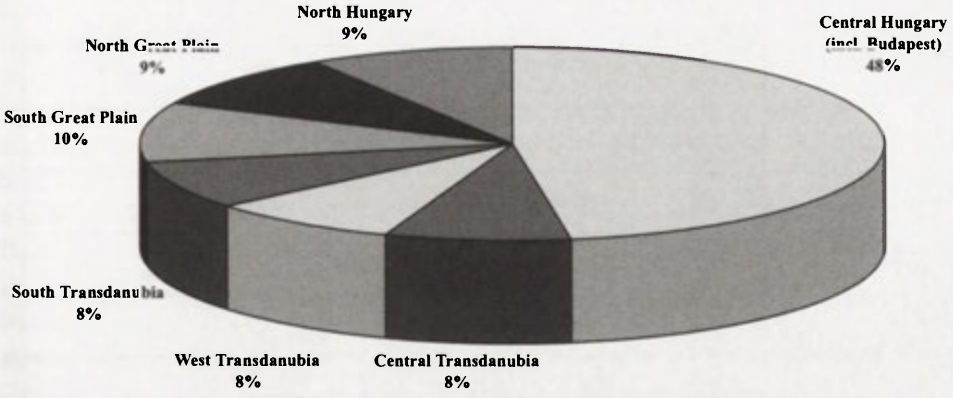


Figure 4. Regional distribution of employment in banking (1996).

Hungary is still the most centralised branch of the economy, with a definite centre in Budapest. The leading position of the capital is more striking in the financial sectors, especially banking and insurance, than in any others. Consequently local and regional banks are missing from the Hungarian banking system. (However, this strongly monopolistic structure is more in line with international trends, which are characterised by overconcentration at the global level; in contrast to in other transitional economies, such as Poland, where the role of regional banking is significant; Fig. 5).

More difficulties arose from the deformed spatial structure of the banking network of the early 1990s:

– The other marginal pole of the national banking system is the dense network of co-operative savings banks scattered throughout the countryside. The most important disadvantages of these are their weak financial standing (accounting for only 5% of the total balance sheet of banking) and lack of strong centres or headquarters. Despite there being 1,700 co-operative savings banks (accounting for 62% of the total national network), most of these small banks in the smaller towns and villages have a very low capital circulation and can supply only a narrow range of services.

– The lower density of the network meant both the low level of availability of branch offices and the greater structural polarisation of the branch network. On the one hand this meant that the rapid expansion of banking, initially concentrated almost exclusively in Budapest, was not followed by the extension of the branch network at a rapid pace in the countryside. On the other hand the new banks established in 1987 inherited a particular branch network from the National Bank of Hungary, since branches were missing from certain county seats, and this was accompanied by a spatial-regional asymmetry. The structure was even more distorted by the fact that the traditional retail bank (OTP-National Savings Bank) had usually had offices in all settlements where the population

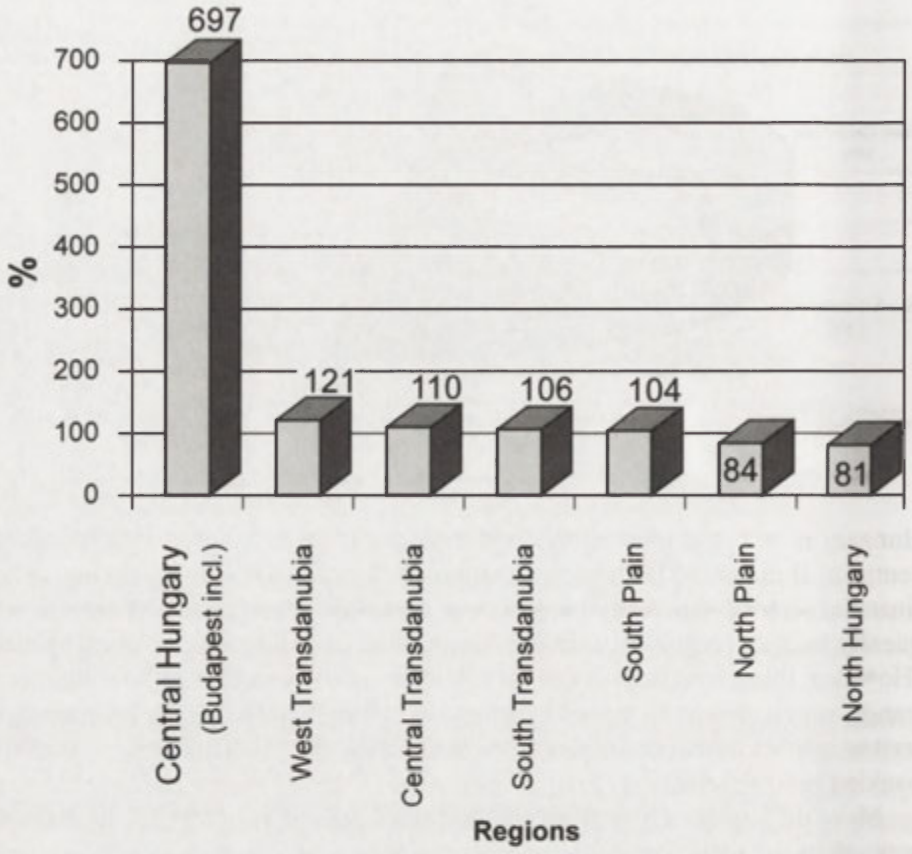


Figure 5. Regional GDP per capita in Hungary's financial-business sector as a percentage of the provincial average in 1997 (Provincial average = 100).

exceeded 5000, but the dynamically developing foreign banks have just started to expand their branch network in the last few years (Gál 1999b).

– The third reason for the polarity is that branches of banks based in Budapest have much less room for making independent decisions than the branches of county seats did during the communist period. Since the Hungarian banking system is characterised by an overcentralised management, controlling and structural system, branches are not in a real decision-making position, at least in part because they have only limited information. Most of the banks offer the same services all over the country and do not have a local advertising strategy. Banks do not usually lay stress on the uniform appearance of their branch offices; so appearances are only more greatly dependent on the hierarchical position of a certain bank.

The start of the 1980s and 1990s was the first period of boom in the establishment of banks: 17 commercial banks founded about 350 branches, concen-

trating 85% of the new offices into the provincial cities. The next period of the two-tier banking system between 1992 and 1996 saw the network restructured considerably. Expansion of the banking system was greatly restricted by the huge inherited debt imposing a large burden on the institutions. The smaller banks went bankrupt (Ybl Bank), while others were liquidated (Dunabank, Iparbankház) or merged. The big banks rescheduled their policy of network building and a few closed some of their branch offices, but the other banks such as Budapest Bank and Postabank began spectacular growth in network expansion. Accordingly, between 1992 and 1996 the number of banking institutions decreased due to bankruptcies, mergers and the purchasing processes of privatisation. In the two years 1995–1996 alone, six banks were liquidated or merged into other commercial banks.

The foreign-owned banks began to expand their branch network (by purchase through privatisation and opening new branches) more cautiously and then only after 1995. There are different reasons for this more cautious policy. On the one hand, these banks were strong enough in terms of capital intensiveness, so they could adjust the pace of network building to their own pace of development. On the other hand, foreign-owned banks were above all interested in corporate banking, supplying services to joint-stock companies. The boom period for the establishment of joint-stock companies was in 1990–1991 and thereafter the corporate market started to become saturated. However, the foreign-owned banks switched to rapid expansion through building their extensive branch network, gaining both larger market shares and leading positions in terms of profitability, and growing more rapidly than the bigger banks (Fig. 6).

Recently, the trend towards concentration has diminished due to the successful expansion of foreign-owned and medium-sized banks.[1] The balance of power in the banking system which held sway at the end of the first decade of two-tier banking will be expected to readjust in line with to growing competition for larger market shares. According to surveys, a shift from moderate deconcentration will emerge and the few large banks (from the group of medium-sized and the foreign-owned ones) of more major financial standing will dominate in the retail market. Besides these, 10–15 banks will play an important role in the banking system.

[1] While the share of the five largest banks in 1990 was 83% of total banking sector assets, by 1997 the five largest and oldest banks' share had decreased to 54%, and the ten largest banks, including newly-established and more dynamic institutions, accounted for 72% of banking assets.

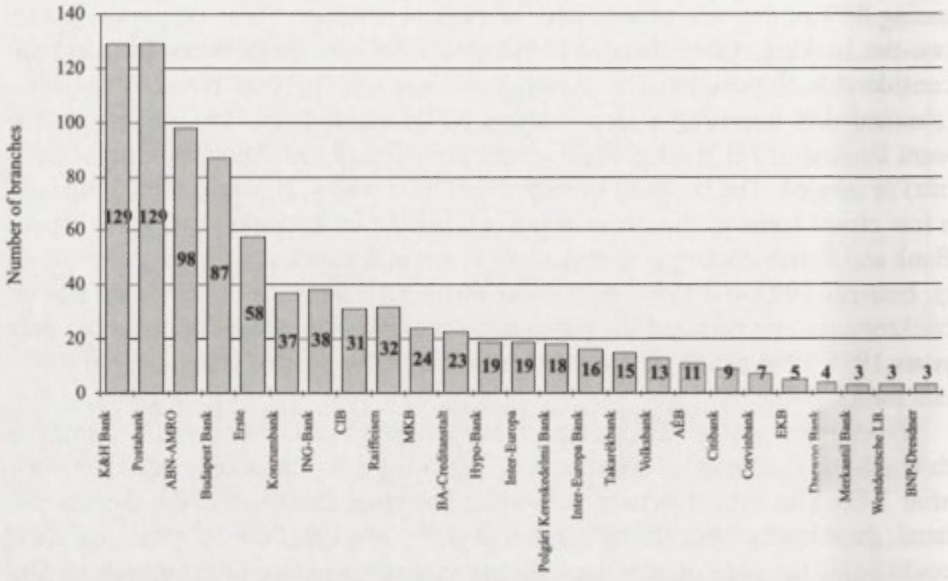


Figure 6. Branch networks of the Hungarian commercial banks in 1998 (429 branches of OTP Bank are not included).

SPATIAL DEVELOPMENT OF THE HUNGARIAN BANKING NETWORK

Regarding the diffusion of the banking network, it is very important to survey the geographical location and different hierarchical types of settlement in which banks are located.

At the birth of the two-tier banking system the network was characterised by a certain spatial balance due to the evenly-allocated branches of the OTP Bank (National Savings Bank), located in more than 270 settlements. After the foundation of the new commercial banks significant spatial asymmetry arose within the country, as certain banks were missing from particular regions and county seats: KHB (Commercial Credit Bank) dominates in the Great Pain region, MHB (National Credit Bank) in Northern Transdanubia and BB (Budapest Bank) around Budapest.

The spatial appearance and regional diffusion of the new branches of banks reflected Hungarian economic processes in the 1990s:

– The prevailing majority of economic associations, among then the joint-venture companies and accumulated capital outside Budapest flowed into the Transdanubian region, above all into the north-western part. All these are underpinned by indices for corporations, associations, household savings and by figures for indebtedness in the population.

– The structure of diffusion of the banking network had followed this spatial

pattern for the first time by the beginning of the 1990s. At that time banks were mainly interested in building up branches in the Transdanubian region. This was evident because the largest unexploited territories for financial services were situated in Western Hungary.

– Significant differences across the greater regions had practically been evened out, except in Northern Hungary, by 1990, and the disadvantaged status of the Transdanubian region came to an end. From the mid-1990s, following the saturation of Transdanubia, the larger cities of Eastern and Southern Hungary became the main targets of branch network expansion (Gál 1998).

There were significant differences behind the well-balanced greater regions concerning network density within the regions and counties. In some counties the number of new branches exceeded ten between 1978 and 1990 (Győr-Moson-Sopron, Baranya, Hajdú-Bihar), while in other places only a few branches were opened (Fejér, Komárom-Esztergom, Tolna) and in some counties they were confined to county seats (Borsod, Fejér, Szabolcs-Szatmár-Bereg). An extreme exception was Esztergom-Komárom County where no branch was opened between 1987–1990 in Tatabánya, the county seat, where economic depression affected its heavy industrial background. For instance, during the short period between 1995–1997 there was no increase in the number of branches in North-Western Transdanubia, as it was viewed as a saturated region.

A general characteristic of the period between 1992–1996 was the growing importance of Budapest in the expansion of the branch network (319 bank offices made up 26% of the national network in 1997). All banks starting to open new branches have opened 2–3 new offices in the capital city in the past five years, and last year 20 banks had branches there (Fig. 7).

Within Budapest most of the principal offices of banks are based in the inner districts. The spatial concentration of the institutions gives a strong impetus to the formation of a central business district, where the office buildings of banks became an important functional-morphological element of the townscape. In

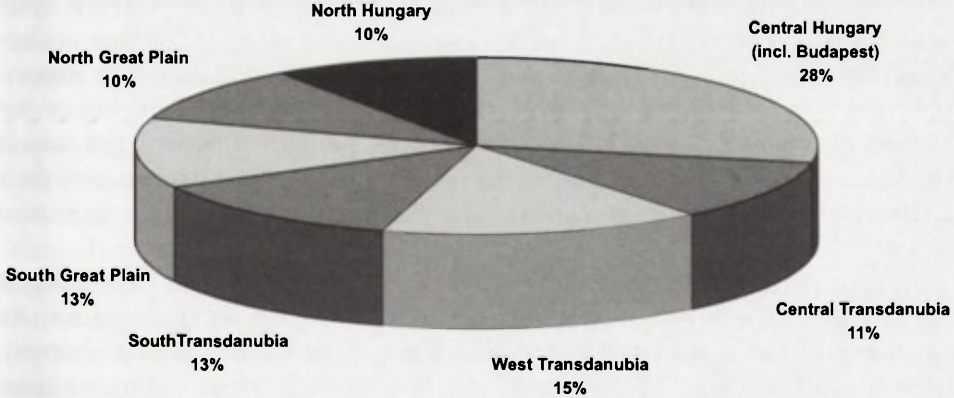


Figure 7. Distribution of the banking network among the regions of Hungary, 1996.

1990 about two thirds of the financial organisations were based in the 5th District, namely in the core area of the city centre itself, which is still the most popular domicile for new banks. By the end of the 1990s business (financial) functions of the 5th District had become saturated and a few years ago financial organisations started to diffuse towards the surrounding inner-city districts. Despite the expansion of banks, there is surprisingly low density of the network in Budapest, at one office per 7,758 inhabitants (15,000 without the OTP). This demonstrates unambiguously the limited degree of extension of the banking network in the capital city. The lack of banking services is more striking in the outer area of Budapest, resulting in overcrowded city centre branches.

The number of banking institutions in Hungary is 1,319, together with 1,700 co-operative savings banks, taking the total to about 3,100. Taking the figures for network density into account, there is one office per 3,200 inhabitants, which is still a much lower density than in Western European counterpart countries, where there is one bank per 1,400–1,500 inhabitants. In spite of the boom in the founding of new branches (a branch-office opening ceremony took place every week on average last year), mainly by foreign and joint-ventures banks, there are still too few branches in Hungary, although spectacular progress has been made, especially since 1996.

In surveying the banking network according to network-density figures, it is possible to find a few counties with lower density of banks. Szabolcs-Szatmár-Bereg and Pest counties are the most undersupplied areas, reaching only half of the national average in 1995. In the case of the former, the economic and geographical situation, activity of entrepreneurs, low level of foreign capital inflow etc. would seem to account for the smaller interest shown by the banks. In the case of Pest County, the capital city causes backwash effects which influence the development of the banking network. Relative to population, Hajdú-Bihar, Borsod-Abauj-Zemplén, Komárom-Esztergom, Nógrád and Fejér counties vere also badly supplied with banking services. These counties could be the main target areas for expansion in the near future (Fig. 8).

Surveying the distribution of the banking network by settlement type is more expedient than investigating at county level; all the more so as banking institutions have more links to the cities and towns, therefore capital flow is an important indicator of the different urban processes. Since, by the beginning of the 1990s, the number of branches had exceeded the number of larger cities, which had been the main targets of the earlier expanding banks, these banks tumed their interest towards smaller settlements. The first branches in villages vere also opened (Tab. 3).

Those banks have just recently started to develop their network - most of them are foreign-owned – situating themselves solely in regional centres. As a consequence of this, certain larger cities (Pécs, Győr, Szeged, Székesfehérvár), despite not being seats of a regional bank, have started to play significant roles in the operation of financial services in which different organisations of the

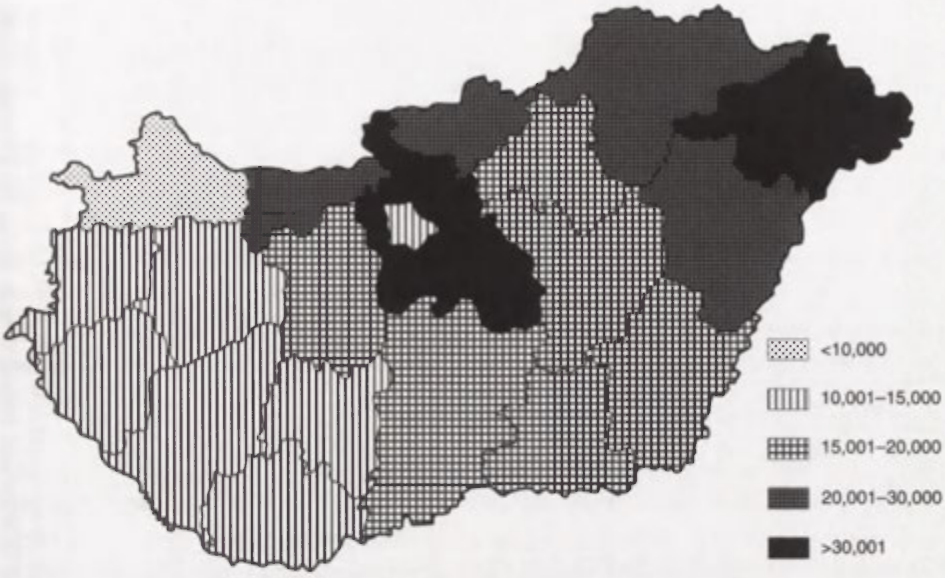


Figure 8. Density of the Hungarian banking network in 1995
(Number of inhabitants per office).

financial sector (banks, insurance companies, consultancies) attract each other mutually. This also induces increased competition in the local-regional market.

At the beginning of the 1990s the banking network was rather more polarised, both hierarchically and regionally, than nowadays. A more developed network existed in the county seats and in the cities of Western Hungary (which were targets of foreign companies and banks); while in Northern Hungary and in the northern part of the Great Plain the banking network is less developed than in Pest county, where the central role of Budapest counterbalances its disadvantage. In recent years a shift has taken place, levelling out the expansion of the banking network in favour of the eastern parts of the country. During these years the number of branches in the cities of Eastern and Southern Hungary increased more rapidly than in the western counterparts which were previously the most saturated parts of the country, in terms of the number of branches. The network building expansion of branches can initially be said to have followed the pattern of the spatial-economic division of the country, as banks were mainly opening branches in Western Hungary. Since the mid-1990s, following the relative saturation of West-Hungary and owing to the process of nivellation, banks have started their expansion towards the eastern and southern parts of the country along the urban hierarchy. While two years ago, Győr, Pécs, and Székesfehérvár were considered the largest financial centres outside Budapest, Miskolc has recently gained the leading position in terms of the number of branches (37), followed by Győr and Kecskemét (each with 32 branch offices), then Pécs and Szeged (31–31), and

Table 3. Banking services in Hungarian cities, 1996.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Hungary	10,212,300	255,043	305	3,216	630	3,175	16,210	79	405	249.7
Budapest	1,906,798	108,865	64	253	125	7,537	15,254	430	871	570.9
Countryside	8,305,502	146,178	278	2,548	505	3,260	16,447	57	289	176.0
Cities	2,082,932	61,951	–	436	274	4,777	7,602	142	226	297.4
Cities of the Danubian counties*	736,898	22,225	127	199	305	4,107	8,215	118	230	296
Debrecen	214,228	5,481	20	29	16	7,387	13,389	189	343	255.8
Dunaujvaros	57,970	1,508	5	8	4	7,246	14,493	189	377	260.1
Hodmezovasarhely	50,631	4,374	5	8	4	6,329	12,658	547	1,094	863.9
Miskolc	193,905	4,280	19	31	18	6,255	10,773	138	238	220.7
Tatabanya	75,258	1,679	9	13	5	5,789	15,052	129	336	223.1
Salgotarjan	48,488	1,060	8	9	6	5,388	8,081	118	177	218.6
Szolnok	81,336	2,213	10	16	10	5,084	8,134	138	221	272.1
Nagykanizsa	55,624	875	6	12	6	4,635	9,271	73	146	157.3
Nyíregyhaza	119,317	3,650	16	26	14	4,589	8,523	140	261	305.9
Szekesfehervar	108,543	3,625	17	25	17	4,342	6,385	145	213	334.0
Kaposvar	70,529	2,135	11	17	10	4,149	7,053	126	214	302.7
Szombathely	84,745	1,996	16	22	12	3,852	7,062	91	166	235.5
Szeged	173,820	7,267	21	46	20	3,779	8,691	158	363	418.1
Pecs	164,872	5,595	25	45	23	3,664	7,168	124	243	339.4
Kecskemet	108,345	3,413	19	30	15	3,612	7,223	114	228	315.0
Zalaegerszeg	62,077	1,663	13	19	14	3,267	4,434	88	119	267.9
Gyor	130,244	3,725	26	43	23	3,029	5,663	87	162	286.0
Eger	60,445	1,345	15	20	10	3,022	6,045	67	135	222.5
Bekescsaba	67,621	1,638	13	23	13	2,940	5,202	71	126	242.2
Sopron	54,311	1,338	12	19	9	2,858	6,035	70	149	246.4
Veszprem	63,553	1,749	14	25	15	2,542	4,237	70	117	275.2
Szekszard	37,070	1,342	12	16	10	2,317	3,707	84	134	362.0

* Dunaujváros, Tatabánya, Székesfehérvár, Pécs, Kecskemét, Győr, Sopron, Szekszárd.

(1) Population as of 1st January, 1996, (2) Number of enterprises with legal and non-legal entity, (3) Number of banks, (4) Number of all institutions of banking network, (5) Number of all institutions of banking network counted without OTP Bank and mutual savings banks, (6) Network density 1: population decline per branch, (7) Network density 2: population decline per branch without branches of OTP Bank and mutual savings banks, (8) Network density 3: enterprise decline per branch, (9) Network density 4: enterprise decline per branch without the branches of OTP Bank and mutual savings banks, (10) Enterpreneurial activity: enterprise decline per population of 10,000.

Sources: *Statistical Yearbook 1994*, Hungarian Almanac of Finance and Capital Market 1995–1996.

finally Debrecen (28).[2] The main targets are the large cities in East and South Hungary, such as Miskolc, Szeged, Debrecen, and Nyíregyháza which have gained a temporary leading position in size of the local network (Fig. 9).

Because of the centralised structure of banking, the Hungarian banks, aim at complete coverage of the relatively small banking market. This tendency promotes equalisation among the different parts of Hungary. However, differences occurring in the number of branches do not mean differences in the quality of banking services. The latter is much more dependent on territorial embeddedness, which can induce mutual attractiveness toward other types of financial and business services. The agglomeration of financial services not only generates competition in the local market but might also result in the performance of certain regional financial centre functions in larger cities even despite a lack of locally-based institutions.

In conclusion, different banks can be seen to have become situated on different levels of network construction in the recent period of development. The share taken by larger cities within the banking network increased markedly from 35–40% to about 50% (Budapest 66%) between 1987 and the early 1990s, owing to the fact that at least a dozen new banks came onto the market and started network development. Bankruptcies and the rationalisation policy of network development in the following period mainly affected these larger cities, as the major beneficiaries of the boom in banking expansion. Despite the opening of several new branches, the proportion of larger cities within the banking network fell to 43% (or 63% including Budapest), in parallel with diffusion of the network towards smaller settlements.

Taking the expansion of the banking network into account, some experts believe that the spread of electronic home banking will counterbalance the traditional means of branch office building. According to others, whose opinion I share, there is a brighter future for the traditional expansion of the branch network since customers are much more devoted to a personal style of administration and rely more on branch offices. Although virtual banking is likely to be widespread in the future, the building of a more cost-intensive branch network is still very important. Moreover, about 40% of the population as yet has no contact with banking. The figures for the year 1998 justify both these theories on the future prospects of banking (Bonin et al. 1998).

THE CHALLENGES OF REGIONAL BANKING

Study of banking history reveals the wide variety to the development of different national banking systems. These systems currently express spatial diver-

[2] In other banking institutions such as the Treasury, representative branches of NBH, regional headquarters, single branches, representative or cash offices, correct spatial differences derive from the number of banks or branches. Regional centres (Győr, Pécs) and greater county seats (Kecskemét, Székesfehérvár) usually have the complete range of these institutions.

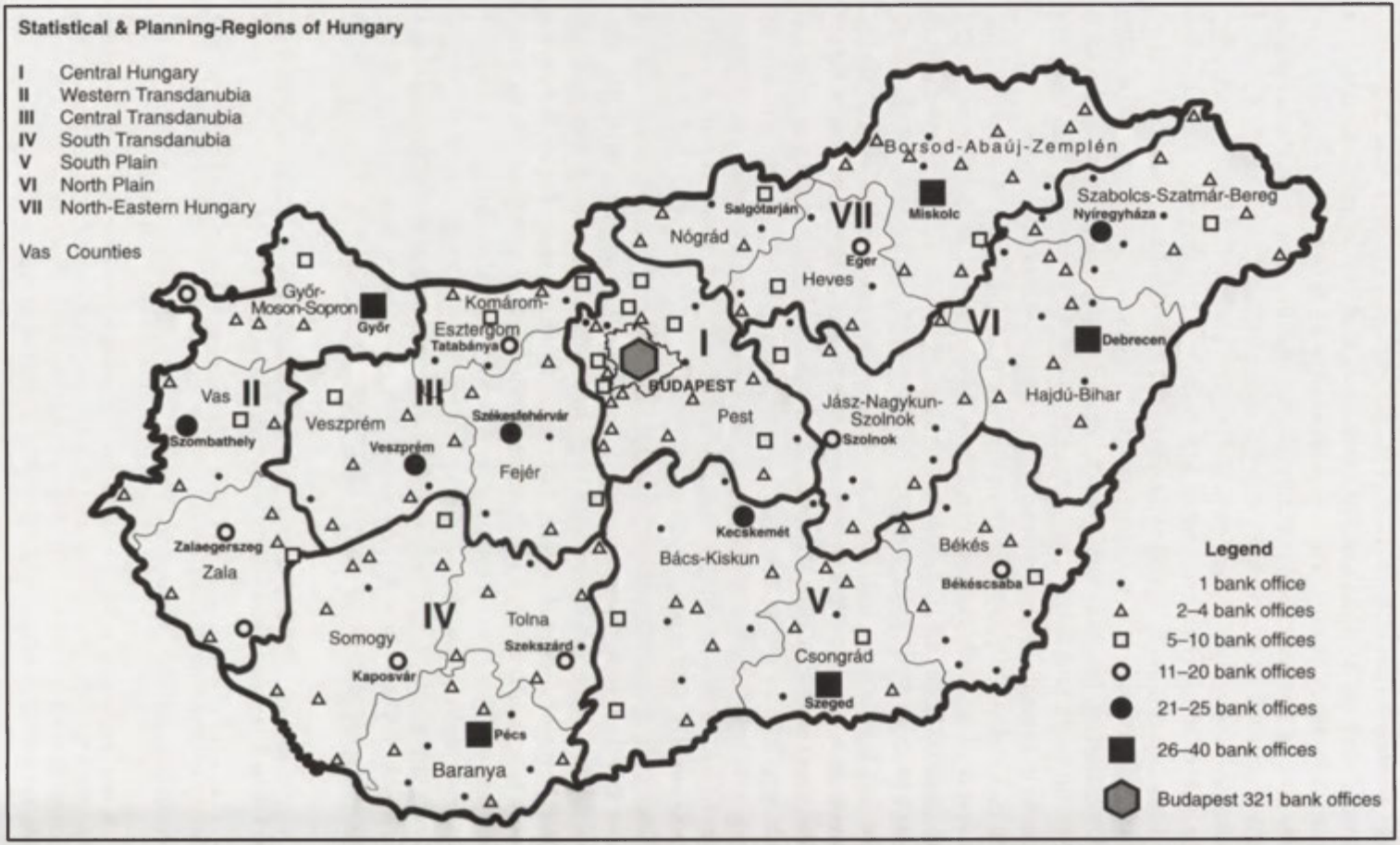


Figure 9. Regional distribution of branch network in the Hungarian towns, 1999
(included banks headquarted in Budapest, excluded cooperative savings banks).

sity arising from the particular location of a distinctive financial centre and from differences in spatial structure and the origins of particular national and regional banking. The Hungarian banking system is characterised by a lack of regional banks and strong agglomeration of services in the financial centre of Budapest that can be partly explained by accommodation to the conditions of the more concentrated international financial markets which are currently undermining and diminishing the role of the local markets. However, recent moves, both toward supra-national economic, political and monetary unions and towards secession and regional autonomy, have tended to undermine the usefulness of the nation state and simultaneously strengthen the role of locally-based regional units. In contrast to the concentration processes in the global markets, the growing significance of European regionalism requires establishment of regional money markets and institutions financing regional policies. Globalisation and the emergence of global financial spaces may actually serve to open up opportunities for local--regional alternatives (Lee 1999; Porteous 1996):

1. Money is sensitive to local differences in economic return (crises still have a distinctly local origin) and local, regional banking systems tend to be more rooted in and committed to the local economy and community than local branches of centralised national or international banks.

2. The effects of financial crises increase the instability of global markets and seriously question the regulative and controlling role of the nation-state over capital flows.

3. Decentralised banking systems have been important in certain European countries, such as Germany, Italy, France, and Poland.

4. The boom in private enterprise, and privatisation, the necessity of a presence in the local markets and competition for retail markets also require the expansion of the banking network in the regions. Regional banks may well be serving the interests of local economies and SMEs better than financial-centre banks whose priorities relate more to national and international markets.

5. Besides the corporate and the retail market, project financing will be another business for banks which support regional development programmes through financing infrastructural, power and telecommunications investment in co-operation with regional and local administrations.

A survey of how money moves between locations and regions raises the problems of integration between the global and local level, or between centre and periphery, that concern an irregular financial division of labour between central and peripheral areas. The emergence of uneven economic development among regions is to a large extent caused by uneven interregional capital flows. Capital is mobile across regional boundaries and usually flows from regions with lower profitability into regions offering a higher rate of return. Consequently, capital is concentrated into the financial centres of the core areas, a phenomenon which can result in regional inequalities within the single European markets as well (Porteous 1996; Leyshon and Thrift 1997).

The extremely concentrated national and international financial markets and the lack of strong regional markets might slow down regional economic development in the long run because of different factors stated below:

– Certain national and international banking centres discriminate against particular regions. This refers usually to credit discrimination which means that nation-wide banks are less prepared to make credit available to agents in the periphery because they allocate loanable funds based on an implicit regional reserve ratio. Core regions and financial centres may through the centralised branch banks invest the savings drained out of the peripheral regions in favour of lending in core regions, slowing down development in the peripheries and resulting in spatial polarisation of the regions (Chick and Dow 1988).

– National banks can only slowly acquire local embeddedness. The distance between decisional and operative centres within a national branch bank structure reduces the availability of information about local firms and local growth prospects. The uniform lending standards by nation-wide banks affect certain regions disproportionally. Credit is made available on the same terms in different locations through the branch system regardless of specific regional requirements and conditions.

– In a national branching system, local branches may adopt a more cautious and restrictive lending policy as they are likely to be constrained by head office in their degree of freedom, as most of the strategic decisions are made at headquarters. Because of the centralised structure the decision-making autonomy is limited: large loans require head office approval. Local branch management of national banks is often in the hands of directors only temporarily committed to that branch who tend to be very risk averse, opting for safe large investment, rather than riskier smaller investment, even to the detriment of important innovative projects for the growth of the local economy. On the other hand, if a national bank is seeking to rationalise its operations, it is likely that the branches in peripheral and economically-declining areas are the first to be closed down (Porteous 1996).

– Within a centralised banking network, information asymmetries often also occur: the headquarters very often assess a higher risk due to poor information on small borrowers in remote or peripheral regions, and because of market segmentation. The larger distance between peripheral regions and the core centres gives rise to larger costs of transactions and monitoring and may result in more expensive credit. (McKillop and Hutchinson 1991). Information and transaction costs of the supply side are higher in relatively isolated regions for lenders based in the core. Although the cost of credit may be equalised across regions in an integrated banking system, the availability of credit may differ, something which continues to limit access to the credit supply.

The solution in the longer term should be the reorganisation of the institutional and managerial structure of the banks and the determination of the kinds of banks (local, regional or national) that are best suited to fostering the develop-

ment of peripheral regions. On the other hand there is a strong need to create decentralised financial sources and establish regional financial centres in order that the interests of local economies might be better served, since regional banking systems represent a link between local economies and financial centres. This highlights the problems of integration between the global, national and regional level. The integration model is a kind of reaction to theories focusing exclusively on "localism" and "globalisation". In the first, it is considered to be detrimental for banks from outside a region to set up, through the opening of branches, mergers, or the purchase of local banks. The localism theory is based both on the notion of local segmentation of financial markets and the idea that the savings of a region be kept within that boundaries of that region. This latter can be counter-productive as savings must be free to move in search of the best investment opportunities and returns in a wider unified monetary system. Local banks have a more limited ability to invest savings in the same area in which they were collected, as local banks very often tend to have greater willingness to export and invest capital – under better conditions of return – out of the region, than do local branches of external banks. Therefore, the main challenge for a region is to offer the best opportunities for investment, attracting capital. On the other hand, the theory on globalisation leads to the argument that global integration of financial markets removes regional disparities in financial structures and capital availability. In fact, the advantages of globalisation are not distributed evenly among regions, as they tend to be located in the stronger and better-organised ones. So, in the absence of corrective policies, regional disparities could become wider rather than narrower (Alessandrini and Zazzaro 1999).

The best way to solve the integration problems is through the co-existence, complementarity and interaction of different regions, and between the centre and the periphery. This reorganisation can take place through the passive integration that arises outside the region and means not only capital inflow, but also the entry of non-resident banks opening new branches or incorporating local banks. On the other hand, regional, local banks, can participate in an active way in inter-regional expansion, which allows local banks to open up to the outside without abandoning their own regional hinterlands. It is important for a regional banking system to compete with those in other areas in order that the benefits of both regional and sectoral diversification might be obtained. The inter-regional integration of the banking system is the most suitable model for future development in Europe and can be adopted in part by Hungarian banking. It offers a perspective of the development of regional finance partly through branches of national banks channelling innovation and providing a wider range of services and a strong equity background, in order that less-prosperous local branches might be protected. On the other hand, the modernisation of existing local banks and the creation of new institutions in the regions meets the challenges posed by technological, institutional and regulatory changes that are transforming the world of finance and serving the needs of the local economy and communities. The most

important step towards this system is the establishment of a strong regional financial centre that can serve the interests of a particular region:

– Regional centres have an ability to capitalise on localised information spillovers, thereby reducing the costs of local lending, and they are closer to the actors in the local economy. The more centralised firms lose their competitive edge if their headquarters are located far from the region where they intend to operate.

– A prosperous regional centre can prevent capital drainage from the region by the national financial centre.

– A regional financial centre which is closer to its hinterland may have lower fixed costs and therefore it be better able to serve SMEs.

– The traditional national financial centres face decentralisation forces as their costs of operation increase (high labour costs, real estate and renting prices), and as rapid development in communications technology appears to favour decentralisation by allowing large volumes of information to be accessed ever more easily from remote locations at low cost.

The Hungarian banking system is characterised by strong spatial concentration. The leading role of Budapest is unique even in the European context. The fact that every bank is headquartered in Budapest results in a deformed structure in the banking system. Banking in Hungary is still the most centralised branch of the economy. The conditions of capital concentration are unfavourable outside Budapest and the most developed regions of the country. There is a threat of new kind of dependence between the capital city and the regions: filtering-down persists, the central region, making use of advantages of its location, filters the most profitable lines of banking (corporate, portfolio management, risk management, private banking) and diverts to the peripheries the traditional, uniformed and less profitable services. Taking into account capital transfers between Budapest and the regions (regarding banking capital and the central budget flows) it can be said that there is capital transfer at the expense of the regions.

There are four principal tasks on the agenda for the development of a more-decentralised banking network in Hungary. First, it is necessary to expand further the density of the branch network and to extend the range of branch services of commercial banks in there should be regions. Second, there should be formation of regional and municipal banks. Third, regional branches of the Hungarian Regional Development Bank should be set up. Fourth, the institutional connection of Budapest to international money markets must be assured (Gál 1999a).

1. In Hungary the local-regional credit supply operates through the centralised national branch-banking system, with local savings banks operating in restricted rural and urban areas. National banks with branches do not usually provide an adequate credit supply for local SMEs and do not finance municipal projects and infrastructural investments in the regions. National banks are not interested in these less profitable and prudent businesses as they have a different orientation of profile and tasks. Therefore they often seem to discriminate against

particular regions. Recently, commercial banks with a larger network have reorganised their institutional and managerial structure to form a hierarchically-built domestic network of branch offices, decentralising certain control functions (Fig. 10). Banks are starting to pay much closer attention to the geography of their distribution networks. Besides a national head office they form regional control offices and local branch offices with various functions, in order to rationalise their dispersed structures and take the first steps towards decentralisation and inter-regional integration of banking, as well as to use their resources in a more productive way. Regional control offices play the role of intermediate tiers between head office and local branches, having authority over local branches in their geographical areas. These newly-created, regionally-controlled territories for large banks are different from each other, and have not overlapped the territories of the statistical regions, but their regional headquarters have in most cases been concentrated in regional centres. Recently banks have been in the expansion phase of their branch-building process. In a few years time they will have to concern themselves with network restructuring, which is sensitive to geographical variation in profitability, risk, debt and social conditions in a particular area (Geenhuizen 1999).

2. Within the banking system, the formation of locally-based regional and municipal banks independent from the national branch network would be very important with a view to the regional interests of that region in which they operate being served and regional economic development assisted. This requires the amendment of the Banking Act, which would allow municipalities, chambers of commerce, economic organisations and private enterprises to establish banks with strong state support, in order that provision of public duties be made to derive from state commission. This can be the financing of public investment, credit supply for local governments, and the intermediation of EU structural funds. These banks can follow the pattern of the German *volksbanke or sparkasse* and might be based on stronger Hungarian co-operative savings banks. These institutions will be the major financial agents of the municipalities, and will be better able to serve smaller businesses and promote the direct integration of regions located at longer distance from the cores into the global economy. Only a strong regional bank can ensure adequate credit circulation and prevent mass capital outflow from a particular region (Illés 1993).

3. Because of the different interests of commercial banks as stated above, only a decentralised developing regional bank network can promote regional development effectively. The establishment of regional branches of the Hungarian Development Bank Ltd. ('MFB') is therefore necessary. MFB, formed on 1 January 1997, has been the leading institution as the legal predecessor of the wholly state-owned credit institution formed in 1991. MFB's original task was to facilitate the modernisation and invigoration of the Hungarian economy, to participate in regional development, to manage and mediate state funds and those allocated for development purposes, and to raise funds in international markets

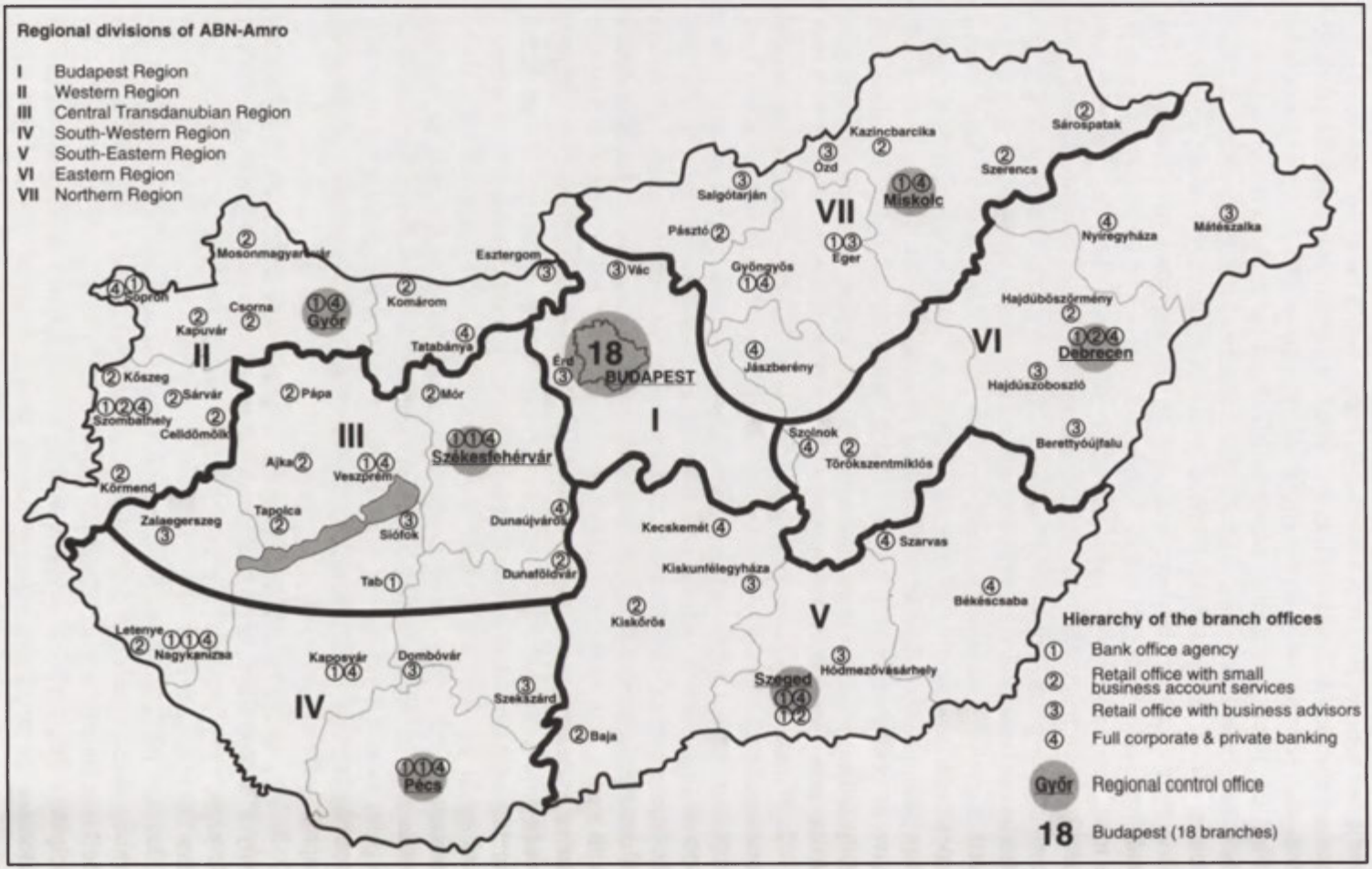


Figure 10. ABN-Amro Bank: The Hungarian branch network and its regional divisions, 1999.

and finance the regional development programmes of municipalities. Furthermore, MFB pays special attention to supporting small and medium-sized enterprises, a sector that plays an important role in the development of the Hungarian economy, carrying out large-scale capital expenditure projects in the fields of infrastructure, telecommunications, and the energy sector.

The banks operating and mediating regional development funds (PHARE, ISPA, SAPARD etc.), have to form a network of regional development banks at regional level, in order to mediate European development funds into the level of use. Here it is not enough to create different frameworks of support; there is also a strong need to establish institutions mediating funds between the EU/state level as the provider and regional levels as the destination. In addition, one of the main tasks of the regional development banks is the promotion of regional projects and investments from the financial side. This includes the provision of medium and long-term loans to SMEs, portfolio management and investment advisory services and promotion of the creation of technology centres, incubators and innovation-oriented enterprises.

4. A question as to whether Budapest will become a real financial centre by international standards thus arises. Recently, large cities and different regions, rather than simply different nations, have been in competition with each other in the field of the global world economy in order to gain investment capital, and to connect with sources of information. As a consequence of the rapid restructuring and modernisation of Budapest's economy, the capital city has become one of the most important innovation-centres of the region, one which might serve as an important bridge-head in foreign capital inflow and investment within Central and Eastern European countries. Budapest has a traditional metropolitan townscape, adequate infrastructural background and stable economic environment – all quite important attractive forces for the investment of multinational companies. Many multinational companies (Pepsi Co., Kodak, Nestle, Xerox, Shell) have built regional bridgeheads facing Eastern Europe from Budapest. On the other hand, there are certain limits to the growth of such international financial functions in Budapest since telematically-based concentration processes, which are characteristic of global money markets, could overcompensate for the advantages of geographical proximity. Multinational companies are most likely to utilise only the simpler financial services in the Central European region, while the services requiring more resources will be utilised in traditional Western European and overseas financial centres in the future also. In addition, the smaller size of the national financial market, the weakness of domestic capital (the activity of the black economy), the low level of economic interactions within the regions, the limited activity of the banking system abroad and the consequently smaller size of banks (smaller provisions for expected liabilities) will all render Budapest unsuitable for the role of regional financial centre. Foreign banks that have opened their subsidiary banks in Hungary have established branches and subsidiaries in other Central European countries, too. Consequently, foreign

banks concentrate more on covering each national market than on establishing a single regional banking centre, for instance, in Budapest.

According to some banking experts, Budapest could only successfully apply for the position of a subordinate offshore-like regional financial centre specialised in certain services. Subject to these conditions, services which require smaller amounts of capital and highly-qualified employees will come into prominence. To render these, it will be necessary for the banking system to be augmented by business-like intervention on the part of the state, but the exact date of integration into the EU and EMU may influence the development of the Hungarian banking system and the international role of Budapest (Bellon 1998).

CONCLUSION

In evaluating the competitiveness of the Hungarian banking system, it can be said that the banking sector has been strengthened since 1994 and has become a more profitable sector. However, the progress in banking is significant only as compared to the previous state of the banking system: by international standards the quality of the sector is still very low. The proportion of outstanding claims, despite a significant decrease, remains higher than in any country of the European Union. The role of the Hungarian banking system in the economy lags far behind that in more-developed countries. The ratio of balanced-sheet footings to GDP is 72%, a very low percentage, but one which indicates an enormous opportunity for progress in the banking market. Despite the general recovery of banking, the sector has remained polarised. In 1995, banks with foreign ownership accounted for one quarter of the market while producing 70% of the profit after taxes. Their profitability and efficiency was twice as high as at banks of Hungarian ownership (Várhegyi and Gáspár 1997).

A survey of the spatial characteristics of the Hungarian banking system shows that economic changes are very much dependent on financial services, which reflect the processes of economic transformation. Financial services have become the key sector of business services differentiated by spatial and regional development characteristics as well. The spatial structure of the banking sector in Hungary is characterised by a large-scale concentration in Budapest, but a boom in the founding of branch offices is also typical in the countryside, as the necessity of a presence on local markets (collections of resources and credit allocation etc.) is clear, as is the competition for the retail market to stimulate banks to build up their national networks.

From the mid-1990s on, following the relative saturation of West Hungary, the main targets have been large cities in East and South Hungary. Because of the centralised structure of banking, the Hungarian banking-institutions aim at covering the relatively small Hungarian banking market completely. This tendency promotes a kind of spatial equalisation among the different part of Hungary.

At the same time, the commercial banks which have their headquarters in Budapest have largely confined themselves to the collection of deposits in their national network, with the result that there is capital drainage and net capital loss from most of the regions. The presence of the centralised capital market and the lack of a decentralised regional banking system can restrain and slow regional development in the long run. In Hungary local and regional credits can only be received through the centralised network of bank offices and to a limited extent through the weak local co-operative saving banks. The Commercial banks do not provide local entrepreneurs with adequate credits, and do not finance larger local municipal projects and regional infrastructural investments either.

Within the centralised banking structure, the regional decentralisation of certain commercial banking services is possible, without questioning the pre-eminent role of the national banking centre, but contributing to more efficient operation of the network. The example of the British banking system demonstrates the development of a series of provincial financial centres, through the decentralisation of certain back-office functions into particular cities serving the operation of a more efficient and autonomous banking network which contributes to the success of regional economies (Tickell 1996).

In Hungary, the formation of regional banks – or at least the functional and institutional reorganisation of the national branch network of the large commercial banks – is very important at regional level if the regional interests of that region in which they operate are to be served and regional economic development facilitated. The pattern of German banking well demonstrates a decentralised banking system in which small local or regional banks (Volksbanken, Raiffeisen banks) and large banks have coexisted for a long time. Banking activities were restricted by law to particular regions, which limited the capacity of banks to expand through their geographic hinterland and so prevented the gradual centralisation of control from outside. In addition, regionally-based banks have a fruitful interregional co-operation with larger banks, resulting in the availability of a cheaper credit supply for regional financial markets.

Larger countries in Europe (Germany, Poland, Italy etc.) show a rise of strong regional centres, based upon independent regional banks or decentralised back-office functions in certain provincial centres holding firm positions in large regional markets or based upon specialisation and complementarity. In contrast to this, Hungarian banking exemplifies the pre-eminent position of the national financial centre, in part because of the much smaller market size and the weakness of regional economies. It seems plausible that there is no place for such strong regional financial centres in a small domestic market and the small geographical areas of the created regions, but finding the right way for a certain decentralisation in the banking sector is a necessity.

The importance of blocks of financial intermediaries being based in the regions is not solely a reflection of the employment generation effects, but is also to do with the creation of a critical mass of financial activity. This in turn attracts

other financial institutions and eventually leads financial sector growth to become self-perpetuating, thereby stimulating regional economic growth.

Summarising the experiences of Hungarian banking, it can be said that the openness of this sector compared to others has contributed more to the modernisation and competitiveness of the overall banking system. Until now, the activity of foreign banks has depended on their subsidiary companies, which have to be established before operations start. In the last instance, the question becomes whether the national banking system is ready for full liberalisation and is able to withstand increasing competition within the European Union as cross-border banking is introduced.

REFERENCES

Ábel I., Siklós P., Szekely I., 1998, *Money and Finance in the Transition to a Market Economy*, Cheltenham, UK, Northampton, USA, Edward Elgar.

Alessandrini P., Zazzaro A., 1999, *A „Possibilist" Approach to Local Financial Systems and Regional Development: The Italian Experience*, [in:] Martin R. (ed.), *Money and the Space Economy*, John Wiley & Sons, Chichester, 71–92.

Anderson R., Kegels C., 1998, *Transition banking, Financial Development of Central and Eastern Europe*, Clarendon Press, Oxford.

Az Allami Penz-es Tokepiaci Felügyelet éves jelentése, 1996, (*Annual report of the Bank and Capital Market Advisory Board, 1996*), Budapest, APTF, June, 1997.

Bácskai T., 1997, *A magyar bankrendszer: Már tíz éve ketszintű, (The Hungarian banking system: A two-tier system already for a decade*), Bankkorszak 1987–1997. A "Figyelő" különszáma, 7–9.

Bellon E., 1998, *Lesz-e Budapest nemzetkozi pénzügyi kozpont? (Will Budapest be an international financial centre?)*, [in:] Glatz F. (ed.), Budapest – International Cty, MTA-HAS, Budapest, 59–82.

Bonin J., Mizsei K., Szekely I., Wachtel P., 1998, *Banking in Transition Economies, developing market Oriented Banking Sectors in Eastern Europe*, Cheltenham, UK, Northampton, USA, Edward Elgar.

Central Europe's Largest Banks, Central European Economic Review, 1999, October, 7.

Chick V., Dow S., 1988, *A Post-Keynesian Perspective on the Relation Between Banking and Regional Development*, [in:] Arestis P. (ed.), *Post-Keynesian Monetary Economics*, Aldershot, Edward Elgar, 219–250.

Dow S., 1999, *The Stages of Banking Development and the Spatial Evolution of Financial Systems*, [in:] Martin R. (ed.), 1999, *Money and the Space Economy*, John Wiley & Sons, Chichester, 29–30.

Gál Z., 1998, *The Questions of Financial Service Sector in Hungary from the Point of View of Regional Policy*, Tér és Társadalom (Space & Society), 12, 43–67.

Gál Z., 1999a, *The Hungarian Banking Sector and Regional Development in Transition*, [in:] Hajdú Z. (ed.), *Regional Processes and Spatial Structures in Hungary in he 1990s*, Centre for Regional Studies, Hungarian Academy of Sciences, Pécs, 180–221.

Gál Z., 1999b, *The Hungarian banking system and regional development in transiton* (1987–1997), [in:] Dúró A. (ed.), *Spatial Research in Support of the European Integration*, Discussion Paper Special, Centre for Regional Studies, Hungarian Academy of Sciences, Pécs, 69–85.

Geenhuizen M., 1999, *Shifts in Urban Hierarchy? The Case of Financial Services in the Netherlands*, Geographia Polonica, 72, 107–124.

Illés I., 1993, *Bankrendszer és regionális fejlodés, (Banking system and regional development)*, Kézirat (Manuscript), 15.

Jakus I., 1997, *Bankfiókhálózatok: A végek ígérete, (Branch network of banking)*, Heti Világgazdaság, 33, 77–78.

Lados M., 1991, *A pénzintézeti rendszer kialakulása Magyarországon, (Formation of banking system in Hungary)*, Területi terjedés, területi fejlodés (Pünkosdfürdő), 1–13.

Laulajainen R., 1998, *Financial Geography*, Series of Gothenburg University, Departments of Geography, 93, School of Economics and Commercial Law, Goteborg.

Lee R., 1999, *Local Money: Geographies of Autonomy and Resistance?*, [in:] Martin R. (ed.), *Money and the Space Economy*, John Wiley & Sons, Chichester, 71–92.

Leyshon A., 1995, *Geographies of money and finance I*, Progress in Human Geography, 19, 531–543.

Leyshon A., Thrift N., Toomey C., 1988, *The south goes north? The rise of the British provincial centre*, Working Paper on Producer Services, 9, Department of Geography, University of Bristol.

Leyshon A., Thrift N., 1997, *Money/Space: Geographies of Monetary Transformation*, London, Routledge.

Martin R. (ed.), 1999, *Money and the Space Economy*, John Wiley & Sons, Chichester.

Martin R., 1999, *The New Economic Geography of Money*, [in:] Martin R. (ed.), *Money and the Space Economy*, John Wiley & Sons, Chichester, 207–224.

Mazucca R., 1993, *A bankhálózat területi szervezete Olaszországban, (The spatial structure of the banking network in Italy)*, [in:] Horváth Gy. (ed.), *Régiók és városok az olasz modernizációban*. Pécs, MTA RKK, 0 307–317.

McKillop D. G., Hutchinson R. W., 1991, *Financial Intermediaries and Financial Markets: A United Kingdom Regional Perspective*, Regional Studies, 25, 543–554.

Porteous D., 1996, *The Geography of Finance*, Aldershot, Avebury.

Területi Statisztikai Évkönyv, 1995, (Regional Statistical Yearbook, 1995), Budapest, KSH, 1997.

Tickell A., 1996, *Taking the initiative: The Leeds financial centre*, [in:] Haughton G., Williams C. (eds), *Corporate City?*, Avebury, Brookefield, 103–118.

Várhegyi E., Gáspár P., 1997, *A tőkemozgások szabadsága – az Európai Unió kihívásai, (Freedom of Capital Movements – Challenges of the European Union)*, Pénzügykutató Rt. Perfect Rt. Budapest.

Várhegyi E., 1997, *A magyar bankrendszer privatizációja, (Privatisation of the Hungarian banking)*, Külgazdaság, 4–18.

Wachtel P., 1997, *A külföldi bankok szerepe a kozép-európai átmeneti gazdaságokban, I–II, (Role of foreign banks in transitional countries of East-Central Europe, I–II)*, Kozgazdasági Szemle, 1. 13–30; 124–141.

Warf B., 1999, The *Hipermobility of Capital and the Collapse of the Keynesian State*, [in:] Martin R. (ed.), *Money and the Space Economy*, John Wiley & Sons, Chichester, 227–240.

GEOGRAPHIA POLONICA, 73, 1, Spring 2000
PL ISSN 0016-7282

INTERNAL MIGRATION IN TODAY'S JAPAN

SATOSHI NAKAGAWA
Department of Economic Geography, Kobe University
Rokkaodai-cho 2-1, Nada-ku, Kobe, 657-8501 Japan
e-mail: satoshin@rose.rokkodai.kobe-u.ac.jp

ABSTRACT: Studies of Japan's internal migration have developed over the last few decades in accordance with changing research methodologies. In particular, the migration between metropolitan and non-metropolitan areas has drawn the attention of researchers. Japanese geographers have observed changing migration patterns, i.e. migration turnarounds in post--war Japan, and have tried to explain them in connection with economic determinants and cohort size in young adults. The obvious inflow into metropolitan areas before the mid 1970s was mostly explained by economic indicators, while the changing cohort size of young adults based on the baby boom in the late 1940s affected the neutral or slightly negative net-migration for metropolitan areas during the mid 1970s and mid 1980s. The "reurbanization" trend after the mid 1980s can be attributed partly to Tokyo's change into a "global city". The traditional cross-sectional approach to migration studies seems, however, to be losing its efficacy gradually because of the diversified behaviours of migrants. The author has therefore presented a longitudinal study of migrants and pointed out the increasing importance for the recent migration trend of non-economic factors, such as education and marriage.

KEY WORDS: internal migration, metropolitan area, non-metropolitan area, longitudinal approach, cohort, Japan.

INTRODUCTION

Migration and its consequences, changes in population distribution, at the national level have been one of the major focuses in urban and population geography for decades. Many important ideas such as urbanization, counterurbanization, reurbanization and migration turnarounds have been investigated and discussed in this research field.

The philosophy of the research has also been developed and reconsidered in the migration studies. From the perspective of neo-classical economics, employment opportunity and income differentials are the determinative factors of migration. People move out from regions of low income and high unemployment, and regional disparities are reduced as a result. Behaviourists have emphasized the importance of individual decision-making, which does not necessarily agree with the above-mentioned economic principle. Subjective "quality of life" also plays an important role in migration. Those interested in a life-course approach pay attention to changing individual residential preferences in accordance with the

aging process, which causes migration. They, therefore, investigate relationships between migration and important life events, such as leaving the parental home, marriage, childbearing and retirement. On the other hand, structurists emphasize constraints, instead of the free decision-making of individuals. According to them, mobility/immobility is the consequence of a limited housing market, a government policy or a failure of capitalism. We should realize, however, that these approaches are not always exclusive, but may be complementary.

Within the quantitative movement of the 1960s and the behavioural and structural approaches of the 1970s, urbanization and migration studies were one of the leading research fields affecting the development of methodology in human geography. Since the 1980s, however, the humanistic approach and more recent approaches in geography have not influenced this field so much (Boyle et al. 1998).

On the other hand, some sophisticated statistical methods have been introduced in order to grasp the determinants of migration turnarounds in recent decades. The term "migration turnaround" indicates a shift in the major stream of internal migration, either from periphery-to-core migration (urbanization) to core-to-periphery migration (counterurbanization) or from core-to-periphery migration (counterurbanization) to the periphery-to-core migration (reurbanization).

Moreover, researchers in the field tend to use not only cross-sectional migration data from official statistics, but also some kinds of longitudinal data. For example, the British statistical office provides one per cent longitudinal census data, from which we can obtain precise information concerning the same person in the 1971 and 1981 censuses as well as the 1981 and 1991 censuses.[1] There are some other affordable panel-data packages showing changes in location of residence and other indicators of the same persons during the given period. Furthermore, some researchers are conducting original questionnaire surveys concerning retrospective migration experiences in individual life courses. Cross-sectional analysis tends to relate changes in migration to changes in the economic and social situation that occurred simultaneously, while, from the longitudinal approach, the changes in migration are regarded as the accumulation of the past behaviours of individuals. Their past behaviours are then investigated in connection with past changes in the economic and social situations in which they experienced life events, such as job-finding and marriage. The longitudinal and cohort viewpoints have gradually become of major interest both at an aggregate and a non-aggregate level in migration studies.[2]

The purpose of this paper is to examine past Japanese migration trends mainly from the cross-sectional approach, and to present a preliminary study from the longitudinal perspective.

[1] The major outputs of the longitudinal analysis of the 1971 and the 1981 Censuses by British geographers were published in two volumes (Champion T. and Fielding T. 1992 and Stillwell J. et al. 1992).

[2] There have been controversial discussions between advocates of the cross-sectional and longitudinal approaches in migration studies (Davies R. B. and Pickles A. R. 1985; Pickles A. R. and Davies R. B. 1991; Clark W. A. V. 1992; Dieleman F. M. 1992).

MIGRATION AND MIGRATION STUDIES IN JAPAN

On the one hand, migration studies in Japan have obviously been influenced by the changing research perspectives in the English-speaking countries. For example, Ishikawa (1992; 1998; 1999) has contributed vigorously to international journals about migration turnarounds in Japan from the comparative perspective, using newly developed quantitative methods. On the other hand, Japanese researchers on migration have also developed a cohort perspective domestically in their studies to interpret the migration turnaround of the 1970s (Kawabe 1983; Itoh 1984).

The rest of this chapter describes some characteristics of the internal migration of Japan, which led to the spontaneous development of the longitudinal approach in Japanese migration studies. As shown in Figure 1, Japan consists of 47 prefectures (*Ken*), which correspond to the highest-level local administration unit. In this paper we will focus on the migration between metropolitan areas (as a core region) and the non-metropolitan areas (as a periphery region). There are three metropolitan areas, namely the Tokyo Metropolitan Area (MA), Nagoya MA and Osaka MA. Each metropolitan area includes three to six prefectures (Fig. 1). The number of prefectures included in the metropolitan areas is still under discussion because of the continuous suburbanization process. Here we regard 13 (4+3+6) prefectures as the total metropolitan area, while some geographers adopt only 10. As tested by Ishikawa (1994), the differences in observed migration

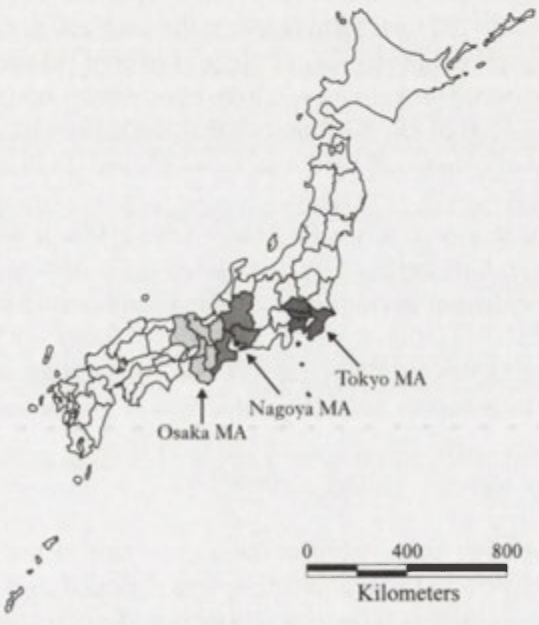


Figure 1. Three Metropolitan Areas (MAs) of Japan.

Notes: Japan consists of 47 prefectures (provinces). Tokyo Metropolitan Area (MA) includes four prefectures (Tokyo, Saitama, Chiba and Kanagawa) and Nagoya MA consists of Aichi, Gifu and Mie prefectures. Osaka MA includes Osaka, Shiga, Kyoto, Hyogo, Nara and Wakayama prefectures.

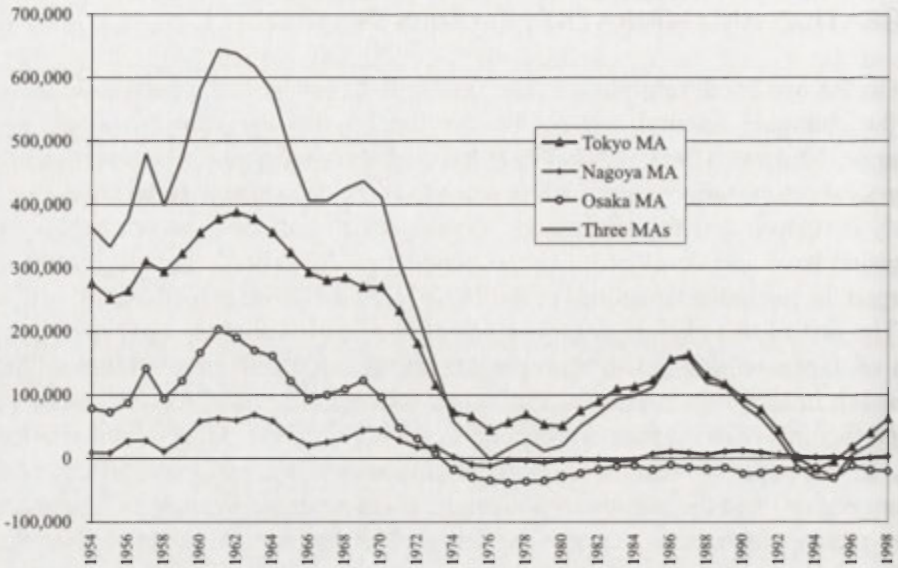


Figure 2. Net Migration for Tokyo, Nagoya and Osaka Metropolitan Areas (1954–1998).
Note: Figures for the years from 1954 to 1972 do not include migrants to or from Okinawa.
Source: Annual report on the internal migration in Japan derived from tha basic resident registers 1998.

trends arising from the various definitions of metropolitan prefectures are negligible if we discuss internal migration between the core and periphery regions.

Figure 2 shows net migration for the three metropolitan areas in the last 45 years. Migrants are those persons who crossed prefectural boundaries in changing their addresses. First of all, it is obvious that the metropolitan areas have net inflow, as against the non-metropolitan areas almost throughout the period shown. There are, however, substantial changes in the migration trend, and in particular a shift of the trend in the first half of the 1970s is apparent. The first period is characterized by an extreme amount of net inflow to the three metropolitan areas, while the net in-migration in the second period was very modest, and there was net out-migration from the areas in 1976 and 1993–1995. The sum of net in-migrants in the first 20 years (1954–1973) amounts to almost 8.7 million, while the total for the next 25 years (1974–1998) is only 1.3 million.

THE FIRST PERIOD: BEFORE THE MID 1970s

Japan experienced its baby boom in the second half of the 1940s, just after the Second World War, and overpopulation was regarded as a serious problem by the government and social scientists at that time. They focused, in particular, on the tremendous "surplus" population in rural areas, and established a study group named "*Noson Jinko Mondai Kenkyukai* (Study group on population problems in the rural area)." Even though the total fertility rate dropped significantly

in the 1950s, from around 5 to 2, they still considered overpopulation to be an urgent issue. They feared that the baby-boom cohort would be added suddenly to the working population in the 1960s.

Then, in the late 1950s post-war Japan's drastic economic growth started. Industrialization in the core region attracted many young people from agricultural areas, and the existence of a sizeable young workforce fortunately supported the economic growth. The interests of the researchers changed in accordance with the rapid industrialization and urbanization. They worried about and examined the depopulation process in rural areas, and pointed to the diseconomy of overurbanization, such as in air pollution and high land prices. There were obviously large disparities in employment opportunity and income between the rural and urban areas in the 1960s, and researchers thus could not find any symptom of a migration turnaround.

THE SECOND PERIOD, PART 1: MID AND LATE 1970S AND EARLY 1980s

Kuroda reported a sign of a migration turnaround at the annual conference of the Japanese Sociological Association in 1970 (Okada 1973). As shown in Figure 2, net in-migration for the three metropolitan areas was still observable in the first half of the 1970s, and the existence (or possibility) of a migration turnaround was disputed among Japanese sociologists, geographers and economists.

After the first experience of net outflow from the three metropolitan areas in 1976, the migration turnaround was broadly recognized, but a large amount of net outflow was never observed. The discussion in the 1980s concentrated on seeking determinants of the shift in the major migration flow. Economists proved the causal relationship between the changing migration patterns and economic indicators, such as income differentials and job opportunities, using mainly cross-sectional analysis. However, some demographers and population geographers emphasized the changing size of the population of young adults (Kawabe 1983; Itoh 1984).

Kawabe (1983) adopted a cohort perspective and investigated the numbers of net-migrants for the metropolitan areas by birth-cohort. He concluded that the migration patterns of each cohort seemed relatively constant, and that the observed turnaround was misleading, arising from changing cohort size. Two points characterizing the post-war migration in Japan will be introduced briefly here, so that the backgrounds of the cohort perspective adopted by Kawabe and Itoh might be understood better.

First, Figure 3 shows the age profile of the cohort-changing rate[3] for the

[3] The cohort-changing rate in Figure 3 was calculated as follows; for example, the male population of the Tokyo MA aged 15–19 in 1990 is 1,413,184 and the male population aged 20–24 in 1995, which corresponds to the same cohort five years later, amounts to 1,623,722. The cohort-changing rate of ages 15–19 at the start of the period 1990–1995 is 1.15, the quotient of 1,623,722 divided by 1,413,184. The cohort survival rate of those aged less than 50 is nearly 1.00, and thus, roughly speaking, if a cohort-changing rate is more than 1.00, that cohort shows positive net-migration for the Tokyo MA in the relevant period.

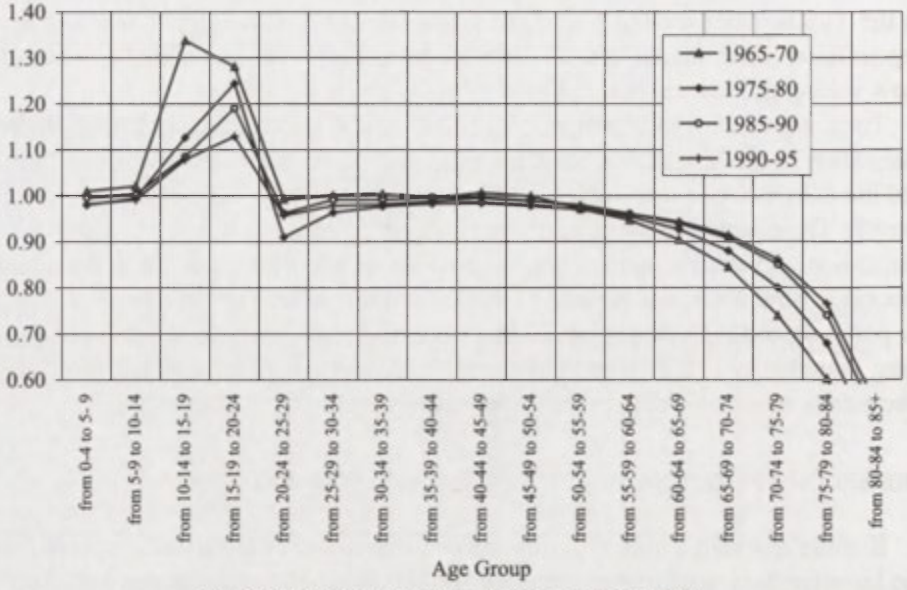


Figure 3. Cohort-changing rate for the Tokyo MA.
Source: Population Census 1965, 1970, 1975, 1980, 1985, 1990 and 1995.

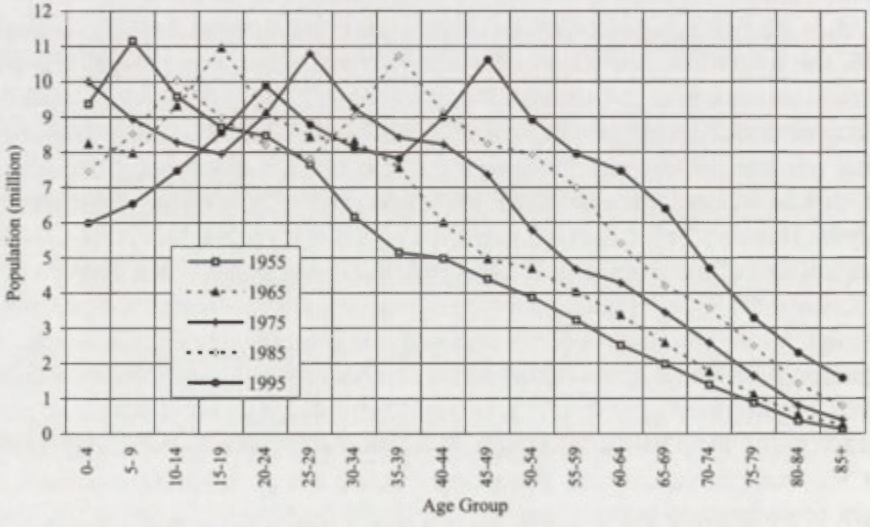


Figure 4. Population size by age group in Japan (1955–1995).
Source: Population Census, 1955 1965, 1975, 1985 and 1995.

Tokyo metropolitan area (MA), which is approximately equivalent to the net migration rate by age group for those aged less than 50. As Kawabe described, the calculated age profiles for the four periods look to a large extent similar. The highest values for the cohort-changing rate are found at ages 15–19 to 20–24,

except for the period 1965–1970.[4] According to Figure 3, it is young people aged around 20 who migrate to the Tokyo MA. And only a few of them move out from it after they become 20–24 or 25–29, a phenomenon which would correspond to a return migration to the non-metropolitan area. Otherwise they remain in the Tokyo MA. In other words, the changing population size of metropolitan areas depends almost completely on the net inflow of those aged around 20 and the net outflow of those in their mid and later 20s.

Second, the size of the baby boom cohort is obviously larger than that of the preceding and following cohorts (Fig. 4). In 1955, the size of the population aged 5–9 born in the period 1946–1950 is of about 11 million, while those aged 0–4 and 10–14 amount to only 9.5 million. As it ebbs and flows, the cohort born in 1946–1950 has affected the distinct age structure of Japan. Those aged 15–19 in 1965, 25–29 in 1975, 35–39 in 1985 and 45–49 in 1995 have exceeded the preceding and following cohorts in number by approximately 15 per cent, a phenomenon which has required new elementary schools, introduced tougher competition for entrance examinations and raised house prices in the relevant periods. After reproductive age was reached, the second generation of the baby boom cohort was born in the early 1970s, that a volume slightly smaller than that of their parents (Fig. 4).

Taking these two points into account, Kawabe attributed the migration equilibrium in the mid and late 1970s to a larger volume of out-migrants equivalent to the baby boom cohort, and a smaller volume of in-migrants in the following baby-bust cohort.

THE SECOND PERIOD, PART 2: AFTER THE MID 1980s

From the mid 1980s onwards, some new trends were observable in both migration flows and migration studies. Ishikawa led the migration research in the late 1980s and 1990s. He examined the determinants of the migration turnaround in the 1970s, adopting the newly-developed quantitative methods, and identifying as the main factor of the turnaround a dispersion of industrialization, in particular of blue-collar manufacturing jobs, toward the non-metropolitan region. On the other hand, he also stressed the effects of the changing size of cohorts for the migration turnaround in the broader internationally-comparable context, and re-evaluated the works of Kawabe and Itoh, which originated from the domestic background. It is widely accepted in recent studies that it is the economic factor and the cohort-size factor together which brought about the migration turnaround of the 1970s (Ishikawa 1992, 1999; Oe 1995). Oe tried to integrate and evaluate the effects of the two factors from a cohort perspective while Ishikawa applied shift-share analysis. They recognized that the economic component effect was

[4] Rising participation ratio in higher education explains this alternation during 1965–1970. After 1975, however, the ratio is almost constant.

stronger in the late 1960s and early 1970s, but it lost its effectiveness in the mid and late 1970s when the turnaround was affected by cohort size instead.

The metropolitan-oriented migration flows of the 1980s (Fig. 2) called for new interpretations. The net inflow increased again toward a peak in 1987 and then declined; net outflow was recorded again in 1993. Ishikawa (1999), considering the similar tendencies in other developed countries in the 1980s, regarded this as a new turnaround, from non-metropolitan to metropolitan-oriented. The net inflow in the 1980s is characterized by the dominance of the Tokyo MA, and the other two MAs remain balanced or with slightly negative net migration (Fig. 2). During the 1960s, all three MAs enjoyed affluent net in-migration. Neither the relative differential in employment and income between the metropolitan and non-metropolitan areas, nor the changing cohort size, can explain this turnaround satisfactorily. Researchers therefore borrowed a new analytical framework, such as the "Global/World city" (Sassen 1991), for interpretation. Based on the free movement of financial capital, Tokyo became one of the most important financial centers, together with New York and London, in the second half of the 1980s. Many new jobs were created thereby, particularly in the tertiary sector, independent of the economic situation in other regions of Japan, and Tokyo enjoyed the so-called "bubble economy" for a few years. Since the shrinking of the bubble economy at the beginning of the 1990s, Japan has been in deep economic stagnation. Moreover, the newest trend in Figure 2 only shows an increasing net inflow for the Tokyo MA, but it still remains economically stagnant.

Ishikawa and Fielding (1998) tried to determine the cause of the migration trends for the Tokyo MA in the 1980s and 1990s, and identified "world city" as a main factor in a rather eliminatory way. However, this problem would still seem to be inside the maze, rather than getting solved, and the limitations of the traditional approaches are thus being revealed.

THE LONGITUDINAL APPROACH: FROM THE 1996 MIGRATION SURVEY

As mentioned in the first chapter, more and more geographers have become interested in conducting original surveys, and in collecting non-aggregate migration histories of individuals along their life courses, with which they propose to supplement and overcome the limitations of the aggregate cross-sectional perspective. Together with other geographers and sociologists, the author conducted a retrospective questionnaire survey on migration in 1996 (Nishioka et al. 1997). Some of the results yielded by the survey will be introduced in the rest of this paper.

We selected 300 districts covering the whole of Japan using a random sampling method, and distributed a questionnaire to all the households in each district,

i.e. to a total of 15,131 households. We succeeded in collecting valid responses from 14,083 households, thereby obtaining retrospective migration histories of 40,400 household members. The available information from an individual migration history is as follows; place of birth, place of residence on finishing junior high school, place of residence at graduation from last school, place of residence on obtaining first job, place of residence just before marriage and just after marriage, place of residence five years before the survey, and place of present residence.

THE MIGRATION PROCESS OF NATIVE NON-METROPOLITANS AT MAJOR LIFE EVENTS

We will examine the process of migration for those born in the non-metropolitan area. In general, a considerable proportion of the native non-metropolitans moved to the metropolitan area in the course of their life. Their migration process is summarized in the ratios for metropolitan residence at each major life event. Figure 5 illustrates the migration process of the native non-metropolitans by sex and cohort. If all the lines move on the same track, we can conclude that

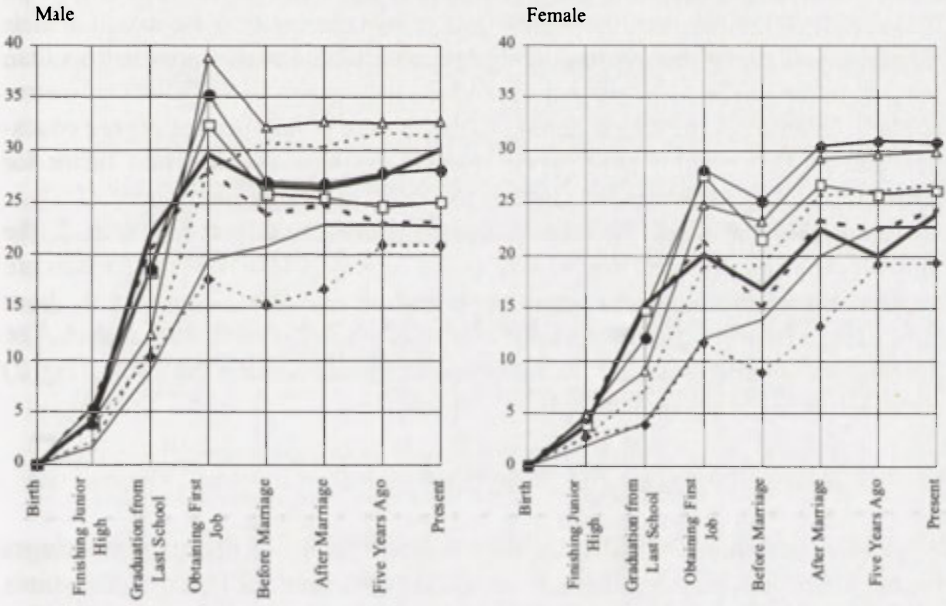


Figure 5. Ratios of metropolitan residents at life events by birth cohort (%). (Married persons born in the Non-Metropolitan Area < N = 14,268 >: 1996 Migration Survey).

Average age at each life event:
Obtaining first job: 19.0 (male), 18,7 (female)
First marriage; 26.9 (male), 23.9 (female)

Notes: Metropolitan Area is the sum of Tokyo MA, Nagoya MA and Osaka MA (Figure 1), and Non-Metropolitan Areas consist of those prefectures not included in the Metropolitan Area.

the distribution of the native non-metropolitans depends only on their age. However, each line traces differently on the graph, an observation which seems to be derived from the different migration behaviours by cohort.

The ratios of metropolitan residence for both sexes show incremental changes toward the point at the first job. Examined in detail, the cohorts born after 1950 trace steeper lines in the period between junior high school and the last school, while the cohorts born before 1945 tend to shift their distribution obviously between graduation and the first job. The rise of the ratios between junior high school and the last school seems to be caused by increasing university participation. The upward changes in the ratio of metropolitan residence from birth to the first job are more remarkable for males.

It is notable that the ratios decline between the first job and the time "before marriage" except for the 1931–1941 cohorts.[5] This downward change seems to be related to return migration to the place of birth or neighboring cities in the non-metropolitan area. When the distribution changes are observed by life event, so-called return migration is found between the first job and marriage. The return migration is more apparent for males as well.

The percentages of metropolitan residents "before marriage" remain without any substantial change for the native non-metropolitan males, while the graph illustrates that the ratios of the female cohort also change to some extent at their marriage, indicating that a considerable number of the native non-metropolitan women move to the metropolitan area when they marry. The factors in metropolitan-oriented net migration among males are the obtainment of higher education and of first employment, while marriage is also an important factor for females. The ratios are almost constant for both sexes after marriage.

Comparing the ratios for metropolitan residents by cohort in Figure 5, the ratios for the cohorts born 1936–1946 for males and 1941–1951[6] for females are highest and reach around 30 per cent, indicating that about one-third of those born in the non-metropolitan area are now residing in the metropolitan area. The tremendous volume of net-inflow into the metropolitan area in the 1960s (Fig. 2) corresponds to the movement of these cohorts.

ECONOMIC AND NON-ECONOMIC FACTORS

In this last section we will discuss the determinants of distribution changes for the native non-metropolitans. Based on the data used for Figure 5, the points

[5] This cohort spent their twenties in the economic boom of the 1960s, when the economic differentials between the metropolitan and non-metropolitan areas widened. They appeared to move to the metropolitan areas not only for a first job, but also after working in villages or hometowns for a while.

[6] The difference seems to be attributable to the age difference at marriage between males and females.

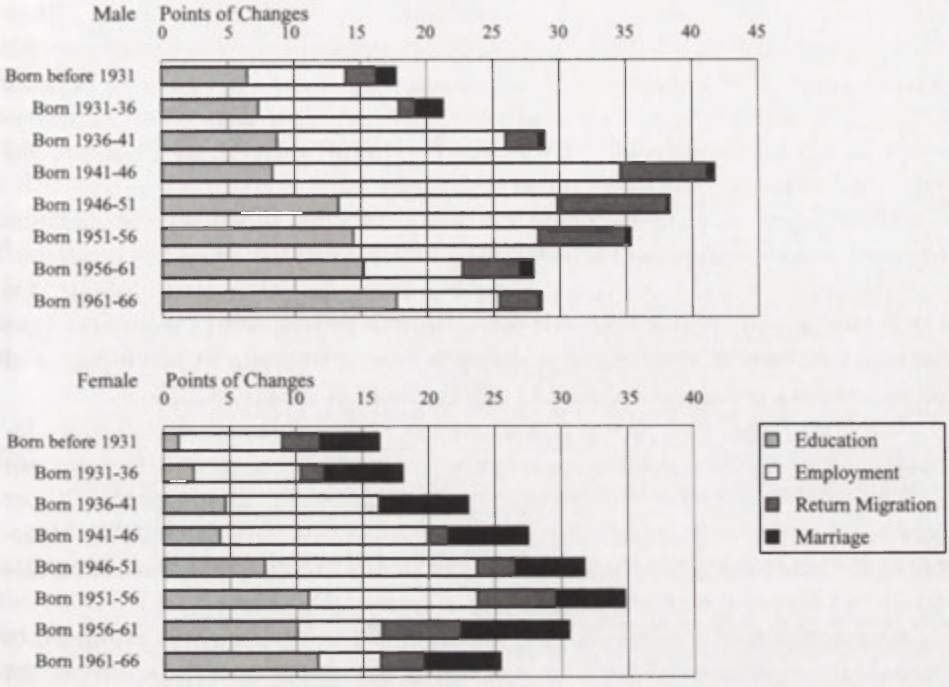


Figure 6. Changing ratios of metropolitan residents between life events from Figure 5. (Married persons born in the Non-Metropolitan Area: 1996 Migration Survey).

of changes in the ratios between life events are calculated and presented in Figure 6. For example, the ratios for metropolitan residents of the 1961–1966 male cohort at the times of finishing junior high, graduating from the last school, obtaining the first job and preparing for marriage are 4.1%, 21.8%, 29.6% and 26.6% respectively. The change for "Education" is thus equivalent to the difference between 4.1% and 21.8%, namely 17.7 points. The difference in the ratios between graduation and first job, 7.8 points here, is regarded as the change for "Employment". The change in the ratios between first job and the time before marriage is assumed here to be a "Return Migration" component (–3.0 points). However because the changes during this period show negative values for most of the cohorts, it is the absolute values that are adopted in Figure 6.

The total points for changes among males are rather greater than for females. For males, the largest value is found for the 1941–1946 cohort and the values diminish gradually for the younger cohorts. In detail, the "Education" component increases both in absolute value and in the share against the total points for the younger cohorts. It seems that younger cohorts born in the non-metropolitan area tend to move to the metropolitan area for higher education. On the other hand, the "Employment" component shows its maximum value for the 1941–1946 cohort (males) and the 1946–1951 cohort (females), being equivalent to that in

the 1960s, and the importance of this component has declined since then. These changes appear to be attributable to the following facts; 1) many young people already reside in the metropolitan area when they find their first jobs, because a considerable proportion study at universities there; 2) the dispersion of employment to the non-metropolitan area after the 1970s reduced the necessity for migration to the metropolitan area in search of a job.

The "Return Migration" component was movement in the reverse direction toward the non-metropolitan area, except in the case of the 1931–1941 cohorts.[7] Considerable "return migration" is observed for the 1946–1956 cohorts. Oe (1995) has already pointed out that the economic stagnation of the metropolitan region in the late 1970s stimulated young job seekers to return to their places of birth, a phenomenon which created large numbers of return migrants.

The "Marriage" component reflects moves at marriage and characterizes female migration. This point is quite new for us, because such data were not available before. The share of this component amounts to approximately 20 per cent for every female cohort. For the younger cohorts born after 1956, "Marriage" becomes the second influential factor in determining their residential distribution between the metropolitan and non-metropolitan areas.

As mentioned in the last chapter, the economic component was estimated to be the major determinant for changing net-migration between the metropolitan and non-metropolitan areas, and many researchers examined the power of this factor against the net-migration change. However, some of the recent studies argue for the declining importance of the economic factor and struggle to find a new interpretation. If we regard "Employment" and "Return Migration" together as the economic factor, and "Education" and "Marriage" as the non-

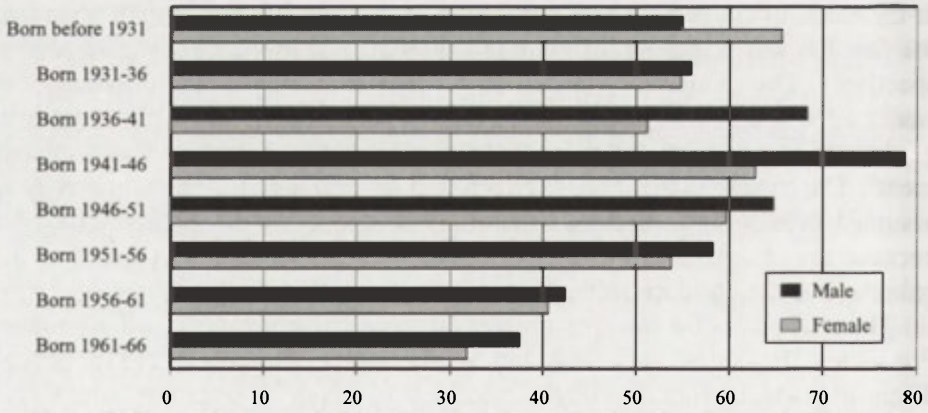


Figure 7. Share of "economic factor" against the total points for changes in Figure 6 (Married persons born in the Non-Metropolitan Area: 1996 Migration Survey).

Note: "Education" and "Marriage" are regarded as a non-economic factor, while "Employment" and "Return Migration" are an economic factor.

[7] See footnote 5.

-economic factor, the shares of the economic factor against the total points for changes are as illustrated in Figure 7. The declining share of the economic factor for younger cohorts, especially in recent years, is remarkable. Almost 80 per cent of the net distribution change was explained by the economic factor for the 1941–1946 male cohort, while only 37.5 per cent of the net migration of the 1961–1966 male cohort is attributed to the economic factor. Considering that it is mostly in its twenties that each cohort tends to change its distribution between the metropolitan and non-metropolitan areas, the economic factor had substantial determinative power for net migration before 1980. But "Education" and "Marriage (for females)" have increased their relative importance in changing residential distribution, while the economic factor has lost power since the 1980s.

CONCLUDING REMARKS

This paper has discussed the changing net migration between the metropolitan and non-metropolitan areas in post war Japan. The first part describes not only the changing migration patterns but also the alternation of perspectives. In order to compliment and overcome the limitations of the traditional aggregate cross-section approach, which are partly revealed in recent studies, a preliminary study from the non-aggregate longitudinal perspective is presented in the last chapter.

Recent studies from the cross-sectional approach come to distinguish the economic factor and the changing size of cohorts as determining changing net migration between the metropolitan and non-metropolitan areas. However, they assume that it is only these two factors that determine the net migration and migration turnarounds, and if their power is not sufficient, an external factor, such as the "World City Hypothesis" tends to be borrowed. However, it appears from this study that the traditional economic factor contains not only a pure economic factor, but also migration for education or migration for marriage, and that the importance of the latter has increased since and the 1980s, something which has reduced the power of the traditional economic factor thus increased and the difficulty of interpretation based upon it.

REFERENCES

Blotevogel H. H., Fielding A. J. (eds), 1997, *People, Jobs and Mobility in the New Europe*, John Wiley & Sons.

Boyle P., Halfacree K., Robinson V., 1998, *Exploring Contemporary Migration*, Longman.

Champion T., Fielding T., 1992, *Migration Process and Patterns*, Vol. 1, *Research Progress and Prospects*, Belhaven Press.

Clark W. A. V., 1992, *Comparing cross-sectional and longitudinal analysis of residential mobility and migration*, Environment and Planning A, 24, 1291–1302.

Davies R. B., Pickles A. R., 1985, *Longitudinal versus cross-sectional methods for behavioral research: a first-round knock out*, Environment and Planning A, 17, 1315–1329.

Dieleman F. M., 1992, *Struggling with longitudinal data and modeling in the analysis of residential mobility*, Environment and Planning A, 24, 1527–1530.

Ishikawa Y., 1992, *The 1970s migration turnaround in Japan revisited: a shift-share approach*, Papers in Regional Science, 71, 153–173.

Ishikawa Y., 1994, *Jinkoido no keiryo chirigaku (Quantitative Geography of Migration)*, Kokon shoin, Tokyo, (in Japanese).

Ishikawa Y., 1999, *Contribution of the demographic factor to the migration turnarounds in Japan, Sweden and Canada*, International Journal of Population Geography, 5, 1–17.

Ishikawa Y., Fielding A. J., 1998, *Explaining the recent migration changes of the Tokyo metropolitan area*, Environment and Planning A, 30, 1797–1814.

Itoh T., 1984, *Recent trends of internal migration in Japan and potential life time outmigrants*, Jinko Mondai Kenkyu (Journal of Population Problems), 172, 24–38 (in Japanese).

Kawabe H., 1983, *The Population Distribution and the Internal Migration since 1960*, Jinkogaku Kenkyu (Journal of Population Studies), 6, 7–14 (in Japanese).

Nishioka H., Nakagawa S., Kojima K., Shimizu M., Oe M., Wakabayashi K., Inoue T., 1997, *Recent trends of migration in Japan: Migration Survey 1996*, Jinko Mondai Kenkyu (Journal of Population Problems), 53-3, 1–30 (in Japanese).

Oe M., 1995, *Cohort analysis of population distribution change in Japan: processes of population concentration to the Tokyo region and its future*, Jinko Mondai Kenkyu (Journal of Population Problems), 51-3, 1–19 (in Japanese).

Okada M., 1973, *A debate on the existence of a migration turnaround*, Chirigku Hyoron (Geographical Review of Japan), 46, 656–665 (in Japanese).

Pickles A. R., Davies R. B., 1991, *The empirical analysis of housing careers: a review and a general statistical modeling framework*, Environment and Planning A, 23, 465–484.

Sassen S., 1991, *The Global City: New York, London, Tokyo*, Princeton University Press.

Stillwell J., Rees P., Boden P., 1992, *Migration Process and Patterns*, Vol. 2, *Population Redistribution in the United Kingdom*, Belhaven Press.

GEOGRAPHIA POLONICA, 73, 1, Spring 2000
PL ISSN 0016-7282

THE APPLICATION OF FOURIER EXPANSION IN ESTIMATING THE ANNUAL VARIATION TO THE VERTICAL GRADIENT OF SOME METEOROLOGICAL PARAMETERS

JANUSZ OLEJNIK, JACEK LESNY
Agrometeorology Department of The Agricultural University in Poznan,
ul. Witosa 45, 60-647 Poznan, Poland
e-mail: olejnikj@owl.au.poznan.pl

FRANK EULENSTEIN
Zentrum fur Agrarlandschafts- und Landnutzungsforschung,
Eberswalder Str. 84, 15374 Muncheberg, Germany

ABSTRACT: Data from 18 meteorological stations in the southern part of Germany were analysed. Long-term average data (1958–1987) were compared with altitude data for meteorological stations in the Main River catchment, which varied from 150 to 920 m above sea level. Linear regression analysis against altitude was carried out for each month for six meteorological parameters: air temperature, precipitation, vapour pressure, saturation vapour pressure deficit, wind speed and relative humidity. The regression shows a very significant correlation for the analysed meteorological parameters. The annual variation to the vertical gradients for these parameters is also shown. To estimate gradient variation a Fourier expansion of the annual cycle is applied (given by monthly values of the gradients) and terms up to the first harmonic retained. Using the proposed functions for vertical gradient estimation, the areal distributions of temperature are calculated and average values for those parameters within the Main River catchment shown at monthly and yearly resolution. The areal distribution to mean annual temperature was determined and presented as a grid map.

KEY WORDS: Fourier expansion, vertical gradient, annual variation.

INTRODUCTION

Many recent publications have described the problem of long-term climate change. Investigators have examined long-term variations in temperature and precipitation (Jones et al. 1986; Angell 1988; Hansen and Lebedeff 1988; Diaz et al. 1989; Jones et al. 1990) with increasing interest in the potential impacts of anthropogenic modification of climate, e.g., increases in the concentrations of greenhouse gases (Hulme et al. 1990; Fabian 1991; Schonwiese and Malcher 1987). One important element of this problem is the investigation of the environmental impacts of climate change on both global and local scales. The evidence for future climate change comes mainly from general circulation models (GCM) based on grids of 5 to 2 latitude by 5 to 2 degrees of longitude (Karl and Wil-

liams 1987; Verstraete 1989), and the impact of climate change is usually analysed at a similar resolution (Olejnik 1988; Solomon and Leemans 1990; Carter et al. 1990; Fabian 1991). To detect the regional impact of climate change the local station network should be taken into account, as should local (e.g. river-catchment) information on land-use conditions, the distribution of agricultural production, phenology etc. Such work should be based on a resolution smaller by at least two orders of magnitude than that of a GCM.

This paper deals with the problem of the spatial averaging of absolute meteorological quantities from stations at different altitude on a local scale (about 200 km by 150 km). The regional, annual and long-term spatial averaging of some meteorological parameters are discussed in what is an initial part of a project concerning to local heat-balance changes due to climate change (Olejnik 1996). A dataset from the Main River catchment (8.2°–11.8° longitude, 48.9°–50.6° latitude) has been examined in relation to the period 1958–1987. Because the meteorological stations involved are situated at significantly-different altitudes, it was necessary to estimate the vertical gradients to different meteorological parameters, before any spatial averaging procedure could be used. The meteorological parameters were assumed to depend linearly on altitude above sea level and the vertical gradients were computed for each month separately, to reveal a strong annual cycle which may be well-represented by the first harmonic of a Fourier expansion. To achieve spatial averaging of meteorological parameters it was proposed that data for all stations be extrapolated vertically to a constant level (sea level), and then interpolated to a regular grid on that level before the value at each grid point was extrapolated to the altitude of that point.

METHODS

The whole river catchment was covered by a grid net at a resolution of 0.1 latitude by 0.1 longitude (thereby giving $40 \times 20 = 800$ grids – longitude $i = 40$ and latitude $j = 20$). The interpolation (and extrapolation) procedure was based on meteorological data from a local network including 18 stations. In this paper the interpolation procedure is described on the basis of temperature, but the same procedure was used for other meteorological parameters (precipitation, humidity etc.). Because of large differences in the altitude of the different stations (from 150 to 920 m above sea level) the interpolation procedure could only be applied after recalculation of temperature data to sea level. Subsequently, to obtain the real value of temperature in all grids the vertical changes to temperature as a function of height above sea level were taken into consideration (Fig. 1). Temperatures in grids G1 and G5 at sea level (T_1' and T_5', Fig. 1) were calculated as follows:

$$T_1^{'} = T_A - \mathrm{grad}T\, h_A\,, \qquad (1)$$

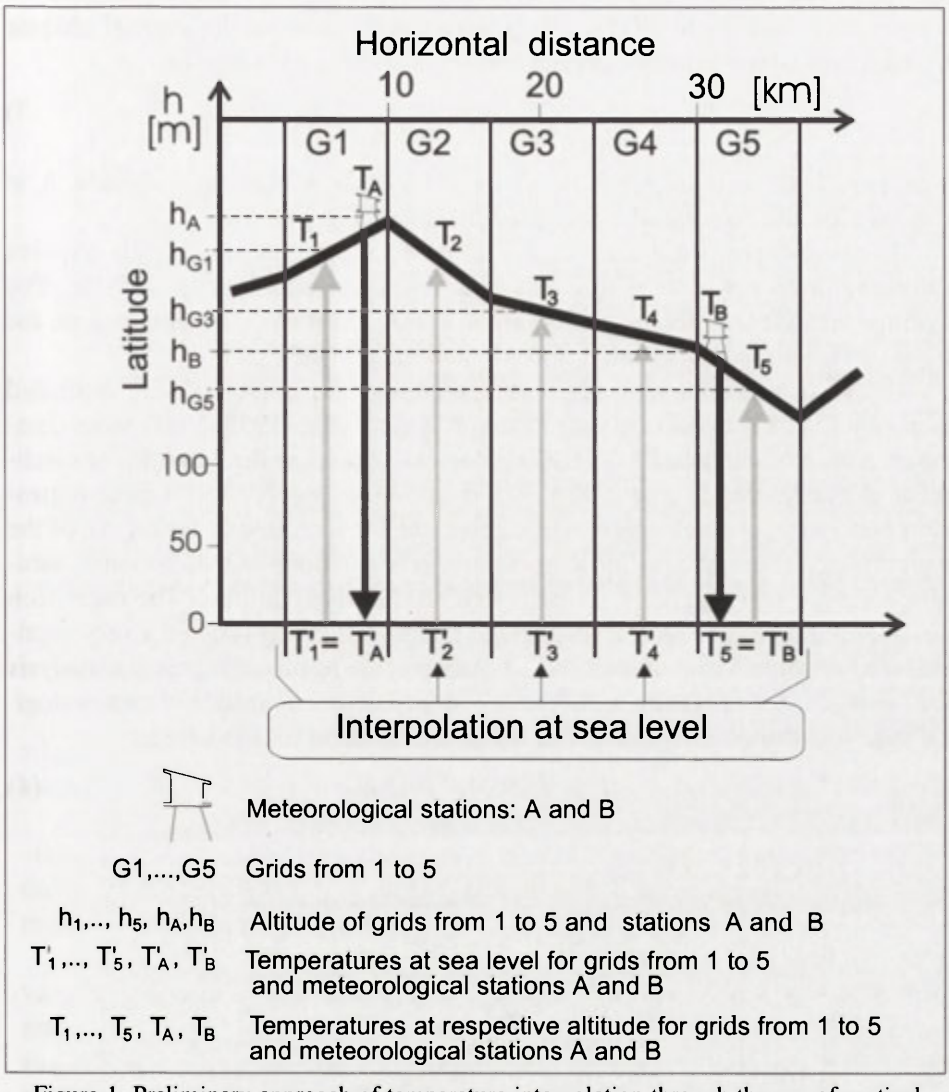


Figure 1. Preliminary approach of temperature interpolation through the use of vertical gradient of temperature.

$$T_5' = T_B - \text{grad}\, T\, h_B \tag{2}$$

where: T_1' and T_5' are the average temperatures at sea level in grids G1 and G5 respectively, T_A and T_B are temperatures measured at stations A and B located within grids G1 and G5 respectively, grad$T$ is the vertical gradient of temperature within the Main River catchment and h_A, h_B are altitudes of meteorological stations A and B respectively.

The interpolation procedure was then used to estimate the temperatures: T_2', T_3', T_4 within grids C2, C3, and C4, but still at sea level (Fig. 1). Afterwards, the

temperatures in all grids (G1, ..., G5) were recalculated for the average altitude of each grid (Fig. 1) using equation (3):

$$T_n = T_n' + \text{grad}T\, h_n\,, \tag{3}$$

where: T_n is the average temperature in grid n at the average grid altitude, h_n is the average altitude of grid n (n varied from 1 to 5 in Fig. 1).

As can be seen, estimation of the average temperature in all grids requires information on the vertical gradient of temperature, and average altitude. The average altitude data for the whole Main River catchment were obtained on the basis of two altitude maps at: 1:300 000 and 1:2 500 000.

The vertical gradients of different meteorological parameters were estimated as follows. The long-term average meteorological data (1958–1987) were compared with altitude data for 18 meteorological stations in the Main River catchment. Altitudes varied from 150 to 920 m above sea level (a.s.l.). Linear regression analysis against elevation was carried out for each month for all six of the meteorological parameters air temperature, precipitation, vapour pressure, saturation vapour pressure deficit, wind speed and relative humidity. The regression analysis was based on long term average monthly data and showed a very significant correlation for all months. As an example, the results of regression analysis for average air temperature in July (T_{July}) as a function of altitude of meteorological station is shown in Figure 2. The regression equation for this case is:

$$T_{\text{July}} = 19.34 - 0.63\, h, \tag{4}$$

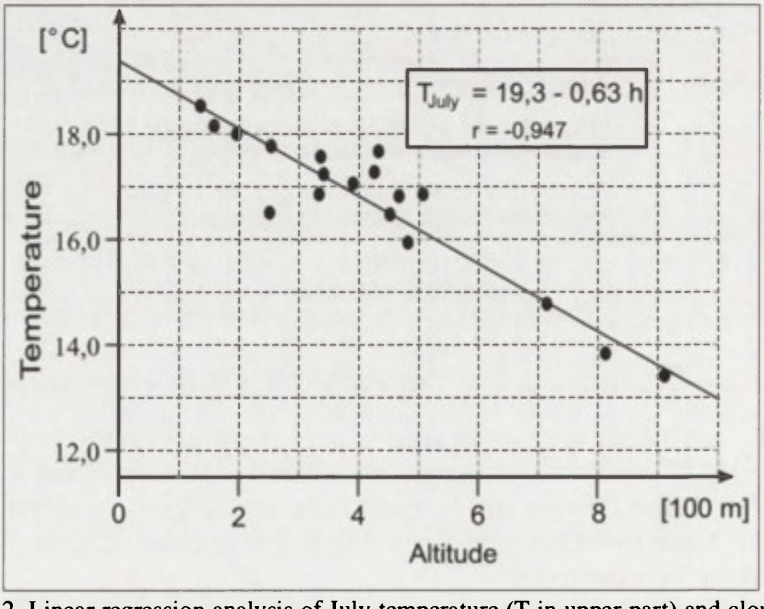


Figure 2. Linear regression analysis of July temperature (T in upper part) and cloudiness (C in lower part) as a function of altitude of meteorological stations.

Table 1. Average monthly values for the vertical gradients of six meteorological parameters. Values obtained on the basis of linear regression analysis of standard meteorological measurements in comparison with the altitude of meteorological stations: r is the correlation coefficient.

Number of months	Air temperature °C /100 m		Precipitation mm/100 m		Vapour pressure hPa /100 m		Saturation deficit HPa /100 m		Relative humidity %/100 m		Wind speed m/s /100 m	
	grad	r	grad	r	grad	r	grad	r	grad	r	grad	r
1	–0.49	–0.89	4.8	–0.77	–0.12	–0.78	–0.11	–0.91	1.2	–0.88	0.19	0.77
2	–0.55	–0.92	2.9	–0.77	–0.12	–0.84	–0.14	–0.93	1.3	–0.92	0.19	0.77
3	–0.65	–0.96	3.4	–0.66	–0.13	–0.89	–0.23	–0.94	1.6	–0.91	0.18	0.79
4	–0.67	–0.95	4.5	–0.88	–0.19	–0.93	–0.31	–0.91	1.4	–0.89	0.17	0.77
5	–0.63	–0.96	4.9	–0.92	–0.24	–0.90	–0.37	–0.91	1.2	–0.84	0.17	0.76
6	–0.65	–0.95	5.4	–0.91	–0.29	–0.89	–0.47	–0.92	1.3	–0.86	0.15	0.79
7	–0.63	–0.95	5.7	–0.95	–0.30	–0.87	–0.49	–0.92	1.2	–0.88	0.15	0.77
8	–0.56	–0.93	5.4	–0.94	–0.29	–0.88	–0.39	–0.90	1.0	–0.85	0.16	0.79
9	–0.49	–0.92	4.3	–0.92	–0.24	–0.85	–0.27	–0.89	0.9	–0.82	0.18	0.76
10	–0.43	–0.94	4.5	–0.76	–0.20	–0.85	–0.16	–0.87	0.7	–0.73	0.21	0.81
11	–0.51	–0.97	3.7	–0.65	–0.16	–0.90	–0.15	–0.93	1.1	–0.87	0.20	0.77
12	–0.51	–0.93	4.5	–0.63	–0.14	–0.87	–0.13	–0.91	1.2	–0.87	0.20	0.79
Annual	–0.57		54.1		–0.21		–0.27		1.2		0.18	

where: T_{July} is the average temperature in July measured at the meteorological station in [°C] and h the altitude of the meteorological station in [100 m].

The vertical gradient to temperature estimated in such a way is equal to –0.63°C/100 m and has a very high correlation of about 0.95. The same procedure was used for all months (from 1 – January to 12 – December) and for all six meteorological parameters. The obtained results are shown in Table 1.

On the basis of the estimated average monthly vertical gradients, the annual variation in them as shown in Figure 3 was determined. As can be seen, the arrangement of the variability of the vertical gradient during the whole year suggests the harmonic distribution in all six cases. Use of the least squares method thus allowed the distribution in time (month by month) of vertical gradients of these meteorological parameters to be found. The first harmonic of the Fourier expansion was applied as follows:

$$\mathrm{grad}T_{(100\ \mathrm{m})} = \mathrm{a} + \mathrm{b}\cos[\pi(\mathrm{x}-1)/6] + \mathrm{c}\sin[\pi\,(x-1)\,6], \tag{5}$$

where: x is the number of each successive month, $\mathrm{grad}T_{(100\ \mathrm{m})}$ the vertical gradient to temperature in [°C/100 m] and a, b, c constants.

The same procedure was used for all the analysed meteorological parameters, and the results are shown in Figure 3.

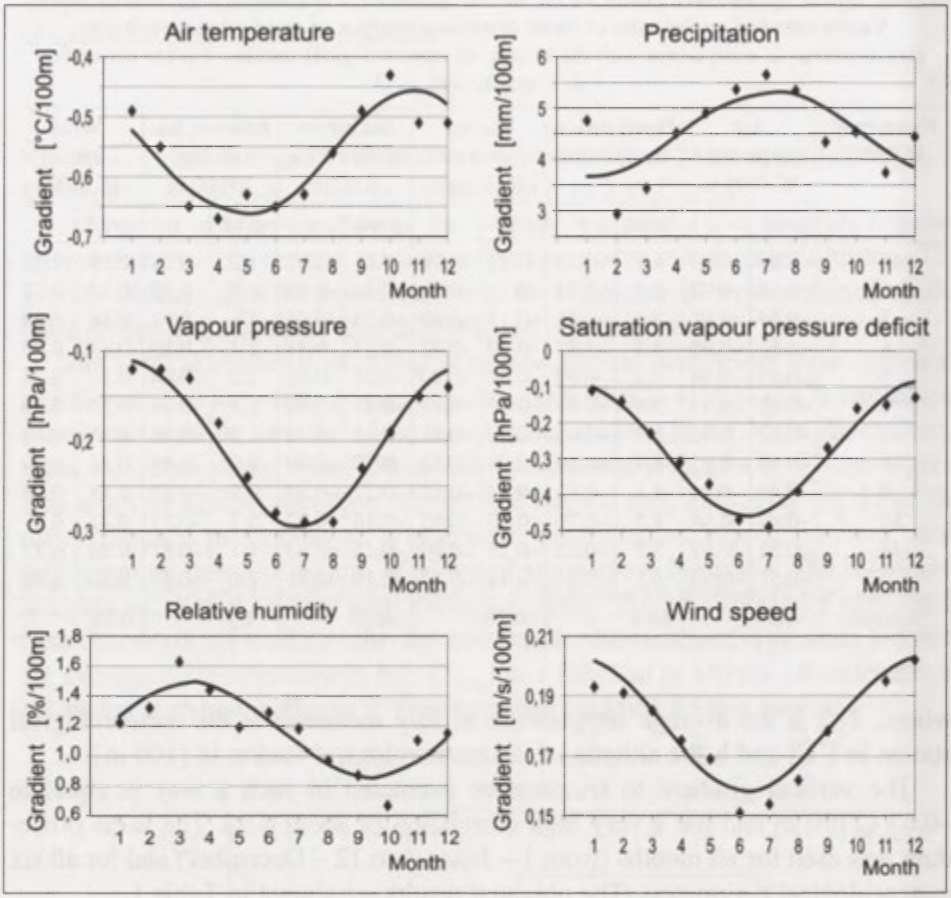


Figure 3. Monthly values for vertical gradients of six meteorological parameters (1 to 12 are numbers of months). Functions show Fourier expansion of estimated gradients on basis of measurements (dots).

RESULTS

Using the described method, equations for vertical gradient estimation, as a function of the number of months only, were obtained for six meteorological parameters. All results and first harmonic functions are shown in Figure 3. The constants (a, b, c in equation 5) for all meteorological parameters are combined in Table 2. Using these equations, vertical gradients were calculated for all six meteorological parameters and for all months. The results are shown in Table 3 in columns headed $grad_e$. Columns headed $grad_m$ detail the gradients estimated previously by linear regression analysis on the basis of measurement. The lower part of Table 3 gives the determination coefficients, again obtained by the least

Table 2. The combination of constants a, b, c (Eq. 5) in harmonic functions of vertical gradient estimation for six meteorological parameters (units as in Table 1), R^2 is determination coefficient (related to Fig. 3).

Meteorological Parameter	a	b	c	R^2
Air temperature	–0.564	0.043	–0.091	0.867
Precipitation	4.490	–0.812	–0.102	0.541
Vapour pressure	–0.202	0.092	–0.008	0.951
Saturation deficit	–0.269	0.169	–0.060	0.948
Relative humidity	1.169	0.092	0.293	0.789
Wind speed	0.181	0.021	–0.008	0.789

Table 3. Average monthly values of vertical gradients for six meteorological parameters: $grad_m$ – values obtained on the basis of linear regression analysis of standard meteorological measurements in comparison with the altitude of meteorological stations, $grad_e$ – values estimated by Fourier expansion (Eq. 5 and Table 2).

Number of months	Air temperature °C /100 m		Precipitation mm/100 m		Vapour pressure hPa /100 m		Saturation deficit HPa /100 m		Relative humidity %/100 m		Wind speed m/s /100 m	
	$grad_m$	$grad_e$	$grad_m$	$grad_e$	$grad_m$	$grad_e$	$grad_m$	$grad_e$	$grad_m$	$grad_e$	$grad_m$	$grad_e$
1	–0.49	–0.52	4.76	3.68	–0.12	–0.11	–0.11	–0.10	1.22	1.26	0.193	0.202
2	–0.55	–0.57	2.94	3.74	–0.12	–0.13	–0.14	–0.15	1.32	1.40	0.191	0.195
3	–0.65	–0.62	3.44	4.00	–0.13	–0.16	–0.23	–0.24	1.63	1.47	0.185	0.185
4	–0.67	–0.65	4.51	4.39	–0.18	–0.21	–0.31	–0.33	1.44	1.46	0.175	0.173
5	–0.63	–0.66	4.91	4.81	–0.24	–0.25	–0.37	–0.41	1.19	1.38	0.169	0.164
6	–0.65	–0.65	5.37	5.14	–0.28	–0.29	–0.47	–0.45	1.29	1.24	0.151	0.159
7	–0.63	–0.61	5.66	5.30	–0.29	–0.29	–0.49	–0.44	1.18	1.08	0.154	0.160
8	–0.56	–0.56	5.35	5.24	–0.29	–0.28	–0.39	–0.39	0.97	0.94	0.162	0.167
9	–0.49	–0.51	4.34	4.99	–0.23	–0.24	–0.27	–0.30	0.87	0.87	0.178	0.177
10	–0.43	–0.47	4.54	4.59	–0.19	–0.19	–0.16	–0.21	0.67	0.88	0.214	0.189
11	–0.51	–0.46	3.75	4.17	–0.15	–0.15	–0.15	–0.13	1.10	0.96	0.195	0.198
12	–0.51	–0.48	4.45	3.84	–0.14	–0.12	–0.13	–0.09	1.15	1.10	0.202	0.203
Average	–0.57	–0.56	4.50	4.49	–0.20	–0.20	–0.27	–0.27	1.17	1.17	0.181	0.181
R^2	0.867		0.541		0.951		0,948		0.789		0.789	

squares method. These determination coefficients derive from comparison of the gradients obtained by the two methods: $grad_m$ estimated on the basis of measurements and linear regression analysis (i.e. temperature versus altitude of meteorological station) and $grad_e$ calculated using the first harmonic of the Fourier expansion. The comparison of these two methods for the yearly variation of vertical gradients yielded a minimum determination coefficient for precipitation (R^2=0.54) and a maximum one for vapour pressure (R^2 = 0.95).

The Fourier expansion was then used to estimate the vertical gradient to temperature for each month. Temperatures at sea level were the calculated as T_A' T_B' in grids 1 and 5 (Fig. 1), from temperatures measured at the meteorological stations and the altitudes of these stations (Eqs. 1 and 2). Sea level temperatures

in all other grids (from 2 to 4 in Fig. 1) can be calculated using one of the standard interpolation or extrapolation procedures (Kriging, Thissen poligons, weight functions). Because of the variation in altitude within the Main river catchment the stations closer to the interpolation point weigh more heavily than those at a greater distance. In this paper the interpolation procedure was applied at sea level using the following equation (Schwarzmaier et al. 1992):

$$T'_{i,j} = \frac{\sum_{k=1}^{4} T'(k) \cdot \frac{1}{d(k)^2}}{\sum_{k=1}^{4} \frac{1}{d(k)^2}}, \tag{6}$$

where $T'_{i,j}$ is the interpolated air temperature in the grid (i, j) at sea level, $T'(k)$ air temperature measured at one of the four nearest meteorological stations and recalculated to sea level (Eqs. 1,2), k the number of one of the four nearest stations ranging from 1 to 4, and $d(k)$ the distance between station k and the grid where the temperature is interpolated.

After interpolation at sea level the values of temperature were again recalculated to the average altitude of each grid (from 1 to 5 in Fig. 1) using Eq. 3 and the previously vertical gradient of temperature.

Using the procedure described above grid maps of temperature for all months, starting in January 1957 and ending in December 1987, were obtained. Annual and long term average data were also analysed (Fig. 4). These grid-map data showed the areal distribution of temperature including the topography configuration of the analysed river catchment. Such an estimation allows for calculation of the average temperature for the whole river catchment over any time period (month, season, annual, long-term). This estimation was made for all months (1958 to 1987) using the equation:

$$T'_{(x)} = \frac{\sum_{i=1}^{40} \sum_{k=1}^{20} T_{i,j} \cdot a_{i,j}}{\sum_{i=1}^{40} \sum_{k=1}^{20} a_{i,j}}, \tag{7}$$

where: x is the number of the month (from 1 to 12 or 13 for the annual) T the average temperature for the whole river catchment, T the average temperature in the (i, j) numbered grid and a the normalised grid area (which varies from 0 to 10, 0 meaning that the grid is located outside the analysed river catchment, 10 meaning that the whole grid is included in the river catchment).

The whole procedure was repeated for each of the six other meteorological

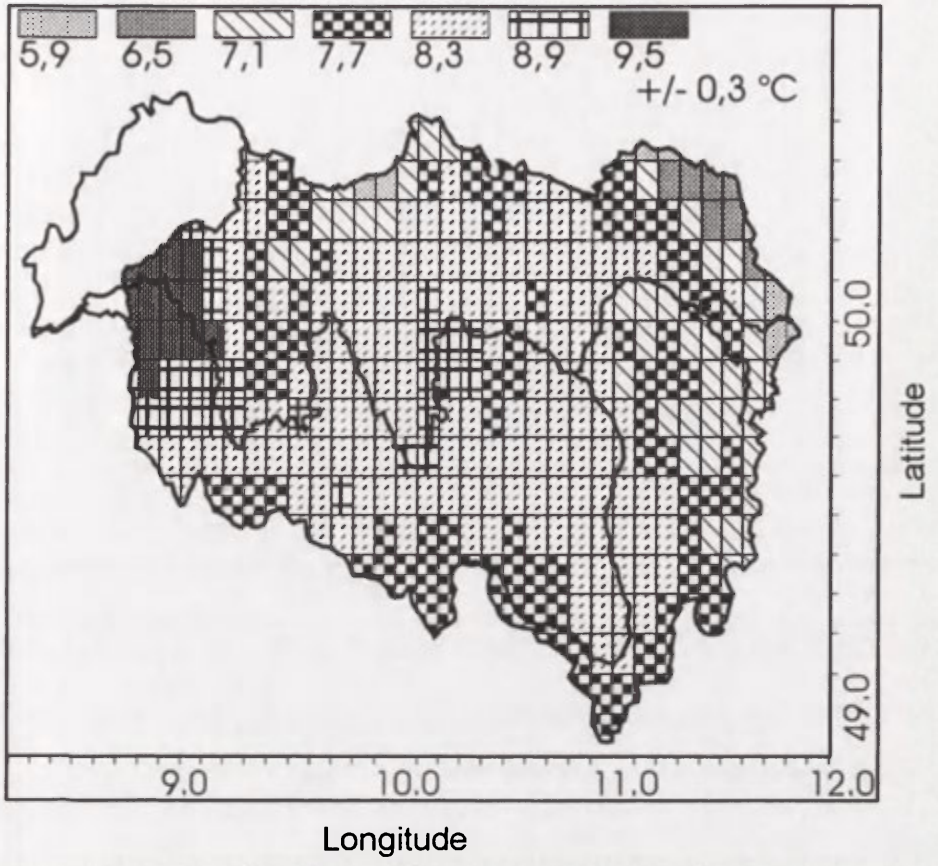


Figure 4. Areal distribution of long-term annual average air temperature in Main River catchment.

parameters. Figure 5a shows the average monthly values of temperature for the whole Main River catchment. On the basis of data obtained as described above the average annual temperature for the whole Main River catchment is shown in Figure 5b.

During the time period considered in this paper the annual temperature (average for the whole river catchment) ranged from 6.9°C in 1962 and 1963 to 8.8°C in 1959 and 1983. Annual precipitation ranged from 570 mm in 1964 to more than 1100 mm in 1965 and 1981.

COMMENTS AND CONCLUSIONS

The method presented here for estimating annual variation to vertical gradients on the basis of the first harmonic of the Fourier expansion, made possible the calculation of the areal distribution of different meteorological parameters in

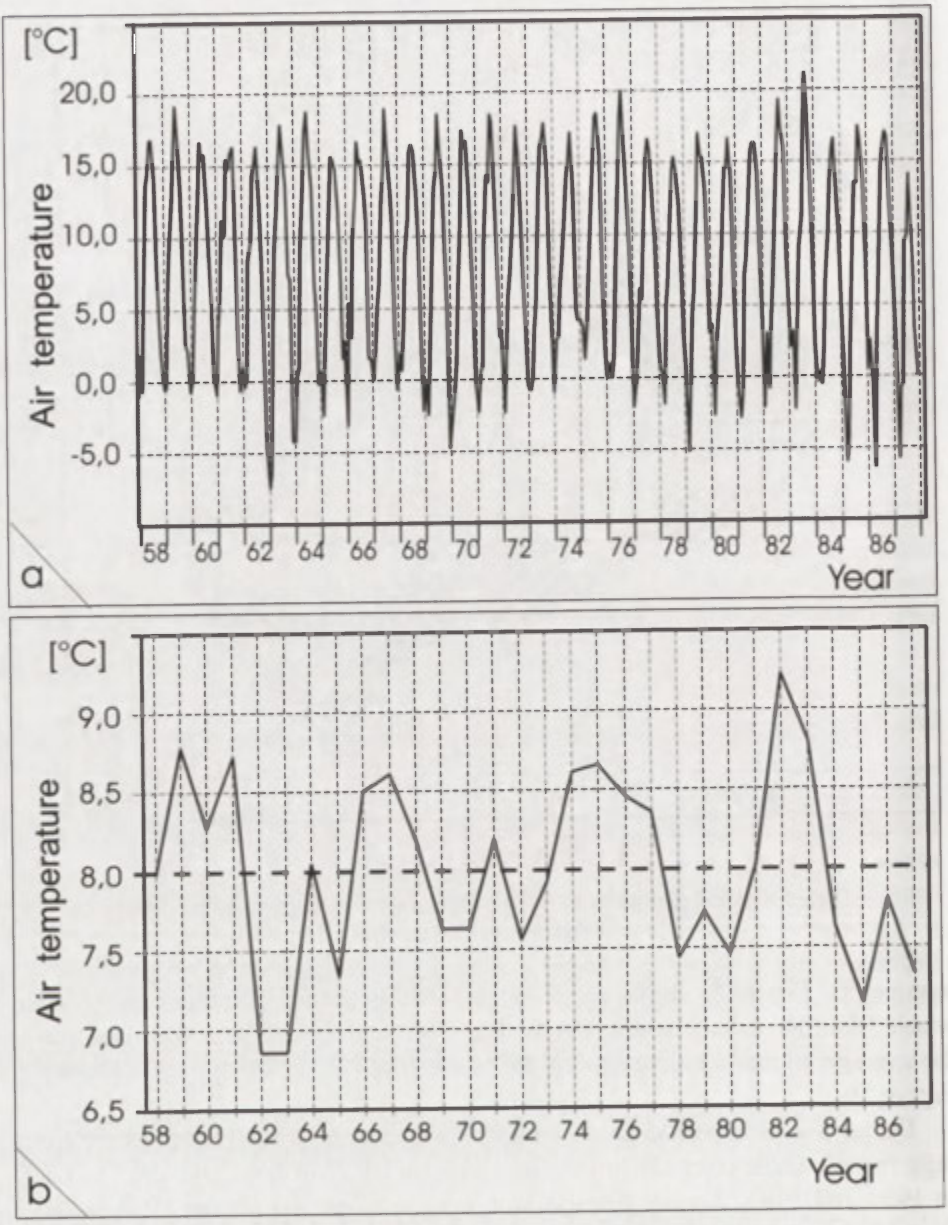


Figure 5. Course of monthly (a) and annual (b) averaged temperature from January 1958 to December 1987 and in Main River catchment – areal average for whole catchment.

the Main River catchment using a grid resolution of 0.1 latitude and 0.1 longitude. Because land use data for this region were available, the heat balance structure for the same period (1957 to 1987) could be estimated using the method proposed by Olejnik and Kędziora (1991). The use of estimated vertical gradients

seems an adequate tool for areal interpolation or extrapolation of different meteorological parameters with respect to altitude data. Using the proposed functions for vertical gradient estimation (Table 2), very high correlation coefficients have been noted for five of the analysed meteorological parameters (0.89 to 0.99). Only for precipitation was the correlation coefficient lower (at 0.69). This happened only because snowfall data from the winter season (Dec. to Feb., Fig. 3) as measured by a gauge are usually subject to higher measurement errors (Larson and Peck 1974). As the relative error to gradient estimation seemed to be no bigger than 0.5 mm/100 m the equation proposed for precipitation gradient estimation (Tab. 2) was also used in later calculations.

Through the use of both the procedure of initial transformation of standard meteorological data, and data concerning climate change in Europe which from GCMs (Hulme et al. 1990), it proved possible to analyse local changes in heat balance structure under the new anticipated climate conditions (Olejnik 1999).

In any models where the analysis of climate data is necessary, the proposed harmonic functions (Tab. 2) would seem to be a helpful tool because it is usually better to use functions in models than numbers (arrays), especially when a simple independent variable is used, here the number of the month. The use of functions instead of numbers is always faster and easier, and reduces the probability of errors. It is clear that the functions proposed in this paper cannot be used in any other climate conditions without earlier correction on the basis of local climate data. But after the estimation of the a, b, c coefficients in Eq. 5 (on the basis of local meteorological data), the Fourier expansion for estimating annual variation to vertical gradients can also be used in other regions.

REFERENCES

Angell J., 1988, *Variations and trends in tropospheric and stratospheric global temperatures 1958–1987*, Journal of Climatology I, 1296–1313.

Carter T. R., Parry M. L., Porter J. H., 1990, *Effects of climatic changes on agricultural potential in Europe*, [in:] Boer M. M., De Groot R. S. (ed.), *Landscape-Ecological Impact of Climatic Change*, IOS Press, Amsterdam, 327–346.

Diaz H. F., Bradley R. S., Eischeld J. K., 1989, *Precipitation fluctuation over global land areas since the late 1800s*, Journal of Geophysical Research, 94, 1195–1210.

Fabian P., 1991, *Klima und Wald- Perspektiven*, für die Zukunft. Forstw., Cbl., 110, 286–304.

Hansen J. E., Lebedeff S., 1988, *Global surface temperatures: update through 1987*, Geophysical Research Letters, 15, 323–326.

Hulme M., Wigley T. M. L., Jones P. D., 1990, *Limitations of regional climate scenarios for impact analysis*, [in:] Boer M. M., De Groot R. S. (ed.), *Landscape-Ecological Impact of Climatic Change*, IOS Press, Amsterdam, 111–129.

Jones P. D., Raper S. C. B., Bradley R. S., Diaz H. F., Kelly P. M., Wigley T. M. L., 1986, *Northern Hemisphere surface air temperature variations 1851–1984*, Journal Climate and Applied Meteorology, 25, 161–179.

Jones P. D., Wigley T. M. L., Farmer G., 1990, *Marine and land temperature data sets: a comparison and look at recent trends chap*, [in:] Schlesinger M. E. (ed.) *Greenhouse-*

Gas Induced Climatic Change: a Critical Appraisal of Simulations and observations, Elesevier, Amsterdam.

Karl T. R., Wiliams C. N., Jr, 1987, *An Approach to Adjusting climatological time series for discontinous inhomogeneities*, Jounal Climate and Applied Meteorology, 26, 1744–1763.

Larson W. L., Peck E. L., 1974, *Accuracy of precipitation measurement for hydrologic modeling*, Water Resources Research, 10, 4, 857–863.

Olejnik J., 1988, *Present and future estimates of evapotranspiration and runoff for Europe*, IIASA Working Paper, WP-88-037, Laxenburg, Austria.

Olejnik J., 1996, *Catchment scale modelling of heat and water balance structure under present and future climatic conditions*, Roczniki Akademii Rolniczej w Poznaniu, 268 (in Polish).

Olejnik J., Kędziora A., 1991, *A model for heat and water balance estimation and its application to land use and climate variation*, Earth Surface Proceeding and Landforms, 16, 601–617.

Schonwiese C. D., Malcher J., 1987, *The CO_2-temperature response. A comparison of the results from general circulation models with statistical assessment*, Journal of Climatology, 7, 215–229.

Schwarzmaier G., Mayer H., Bayer U., 1992, *Climatological criteria for estimation of runoff by linear regression*, International Journal of Climatology, 12, 103–110.

Solomon A. M., Leemans R., 1990, *Climatic change and landscape-ecological response: issues and analysis*, [in:] Boer M. M., De Groot R. S. (ed.), *Landscape-Ecological Impact of Climatic Change*, 293–316.

Verstraete M. M., 1989, *Land surface processes in climate models: status and prospects*, [in:] Berger A., Schneider S., Duplessy J. C. (ed.), *Climate and the Geosciences: A challenge for Science and Society in the 21th Century*, Kluwer Academic Publishers, Dordrecht, 321–340.

GEOGRAPHIA POLONICA, 73, 1, Spring 2000
PL ISSN 0016-7282

DIFFERENTIATION OF EUROPE'S MAMMAL FAUNA AGAINST A BACKGROUND OF BIOGEOGRAPHICAL UNITS, THE AREA OF UNITS AND MAMMALIAN TAXONOMIC RICHNESS

BOŻENNA GRABIŃSKA
Institute of Geography and Spatial Organization, Polish Academy of Sciences
ul. Twarda 51/55, 00-818 Warsaw, Poland
e-mail: b.grab@twarda.pan.pl

ABSTRACT: Analysis considered the relationship between the occurrence of mammals and geographical location in Europe, as well as the size of biogeographical units. It was found that there was differentiation in the qualitative composition of the mammal fauna in the north-south direction, which resulted from the overall zonal distribution of cool, temperate and hot areas. The correlation between the number of species and the area of a unit was not statistically significant in every case considered. Besides area, the regional variability of environmental conditions also has a decisive influence on the distribution of mammals in Europe.

KEY WORDS: Zoogeography, mammals, ranges, regionalization, Europe.

INTRODUCTION

The relation between the area of a region and the number of resident species has long fascinated ecologists (Williams 1995, 1996).

A fundamental problem in ecological research is to explain large scale gradients in species richness. The aim of the work was thus to answer a question regarding whether the species richness of a mammal fauna is a straight function of the area occupied by biogeographic units identified *a priori* and of the geographical location, or whether it is also the result of the action of the whole spectrum of variability of the environment in Europe.

Of primary influence on the distribution of mammals, and for that matter all animals, are temporally-variable internal and external factors. To be included among these are, above all, an ability to spread and overcome barriers, a dependency on climate, and vegetation as expressed most clearly in the zonal arrangement of environmental conditions. In particular climatic-vegetational zones, and within them in different types of vegetation and landscape, mammals even not closely related to one other generate similar morphological and physiological adaptations (Kerr and Packer 1997).

Mammals are a class of very wide geographical distribution. The terrestrial species are met with on all continents and on many islands, even those lying far

from the mainland. The most widespread are of course the bats and the rodents. Marine mammals (cetaceans, pinnipeds and manatees) inhabit all the oceans and seas linked to them.

The relation between the area of a region and the number of resident species has long fascinated ecologists. The areas occupied by different species, genera, families and orders of mammals are often very extensive, a feature linked with their considerable ecological plasticity. The sizes of the areas occupied also depend on the ability to overcome geographical barriers. Thanks to their ability to fly, bats have very wide ranges, while predatory species usually occupy wider ranges than herbivores because they are less dependent on a defined type of vegetation. Steppes pose a barrier to forest mammals, and vice versa, and sometimes even large rivers may limit the ranges of species or subspecies.

The greater or lesser mobility of particular species combines with their varied resistance to environmental factors to ensure that the distributions of particular species, genera and families of mammal reflect in different ways the differentiation of habitats often to a greater degree than the area over which these animals can live (Gaston et al. 1998). For example, relief, climatic conditions or the activities of man may disrupt or link ranges-leading to the differentiation of local faunas, even in areas located close to one another.

It is accepted that the distribution of mammals is conditioned by the differentiation of the natural environment, which creates a basis for the division of areas into spatial natural units (Brown et al. 1996). On the basis of the distribution of animals, conclusions may also be drawn as to the longitudinal-latitudinal gradient in the differentiation of environmental conditions, and the competition between organisms in homogenous environments (Udvardy 1978).

The differentiation of dynamic factors of the physical environment through chain reactions in which both plants and animals are included also influences changes in the biotic environment. Not every species of mammal reacts identically to the changes, but their reactions treated as a whole have an influence on the character of the evolution of the ecosystems which they create (Cox and Moore 1993).

Comparative, quantitative, biogeographic studies are revealing empirical patterns of interspecific variation in the sizes, shapes, boundaries, and internal structures of geographical ranges; these patterns promise to contribute to understanding the historical and ecological processes that influence the distributions of species.

In accordance with some zoogeographical theories, there is a simple relationship between the size of an area and the number of taxa inhabiting it (Margalef 1963, 1968; MacArthur and Wilson 1967). In order to verify these viewpoints in relation to mammals at least, the present work details statistical analysis of the above relationship, accepting as a basis the habitation by mammals of relatively homogeneous biogeogographical units established *a priori* and characterized by different sizes and physicogeographical properties.

METHODS

To obtain a statistical estimate of the number of taxa/area relationship, calculations were made of the numbers of species, genera and families per 10,000 km^2 area of divided biogeographical units. The values obtained were next transferred to a map of the biogeographic units of Europe (Grabińska 1992, 1994). The division applied comprises 58 biogeographical units. The principles used in this division derived from the phytogeographical differentiation of the European continent after H. Meusel (1965), and to some extent also from the boundaries of the physicogeographical division of J. Kondracki (1988). A map of Europe set against this biogegraphical division was used as the basic map upon which to present mammalian taxonomic richness. It is always a controversial matter in biogeography as to where to delimit the units in the division into regions. This work with units distinguished *a priori* and the results obtained for the occurrence of mammals can designate one of many possible ways by which to divide Europe into zoogeographical regions. Another view is that the ranges and areas of occurrence of animals are the basis for zoogeographical regionalisation of J. Pawłowski (1983). A zoogeographical division into regions should first follow from a comparison of the ranges of animals and only then from qualitative evaluation of taxonomic composition in the areas analysed. This division could be very much based on the presence of native species that do not occupy large ranges, with the ubiquitous, eurytopic species being disregarded, along with those whose geographical ranges are very extensive (such as the Eurosiberian element).

The abundance of the studied taxa per 10 000 km^2 is presented on Maps 1–3 (Figs. 1, 2 and 3). In addition, Figure 4 presents the deviation from the mean number of species per 1000 km^2 of unit calculated for the whole area of Europe (Fig. 4). The next stage involved the statistical analysis of material concerning an estimate of the degree of qualitative diversity of the mammal faunas in relation to geographical location and size of unit.

For a better visualization of the discussed relationships, the 58 biogeographical units of Europe were divided into 8 groups of wider range linked with defined phytogeographical zones (Meusel 1965).

Statistical analysis using regression and correlation thus concerned estimates of faunistic differentiation within groups of units and not single units.

These are:

– group I, northern units including mainly arctic-boreal phytogeographical formations (Nos. 1, 2, 3, 4, 5, 6 and 7),

– group II, bringing together units in temperate areas with a predominance of Atlantic and sub-Atlantic phytogeographical formations (Nos. 8, 9, 10, 17 and 24),

– group III, of a sub-Atlantic/sub-Continental nature (Nos. 11, 12, 16, 18, 19 and 20),

– group IV, clearly sub-Continental (forest zone) (Nos. 13, 14, 15 and 21),

Figure 1. Number of mammal species per 10,000 sq.km of unit area.

– group V, sub-Continental/Continental steppe and semidesert zone (Nos. 22, 23, 27, 28, 31, 32 and 33),

– group VI, sub-Mediterranean and near-Mediterranean units (Nos. 35, 38, 40, 44, 47, 48 and 55),

– group VII, Mediterranean units (Nos. 25, 34, 37, 39, 41, 42, 43, 45, 50, 51, 52, 53 and 58).

The European mountain areas were separated out as a separate group VIII, consisting of Nos. 26, 29, 30, 36, 46, 49 and 56.

On the basis of the values obtained, a function was described relating differences in the number of species to the size of an area. Determined with these parameters was a regression of one factor against the other. Calculations analogous to those for species were carried out for taxa of higher order, namely genera and families. This allowed for wider conclusions to be drawn regarding the distribution of mammals. An evaluation of the strength of the relationship between the features was performed using linear correlation calculations.

Figure 5 presents the regression lines describing the relationship between the dependent variable Y (the number of animal taxa in the unit) and the independent variable X (the area of the unit).

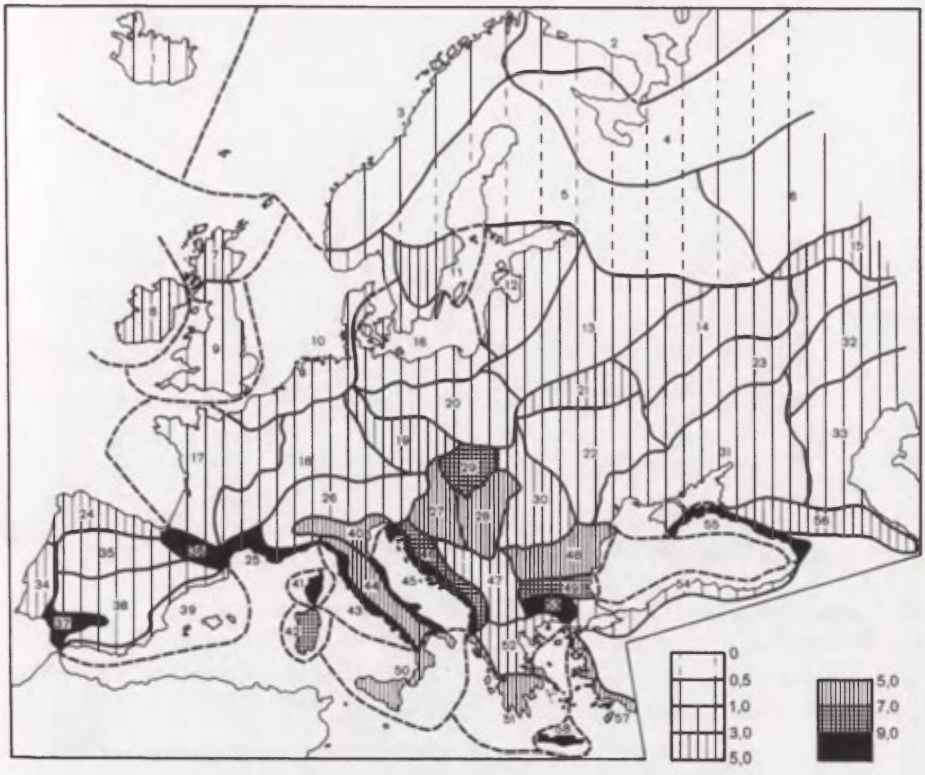


Figure 2. Number of mammal genera per 10,000 sq.km of unit area.

Figure 5 also contains parameters a and b of the regression equations, describing the relationship between variable Y for defined values of variable X. These are at the same time the lines of best fit for the empirical data.

RESULTS

As was to have been expected, the fewest mammal species per 10,000 km^2 occur in the northern regions of Europe and the greatest number in the south (Fig. 1). The values obtained from the temperate zone were of an intermediate nature, and the upland and mountainous regions of central Europe stand out particularly from this point of view. The greatest density of species characterizes the southern, sub-Mediterranean and Mediterranean regions which are most diversified from the physicogeographical point of view.

A similar picture of diversity is also true for the higher taxonomic units (Figs. 2 and 3).

Figure 3. Number of mammal families per 10,000 sq.km of unit area.

Map 4 (Fig. 4) gives a similar information content where the faunistic richness of the discussed biogeographic units is concerned.

The negative deviations from the mean number of species per thousand kilometres are greatest for northern, Arctic, sub-Arctic and boreal Siberian regions, while the greatest positive deviations are noted in southern regions, especially in the near-Mediterranean and Mediterranean areas, the Pontic-Pannonian region and certain mountainous areas.

Regression and correlation analyses show that only group I – of Arctic and Boreal units – has a number of species significantly correlated with the area of the unit (correlation coefficient 0.78) (Fig. 5). In contrast, in southern Europe – where the scale of the variability in biotic and physical environments is greatest – as well as in central parts, there is a lack of any clear relationship between the variables analyzed.

In group I – northern, Arctic and Euro-Siberian boreal units – taxa of higher rank also have a particularly high and statistically-significant correlation. For genera, the correlation coefficient is 0.73, while for families it is 0.79. These relations are best expressed in a linear form (Fig. 5).

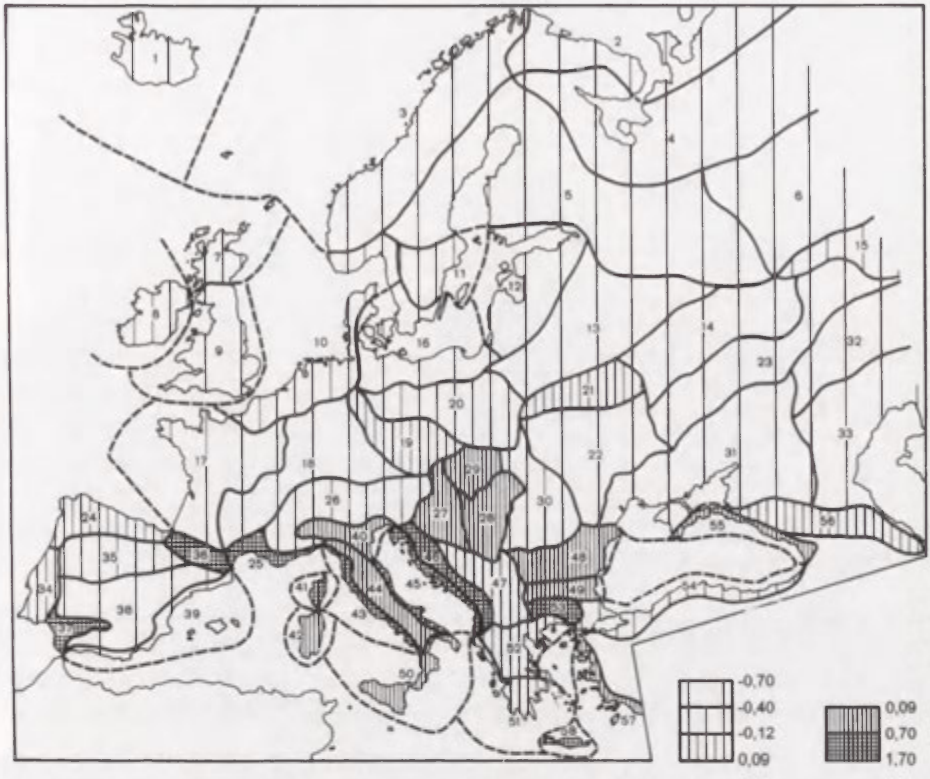


Figure 4. Deviation from the mean number of species per 10,000 sq.km of Europe.

DISCUSSION OF RESULTS

Climatic factors, environmental stability, land area, habitat heterogeneity, historical influences (such as Pleistocene glaciations) and energy availability are the factors most often discussed as determinants of regional variability in species richness (Kerr and Packer 1997).

The picture obtained of the relationships between the mammalian fauna and area is basically in agreement with the latitudinal differentiation of the climate. Disturbing this is the action of local factors having a recurrently greater influence on the structure of the theriofauna than macroclimate or the size of an area (Kowalski 1959, 1973, 1984). There is a projection of vegetation patterns in Europe in the zonal arrangement of the fauna that is particularly clear. The question remaining to be addressed concerns differentiation within the main climatic-vegetational zones (groups I–VIII). The maps presented clearly show that – besides zonal differentiation – mountain areas also have great significance on the diversity of habitats, while the larger islands of the Mediterranean have large numbers of mammal taxa per 10,000 km^2 in spite of their small areas. It is

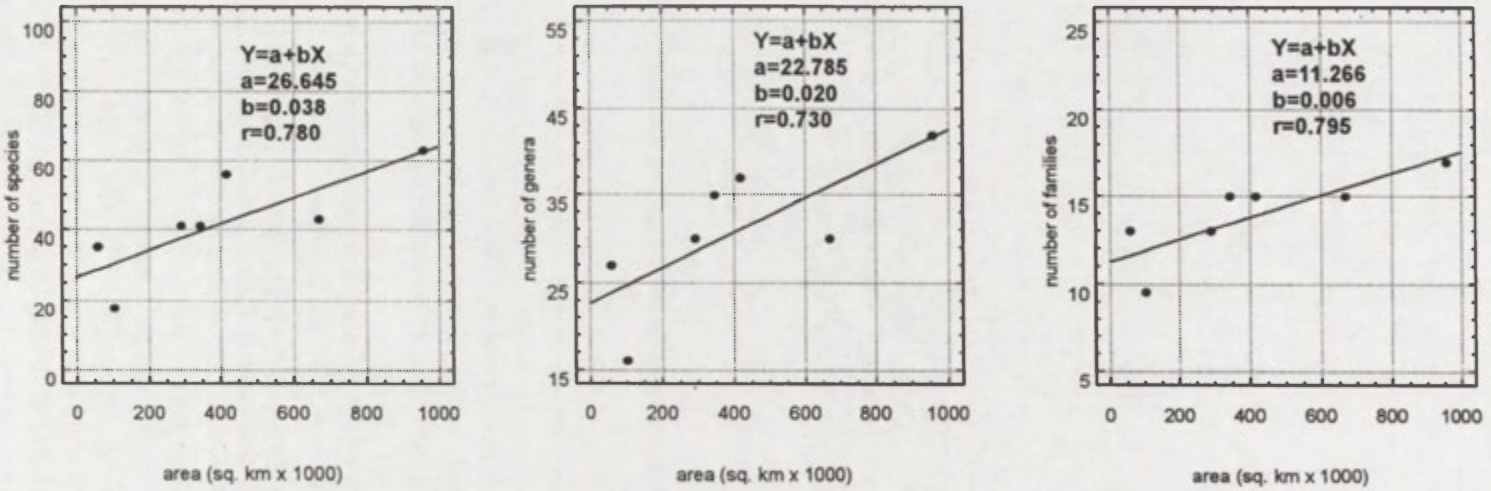


Figure 5. Statistically significant regressions between the number of mammal taxa and the unit area (for the I Groups of units). r – correlation coefficient.

in these very near-Mediterranean and Mediterranean areas, where the diversity of habitats is greatest and there is a passage between coastal and mountain areas, that there occurs the greatest faunistic diversity.

Analysis of variance was carried out to evaluate the link between the area of a unit and the species richness of its fauna. This showed different degrees of significance of this relationship and confirmed its specifically local character. For it emerged that – with enhanced diversity of environmental conditions there is a lack of any correlation between area and the number of taxa.

A variable situation exists in northern areas where severe impacts of the climate are joined by remaining physicogeographical conditions that are quite homogenous. In the north, the size of the area in the sense of food area is a clear factor limiting the number of species. The limiting effect of extreme physical factors in northern and polar environments ensures that the fauna occurring there is poor.

The further south, the more taxa per unit area, attaining a maximum in the Mediterranean zone, as well as in the mountain massifs lying beyond the range of the last Pleistocene glaciations and naturally highly diversified from the point of view of life conditions. This is clearly illustrated on the map showing deviations from the mean number of species per 1000 km^2. Regions with the greatest faunal density in which the number of mammal species per unit area diverges in a positive direction from the mean for Europe as a whole are sub-meridional and meridional areas with climates ranging from the Atlantic to the sub-continental. This confirms information in litt. on mammals (and other faunal groups) which points to these areas as "hot spots" (Gaston and David 1994).

Comparisons of the densities of species with those of taxa of higher rank (Figs. 1–3) show that general similarities do not disguise certain important differences. This results not only from the fact that the ranges of the majority of genera or families are greater than those of single species, but also from historical conditioning which shapes not so much their genesis as their stage of colonization. The broad distribution of taxa of higher rank is thus first and foremost a function of their history. For example, genera including closely related species which underwent differentiation initially in different nearby ecosystems later came to dominate or were driven into fundamentally different ecosystems. The same applies to the structural and ecological properties of genera within families, which may differ greatly in relation to the limiting impact of environmental elements in the ecosystems in which they occur. A particular role in faunistic differentiation is played by long-term isolation, thanks to which adaptive radiation within isolated groups may proceed in different ways (Szarski 1990).

Comparative, quantitative biogeographic studies are revealing empirical patterns of interspecific variation in the size, shapes, boundaries, and internal structures of geographic ranges; these patterns promise to contribute to understanding the historical and ecological processes that influence the distributions of species (Brown et al. 1996).

After collating the facts on the occurrence of mammals in different geographical regions of Europe, a further analysis which should be carried out is research into the relationships between the structure and diversity of the geographical environment and the structure of the fauna. Geomorphological, hydrographic, climatic differences etc., as well as the impact of different forms of human action, determine the real areal presence and thus the zoogeographical diversity of a studied area (Williams 1995,1996). These very detailed studies must however be preceded by macrogeographical analysis which would be a basis for all micro--zoogeographical considerations concerning above all the chorology of particular taxa in full, from their distribution through their genesis, ecology etc. This is because it is known that the distribution and properties of the majority of mammal species vary geographically in accordance with the degree of environmental pressure (Serafiński, Wielgus-Serafińska 1976).

The zoogeographical analysis carried out, along with the distribution map of physicogeographical units and faunistic characterizations, may constitute part of a databank on the natural environment and may be used for variously-directed evaluation work.

In their more diversified life environments and adaptation to defined conditions, besides their general adaptive features, mammals have a character of specialization. They are a group associated with diverse environments marked by enormous physiological plasticity and clear morphological differences connected with the life they lead. Types of morphoethology emerge in connection with this and result in the specifics of regional faunas, which are particularly visible in the group of southern regions where we deal with the greatest physicogeographical diversity (islands, coasts and mountains). For land mammals, the most significant barriers are of course the seas. The ability to pass across them defines the number of taxa in southern Europe in the near-Mediterranean regions (groups VI and VII). Here there is not therefore a close relationship between the occurrence of species and the areas of physicogeographical units, and there is only a weaker correlation where families are concerned.

The seas are joined by many other barriers which stand in the way of the dispersal of mammals. Their influence is however on the expansion of particular species on the microscale in relation to climatic and physicogeographical conditions (Schüle 1993).

The result of the impacts of geocomponents within regions is the internal differentiation of small units. This attests to the fact that adaptive properties may change within zones the distribution of clinal species e.g. in relation to elevation above sea level, latitude or other external internal gradients (Kowalski 1959, 1973, 1984).

CONCLUSIONS

The results obtained allowing the following conclusions to be drawn:

– in the majority of biogeographical units there is no correlation between the number of mammal species or taxa of higher rank, and the size of the area these units occupy,

– a high and statistically-significant correlation between the number of all the analyzed taxa (species, genera and families) and the coresponding area was only noted for the regions on the northern edge of the continent,

– there is a gradient of increasing taxonomic richness of the mammal fauna in the north-south direction. This results from the overall zonal distribution of cool, temperate and hot areas, disturbed by locals-cale enhancement of the geomorphological, altitudinal, hydrographic, climatic and biotic diversity of habitat conditions.

REFERENCES

Brown J. H., Stevens G. C., Kaufman D. M., 1996, *The Geographic Range: Size, Shape, Boundaries, and Internal Structure*, Annual Review of Ecological System, 27, 597–623.

Cox C. B., Moore P. D., 1993, *Biogeography. An ecological and evolutionary approach*, Blackwell Scientific Publ.

Gaston K., David R., 1994, *Hot spots across Europe*, Biodiversity Letters 2, 108–116.

Gaston K., Blackburn T. M., Spicer J. I., 1998, *Rapoport's rule: time for an epitaph?*, Tree, 13,2, 70–74.

Grabińska B., 1992, *Zoogeograficzne zróżnicowanie fauny ssaków w Europie (Zoogeographical distribution of the mammal fauna in Europe)*, Przegląd Geograficzny, 3–4, 311–324.

– 1994, *An attempt at the zoogeographical regionalization of Europe on the basis of the distribution of vertebrates*, Geographia Polonica, 63, 105–118.

Kerr J. T., Packer E., 1997, *Habitat heterogeneity as a determinant of mammal species richness in high-energy regions*, Nature, 385/16, 252–254.

Kowalski K., 1959, *Katalog ssaków plejstocenu Polski*, PWN, Warszawa–Wrocław.

– 1973, *Rozwój fauny w czwartorzędzie Polski*, [in:] *Metodyka badań osadów czwartorzędowych*, PWN, Warszawa.

– 1984, *Klucz do oznaczania ssaków Polski*, PWN.

Kondracki J., 1988, *Geografia fizyczna Polski, (Physicogeography of Poland)*, PWN, Warszawa.

MacArthur R. H., Wilson E. O., 1967, *The theory of island biogeography*, Princeton University Press, Princeton.

Meusel H., 1965, *Vergleichende Chorologie der Zentraleuropäischen Flora*, Veb Gustav Fischer Verlag Jena, vol. I and II.

Margalef R., 1963, *On certain unifying principles in ecology*, American Naturalist, 97, 357–374.

– 1968, *Perspectives in ecological theory*, the University of Chicago Press, Chicago–London.

Pawłowski J., 1983, *Zróżnicowanie faunistyczne województwa Bielskiego*, Folia Geographica, S. Geographica-Physica, XV, Kraków.

Serafiński W., Wielgus-Serafińska E., 1976, *Ssaki (Mammals)*, PWN, Warsaw.

Schüle W., 1993, *Mammals, vegetation and the initial human settlement of the Mediterranean islands: a paleoecological approach*, Journal of Biogeogrphy, 20, 399–412.
Szarski H., 1990, *Historia zwierząt kręgowych (History of vertebrates)*, PWN, Warszawa.
Udvardy M., 1978, *Zoogeografia dynamiczna (Dynamic zoogeography)*, PWN, Warszawa.
Williams H. R., 1995, *An extreme-value function model of the species incidence and species--area relations*, Ecology, 76(8), 2607–2616.
– 1996, *Species area curves: the need to include zeros*, Global Ecology and Biogeography Letters 5, 91–93.

INFORMATION FOR AUTHORS

The editors of GEOGRAPHIA POLONICA invite theoretical and empirical contributions to human as well as physical geography, in broad terms. Articles are expected to be original and not yet published elsewhere unless in languages other than English.

MANUSCRIPTS. Manuscripts, in the English language, should be submitted to the Editor of GEOGRAPHIA POLONICA (Piotr Korcelli, Institute of Geography and Spatial Organization, Polish Academy of Sciences, Twarda 51/55, 00-818 Warsaw, Poland).

As all manuscripts will be reviewed, therefore, authors are kindly requested to provide two hard copies of the complete text (main text, abstract, references, tables, illustrations (one set of illustrations may be in xeroxed form) and one 3.5" diskette in MS Word.

The manuscripts should be arranged in the following order. First sheet: title, full name of author(s), affiliation, full postal address, e-mail address (if possessed). Second sheet: abstract of no more (no less) than 100 words, key words (3–10) at the end of the abstract. Subsequent sheets: the main text of about 20–25 pages double-spaced (in A4 format). Then on separate sheets: acknowledgements (if desired), notes, references, tables, illustrations, captions to illustrations.

The main text may be organized in sections under appropriate headings without numerals.

SPACING. Manuscripts should be printed with double spacing without indents; for paragraphs enter a line space. References and abstracts should be listed triple-spaced.

REFERENCES should be printed in alphabetical sequence at the end of the text. Examples:

Articles from journals: Berry B. J. L., 1960, *The impact of the Expanding Metropolitan Communities upon the Central Place Hierarchy*, Annals of Association of American Geographers, 50, 112–116.

Books: Stren R., White R., Whitney J., 1992, *Sustainable Cities*, London, Jessica Kingsley Publishers.

Chapters from books: Dematteis G., 1996, *Toward a Unified Metropolitan System in Europe: Core Centrality Versus Network Distributed Centrality*, [in:] Pumain D., Saint-Julien T. (eds), *Urban Networks in Europe*, INED, John Libbey, Paris, 19–28.

Theses: Elfring T., 1987, *Service employment in advanced economies*, Unpublished Ph. D. Thesis, Rotterdam, Erasmus University, School of Management.

References should be indicated in the text by listing, in parenthesis, the author's name followed by the data of publication, eg. (Kowalski 1995) or (Kowalski et al. 1996) if there are more than two authors (Note: all the authors should be listed in the references).

NOTES should be referred to by superscript letters.

FOOTNOTES are not recommended and should be used sparingly.

ILLUSTRATIONS (photographs and figures) must be in a camera-ready form suitable for reproduction or scanning. Lines and symbols should allow for size reduction, if required. Figures (maps, graphs, diagrams and drawings) should be in black and white, laser printed or as clear drawings. Illustrations should be sequentially numbered with Arabic numerals (eg. Figure 1) and prepared on separate sheets. Their position in the text should be indicated.

TABLES should be printed on separate sheets and numbered sequentially with Arabic numerals (eg. Table 1). Each table should be provided with a clear descriptive caption at the top and informative column headings.

EQUATIONS and SYMBOLS used in mathematical formulae must be clearly explained. Axes on graphs should be described clearly. Units and abbreviations used by authors should conform to the International List.

MEASUREMENTS should be given in metric units.

ACCEPTED MANUSCRIPTS. After acceptance for publication, authors will receive page proofs (which should be returned promptly) for approval before publication. The revised manuscript should be returned to the Editor by air-mail.

Articles accepted for publication in GEOGRAPHIA POLONICA are not honoured financially. Authors of accepted articles will receive one copy of the journal free of charge.

Authors are responsible for indicating copyright as well as permission from the copyright holder to reproduce any texts, tables or illustrations for which copyright exists.

GEOGRAPHIA POLONICA

Vol. 73 No. 1 Spring 2000

CONTENTS

ISSN 0016-7282