

**Relationship between the Redundancy and the Concentration
based on Petres' Red**

PETER KOVACS (*University of Szeged, Faculty of Economics and Business Administration, Hungary, pepe@eco.u-szeged.hu*)

TIBOR PETRES (*University of Szeged, Faculty of Economics and Business Administration, Hungary, petres@juris.u-szeged.hu*)

In today's globalizing world decision makers have an increased need for information. The great increase in the quantity of data is not automatically accompanied by an appropriate increase in information. Actually, the problem that decision makers have to face today is not the lack but the abundance of information, but this huge amount of data frequently has only a little information content, which means that redundancy is high. Redundancy means "superfluous" data which do not convey new or noteworthy information in terms of the examination. For this reason the information content of metric data is an essential issue in empirical analyses. This is particularly true for the application of linear regression models. In the case of linear regression models, multicollinearity can be interpreted as a type of redundancy. Petres' Red is a new possible indicator of redundancy and thus of multicollinearity. The Red indicator is defined by using the eigenvalues of the correlation matrix of the explanatory variables. In case of multivariate data analysis the question rightly arises what kind of relationship is between the redundancy and concentration. Concentration refers to the extent to which a small number of units account for a large proportion of total. The Herfindahl-index is one of the measure of concentration.

The aim of this paper is to examine the relationship between the redundancy and concentration in a multivariate linear regression model. Our examination is based on Petres' Red.

We present that the redundancy could be express as the concentration of the eigenvalues of the correlation matrix of the explanatory variables. In this case the values of the normalized Herfindahl-index and the square of Petres' Red are the same.

Keywords: concentration, redundancy, multicollinearity