AN INDEX OF THE REGIONAL SPECIALIZATION OF INDUSTRY ON A COUNTY LEVEL

F. Móricz — Gy. Krajkó — Mrs. Abonyi

Introduction

Regional specialization is a category which denotes quantity and refers to the development level of a given economic region. Researches concerning specialization are of great interest with the introduction of an indirect control system. During the past years researches concerning this matter underwent a considerable development, the circle of the researches was extended, the methods developed and new methods — mathematical methods — were introduced.

Different characteristic features and collateral phenomena of the regional specialization were researched by Hungarian as well as foreign experts. Among the foreigners it is the Englishman John N. H. Britton, the Russians Kalashnikova T. T. and I. G. Saushkin, and among the Hungarians it is Csete L., Erdei F., and Kőszegi L. whose works are of special interest and revitalised the technical literature concerning the matter.

Specialization is a form of social division of labour by which the demands of one or more chosen branches of production come into prominence by subordinating or coordinating the others. "The branches of production which make up a considerable proportion of the production of the country and the centre as well as of foreign trade, branches which have a dominating effect in a given area and contribute to the effectiveness of the principle 'minimum labour — maximum returns' belong to the specialization of the region". With the development of the forces of production and the relations of production the specialization is extended, production as such as well as the production of goods increases, manufacturing costs go down and productivity increases. In the framework of socialism, for the sake of planned, proportional development as well as for the sake of the effectiveness of the national economy, the increase of regional specialization is necessary objectively.

Industrial specialization can be researched in different respects. One can observe structural changes concerning the whole industry, as well as the division of labour among certain units of production, certain factors, units of administration, and economic centres. The increase of horizontal and vertical division of labour follows the tendency of the development. Klaaren, Th. A., using the Heckscher—Ohlin model in researching regions in the U.S.A., came to the conclusion that "an economic region is liable to specialize in producing goods for which there is a relative abundance of the parts necessary for the production". The author made comparative and reciprocal investigations by multilateral regression-analysis. As a result of these investigations, he came to the following conclusion: "Highly industrialised regions should be specialized in those branches of industry which have production functions demanding more of the regional factors in abundance."

With this present investigation we want to know more about the regional difference of the industrial specialization connected with counties. In this work we do not want to deal with the non-utilized possibilities of division of labour, the economic success of its development, its financial, technical and material conditions, we do not want to touch the questions of the flexibility of the structurs of production. We only want to present a method of measuring the degree of specialization within the limits of Hungarian statistics.

The Method of the Investigation

In the case of a given regional unit (a county, etc.) Q_i is a quantity referring to the branch i-involved in the investigation (which can be the number of employment working in the given branch, or the value of the investment stocks belonging to the branch) and

$$Q = Q_1 + Q_2 + ... + Q_n ,$$

where n means the number of branches involved in the investigation. It is advisable to use ratios instead of statistical data (i.e. absolute numbers). Then the P_i index referring to the proportion of branch i is given by the formula

$$P_i = \frac{Q_i}{Q} \quad (i = 1, 2, ..., n).$$

Obviously

$$P_1 + P_2 + ... + P_n = 1 .$$

Let $\bar{Q}$ be the *arithmetical mean* of the quantities $Q_1, Q_2, ..., Q_n$:

$$\bar{Q} = \frac{1}{n} \sum_{i=1}^{n} Q_i = \frac{Q}{n} ,$$

and let σ be the *standard deviation* of these quantities:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Q_i - \bar{Q})^2} .$$

The *relative deviation* V is defined by

$$V = \frac{\sigma}{\bar{Q}} .$$

The value of this formula is between 0 and $\sqrt{n}$ in the case of n branches.

The specialization index I of the given area is interpreted as the quotient of the relative deviation V and $\sqrt{n}$:

$$I = \frac{V}{\sqrt{n}} .$$

The specialization index can be counted up from the rations $P_1, P_2, ..., P_n$ in a less complicated way:

$$I = \frac{V}{\sqrt{n}} = \frac{1}{\sqrt{n}\,\bar{Q}} \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Q_i - \bar{Q})^2} = \sqrt{\frac{1}{n^2 \bar{Q}^2} \sum_{i=1}^{n} (Q_i - \bar{Q})^2} =$$

$$= \sqrt{\sum_{i=1}^{n} \left(\frac{Q_i}{Q} - \frac{1}{n} \right)^2} = \sqrt{\sum_{i=1}^{n} \left(P_i - \frac{1}{n} \right)^2},$$

as

$$Q = n\bar{Q} \quad \text{and} \quad P_i = \frac{Q_i}{Q}.$$

Taking into consideration that $\sum_{i=1}^{n} P_i = 1$ results that

$$\sum_{i=1}^{n} \left(P_i - \frac{1}{n} \right)^2 = \sum_{i=1}^{n} P_i^2 - \frac{2}{n} \sum_{i=1}^{n} P_i + \frac{1}{n} = \sum_{i=1}^{n} P_i^2 - \frac{1}{n},$$

we come to the result of

(1)
$$I = \sqrt{\sum_{i=1}^{n} P_i^2 - \frac{1}{n}}.$$

Now we can interpret some characteristics of the specialization index.

1. The inequality

$$0 \leq I \leq \sqrt{\frac{n-1}{n}}$$

is always valid. The low estimation is caused by the fact that the relative deviation V is always non-negative. The high estimation can be interpreted as follows: The specialization index I is maximal only if there is one chosen branch described of the distribution of branches in our chosen regional unit, i.e. if for a certain i $P_i = 1$, while for all the others

$$P_1 = P_2 = \ldots = P_{i-1} = P_{i+1} = \ldots = P_n = 0.$$

Then according to (1) we have

$$I = \sqrt{1 - \frac{1}{n}}.$$

2. If we divide the given mezo-centre into counties, there is an obvious connection between the specialization index of the counties and those of the mezo-regions. To understand this, denote by Q_i^k the quantity referring to the branch number i in the county k, where $i = 1, 2, \ldots, n$; $k = 1, 2, \ldots, m$; m means the number of the counties. It is obvious, that

$$Q_i = Q_i^{(1)} + Q_i^{(2)} + \ldots + Q_i^{(m)} \quad (i = 1, 2, \ldots n).$$

Then

(2)
$$Q^{(k)} = Q_1^{(k)} + Q_2^{(2)} + \ldots + Q_n^{(k)} \quad (k = 1, 2 \ldots m).$$

On all this the contribution of branch i in the county k can be denoted by the formula

$$P_i^{(k)} = \frac{Q_i^{(k)}}{Q^{(k)}} \quad (i = 1, 2, \ldots, n; \; k = 1, 2, \ldots m).$$

The quantities Q_i^k, Q_i^k, Q_i^k, and Q can be seen on the following break-down:

					altogether
1. county:	$Q_1^{(1)}$	$Q_2^{(1)}$	$\ldots$	$Q_n^{(1)}$	$Q^{(1)}$
2. county:	$Q_1^{(2)}$	$Q_2^{(2)}$	$\ldots$	$Q_n^{(2)}$	$Q^{(2)}$
$\cdots$					
m. county:	$Q_1^{(m)}$	$Q_2^{(m)}$		$Q_n^{(m)}$	$Q^{(m)}$
mezo-region	Q_1	Q_2		Q_n	Q
(altogether)					

The arithmetical mean $\bar{Q}^{(k)}$, the standard deviation σ_k, and the relative deviation V_k referring to the county number k are defined by

$$\bar{Q}^{(k)} = \frac{1}{n} \sum_{i=1}^{n} Q_i^{(k)} = \frac{Q^{(k)}}{n},$$

$$\sigma_k = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Q_i^{(k)} - \bar{Q}^{(k)})^2},$$

$$V_k = \frac{\sigma_k}{\bar{Q}^{(k)}}.$$

According to the explanation the specialization index of the county number k is

$$I_k = \frac{1}{\sqrt{n}} V_k = \frac{1}{n\bar{Q}^{(k)}} \sqrt{\sum_{i=1}^{n} (Q_i^{(k)} - \bar{Q}^{(k)})^2} \quad (k = 1, 2, \ldots, m).$$

According to (1)

$$n\bar{Q}I = \sqrt{\sum_{i=1}^{n} (Q_i - \bar{Q})^2}.$$

According to (2) and the obvious fact that

$$\bar{Q} = \sum_{k=1}^{m} \bar{Q}^{(k)},$$

we come to the result as follows:

$$n^2 \bar{Q}^2 I^2 = \sum_{i=1}^{n} (Q_i - \bar{Q})^2 = \sum_{i=1}^{n} \left(\sum_{k=1}^{m} Q_i^{(k)} - \sum_{k=1}^{m} \bar{Q}^{(k)} \right)^2 = \sum_{i=1}^{n} \left[\sum_{k=1}^{m} (Q_i^{(k)} - \bar{Q}^{(k)}) \right]^2 =$$

$$= \sum_{i=1}^{n} \left[\sum_{k=1}^{m} (Q_i^{(k)} - \bar{Q}^{(k)})^2 + 2 \sum_{i \leq k < j \leq m} (Q_i^{(k)} - \bar{Q}^{(k)})(Q_i^{(j)} - \bar{Q}^{(j)}) \right] =$$

$$= \sum_{k=1}^{m} \sum_{i=1}^{n} (Q_i^{(k)} - \bar{Q}^{(k)})^2 + 2 \sum_{1 \leq k < j \leq m} \sum_{i=1}^{n} (Q_i^{(k)} - \bar{Q}^{(k)})(Q_i^{(j)} - \bar{Q}^{(j)}).$$

The double factor can have a very simple meaning if we introduce the correlation coefficient between the data $Q_1^{(k)}$, $Q_2^{(k)}$, ..., $Q_n^{(k)}$ of the county k and the data $Q_1^{(j)}$, $Q_2^{(j)}$, ..., $Q_n^{(j)}$ of the county j:

$$(3) \quad R_{kj} = \frac{\frac{1}{n}\sum_{i=1}^{n}(Q_i^{(k)}-\bar{Q}^{(k)})(Q_i^{(j)}-\bar{Q}^{(j)})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_i^{(k)}-\bar{Q}^{(k)})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_i^{(j)}-\bar{Q}^{(j)})^2}} =$$

$$= \frac{\sum_{i=1}^{n}(Q_i^{(k)}-\bar{Q}^{(k)})(Q_i^{(j)}-\bar{Q}^{(j)})}{\sqrt{\sum_{i=1}^{n}(Q_i^{(k)}-\bar{Q}^{(k)})^2 \sum_{i=1}^{n}(Q_i^{(j)}-\bar{Q}^{(j)})^2}} .$$

With arrangement we gain that

$$\sum_{i=1}^{n}(Q_i^{(k)}-\bar{Q}^{(k)})(Q_i^{(j)}-\bar{Q}^{(j)}) = R_{kj}n\bar{Q}^{(k)}I_k n\bar{Q}^{(j)}I_j,$$

where we took into consideration the formula of the specialization indices I_k and I_j. Finally we gain that

$$n^2\bar{Q}^2 I^2 = \sum_{k=1}^{m}n^2(\bar{Q}^{(k)})^2 I_k^2 + 2\sum_{1\leq k<j\leq m}n^2\bar{Q}^{(k)}\bar{Q}^{(j)}R_{kj}I_kI_j.$$

We can simplify both sides of the equity with n^2 and on multiplying with $\bar{Q}^2$ we get that

$$I^2 = \frac{1}{\bar{Q}^2}\left[\sum_{k=1}^{m}(\bar{Q}^{(k)})^2 I_k^2 + 2\sum_{1\leq k<j\leq m}\bar{Q}^{(k)}\bar{Q}^{(j)}R_{kj}I_kI_j\right].$$

Here we can substitute $\bar{Q}$ with Q and $\bar{Q}^{(k)}$ with $Q^{(k)}$:

$$(4) \quad I^2 = \frac{1}{Q^2}\left[\sum_{k=1}^{m}(Q^{(k)})^2 I_k^2 + 2\sum_{1\leq k\leq j\leq m}Q^{(k)}Q^{(j)}R_{kj}I_nI_j\right].$$

Here we get a weighted square mean, where the amount of the weights is:

$$\frac{1}{Q^2}\left[\sum_{k=1}^{m}(Q^{(k)})^2 + 2\sum_{1\leq k<j\leq m}Q^{(k)}Q^{(j)}\right] = \frac{1}{Q^2}\left[\sum_{k=1}^{m}Q^{(k)}\right]^2 = 1.$$

The formula (4) denotes the connection between the specialization index I of the mezo-region and the specialization index I_k of the counties ($k=1, 2, ..., m$).

3. The following relation is always valid:

$$I \leq \max(I_1, I_2, ..., I_m) = I_{max}.$$

Because on the basis of the formulas (4) and (2)

$$I^2 \leq \frac{1}{Q^2}\left[\sum_{k=1}^{m}(Q^{(k)})^2 I_{max}^2 + 2\sum_{1\leq k<j\leq m}Q^{(k)}Q^{(j)}I_{max}^2\right] =$$

$$= \frac{I_{max}^2}{Q^2} \left[\sum_{k=1}^m (Q^{(k)})^2 + 2 \sum_{1 \le k < j \le m} Q^{(k)} Q^{(j)} \right] = \frac{I_{max}^2}{Q^2} \left[\sum_{k=1}^m Q^{(k)} \right]^2 = I_{max}^2 ,$$

where we took into consideration that the absolute value of the correlation coefficient R_{kj} cannot be greater than 1:

$$|R_{kj}| \le 1 \quad (k, j = 1, 2, \dots, m).$$

But the following relation

$$I \ge \min (I_1, I_2, \dots, I_m)$$

is not always valid, because the counties can be on the same level in each industrial branch.

Application

To present this method, we chose the counties of Hungary. Our calculations are based upon the number of the industrial employees because on the recent level of our economy it is the industry that can be determinant as far as specialization is concerned.

Naturally — like any other branch in the field of social sciences — the whole complexity of this dynamic category cannot be presented either, having used only one arithmetical method. We approached to the problem only on the level of quantity, so the picture shown by our results is real only if we take into consideration the problems of the quality, too.

Our research was based on the following industrial branches:
Mining
Electrical industry
Metallurgy
Machinery
Vehicle-industry
Electrical engineering
Industry of communication and vacuumtechnical industry
Precision engineering
Metallurgical engineering
Building material industry
Chemical industry
Wood-working industry
Paper-industry
Printing
Textile-industry
Leather-fur-shoe-making
Tailoring industry
Home-industry
Food-industry
Other industrial branches

The index of the rate of the specialization level of the industry can be seen on the following breakdown. We formed specialization levels on the gained indexes and on the basis of all these we introduced the categories
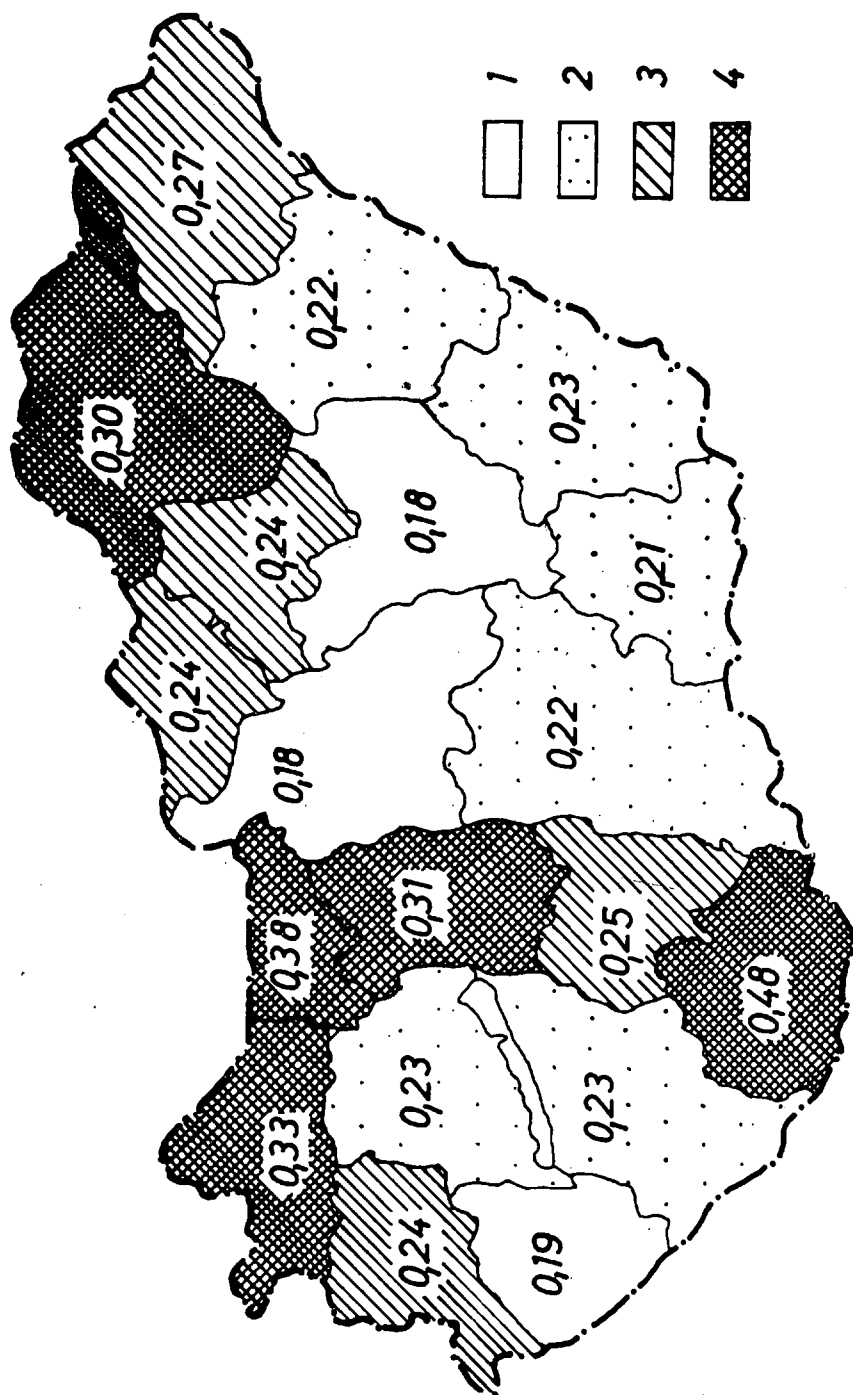
Fig. 1. Specialization of industry in our counties on the basis of the number employed in the industry (1972)
1 = hardly specialized
2 = medium specialized
3 = well specialized
4 = very well specialized

hardly specialized: $I < 0,20$;
medium specialized: $0,20 \leqq I < 0,25$;
well specialized: $0,25 \leqq I < 0,30$;
very well specialized: $0,30 \leqq I$.

Looking at the diagram one can see that in the geographical division of labour the regional differences are great and that they are expressed by the productional profil of the given county. With applying our method, we gave an exact presentation of the relative specialization of the researched regional units. Further researches may come to valuable conclusions approaching to the problem in a similar way.