A classification of peasants attached to land in Byzantium of the 14th century

Khvostova, Xenia V.; Kumekin, Yu. P.

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Historical Social Research, Vol. 16 — 1991 — No. 2, 17-24

A Classification of Peasants Attached to Land in Byzantium of the 14th Century

*Xenia V. Khvostova, Yu. P. Kumekin**

Abstract: Cluster analysis was used to classify peasants attached to land in the 14th century Byzantium on the basis of 1321 property-tax cadastre that registered the landed property of St. Athanasius Lavra in Athos. Peasant households were grouped according to their demographic and economic structures and the taxes they had to pay. We also identified a complex non-single connection between the household groups and their distribution over an administrative unit.

Everybody knows that application of quantitative methods and formalisation in general in medieval socioeconomic history is limited due to a number of factors, the most important of them being incompatibility of data of different medieval sources. Besides, frequently the material is too limited to be quantitatively processed, with inconsistent, random and unclear data. Still, there are situations when socioeconomic information of the Middle Ages can be quantitatively processed. One should not miss such a god-sent opportunity. The Byzantine property-tax cadastres is one of them. (1)

Here we use quantitative methods to study the cadastre registering the landed possessions of St. Athanasius Lavra in Athos compiled in 1321. It includes descriptions of both the domains and the households of dependent peasants found in different regions (katepanykij) of the Thessaloniki Region (theme). The cadastre registered 856 households (the number includes only those data which have been completely explained by the publishers). It supplies information about the number of family members in each of the households, the types of movable property and real estate and the size of tax payed by each family. Our aim is to classify the peasant households according to their demographic, economic and tax characteristics to provide answers to the questions of how far the households differed and whether they formed homogeneous groups with definite economic, demographic and taxing structures. No less important is a problem of possible connection between the economic, demographic and taxing descrip-

* Address all communications to: Xenia V. Khvostova, Academy of Sciences of the USSR, Institute of General History, Leninprospect 32 A, Moscow, 117334, USSR.

tions, on the one hand, and the administrative division of Thessaloniki into katepanykij, on the other. So far, the dominating opinion was that in Byzantium administrative division into regions was little connected with the specific fiscal and economic characteristics of individual households. Naturally enough, our classification takes into account not only the presence of property but also its size and the size of families and the taxes paid.

To reach our aims we used cluster analysis of corresponding data that allowed us to identify groups of households in which the variation of values of considered variables is appreciatively smaller than in other classification groups. To clarify the point let us suppose that each household is described with the help of two variables and it can be represented as a point on the plane. The places of classificational groups are marked by clusters of points the nature of which varies. Some of the points can gather into groups far removed from the others. This is a case of a set of structural heterogeneity. There can be different patterns with the points equally distanced from one another. This allows us to speak of a set of stucturally homogeneous cases without any clusters.

If each peasant household is determined by m-variables cluster analysis should be applied to the m-dimensional space of variables. Here by clusters we mean groups of peasant households computed by a leaf-to-stem hierarchical agglomerate cluster algorithm of data matrix rows, the data matrix columns consist of the variables of households. The data registered in 1321 were used to develop a data matrix for 856 peasant households consisting of 16 columns of variables:

1 - family size;
2 - number of relays;
3-8 - number of donkeys, cows, billocks, goats and swine;
9-11 - number of mills, beehives and fruit trees;
12-15 - sizes of vineyards, orchards, kitchen gardens and arable land;
16 - the size of tax.

The problem of cluster analysis has no unambiguous solution: there are indeterminate forms of four kinds: first, the peasant household variables, i.e. the input data matrix columns, have considerably varied meanings and numerical values of elements depending on the accepted scale of measurements.

We often standardized the data there by renderring all group mean values equal to zero and all group deviation values equal to one. This influences the definition of the distance measure between the two cases before joining.

Second, there are different definitions of distance measures between the two cases before joining. Measures based on either the Euclidean metric or

the Manhattan metric of the multidimensional space of variables are widely used.

Third, the indeterminate form is caused by the choice of the algorithm for joining clusters. The preferred methods are of the nearest-neighbour, the pseudonearest-neighbour and the centroid linkage.

Finally, there is an indeterminate form of analysis of results, that is, the agglomerate algorithms of classification can be used until all the initial objects join into a single whole. Content interpretation of the results obtained at a certain stage of joining clusters points to the clusters that interest the researcher. These indeterminates pose the task of selecting the strategy of research, i.e. selecting the method of classification that would best suit the task posed. Here the selection was determined by a comparison of results of the sample classification obtained through the methods of cluster analysis computed by various algorithms for data of 296 peasant households.

The computations were made on EC-1033 computer (4) by the P2M programme of the statistical packages BMDP of the ABMDP variant. (5) This programme classifies the rows of either the input data matrix or the poststandardization data matrix. The Minkowsky P-th power distance is used as a distance measure, which under P = 1 and P = 2 correspond to the Manhattan and Euclidean metric spaces. Either the L-nearest neighbour method L g 10 or the centroid linkage method that formally corresponds to L = 11 were used as an algorithm for joining clusters. Initially, every row of a matrix (in our case every peasant household) is regarded as a cluster. At any step of the programme two clusters that are separated with the shortest distance are joined to be regarded as a single cluster. This process goes on until all the households are joined into a single whole. Evidently, the overall number of steps will be one less than the number of the rows of the input data matrix.

The main results are printed in the form of a horizontal triangle tree diagram for cluster joining that is used for the concluding content analysis of the classification data. Classification of 296 households was performed by the following algorithms of cluster analysis:

1 - data matrix standardized S, P = 2, L = 11, that is, the Euclidean distances and the centroid algorithm of joining;
2 - data matrix non-standardized NS, P = 2, L = 11;
3 - data matrix standardized S, P = 1,L = 11;
4 - data matrix standardized S, P = 2, L = 1.

We used the data cited in the source that related to households in the Kalamaria katepanykij: the settlements of Gournai, Sarantarea, Pinsson, Loroton, Neohorrion and Karbaioi. Even a visual inspection of the printout of the input data matrix testifies that intensive agriculture (viticulture

and, sometimes, fruit growing) predominated. There were few cows, goats, swine and other non-draft animals, or none at all. The greater part of the households possessed relays. The preliminary inspection allows to classify the larger part of the households in the aboveniamed settlements as poor or average rich, with the average family sizes who payed an average or low tax.

Further classification made these preliminary conclusions more precise and detailed. We compared the results obtained by different algorithms to determine the nature of the input data, on the one hand. Our aim was to demonstrate that these households were a set of structurally heterogeneous cases according to sixteen variables, that is, that they form compact clusters of cases. On the other hand, this comparison has identified the algorithms of cluster analysis that can be successfully used to classify the entire data file for 856 households. In the presence of considerable structural heterogeneity with the clear-cut differentiation of households into groups the results of classification by different algorithms should be similar. Poorly manifested structural heterogeneity of data signifies that inevitably individual households shifted from one classificational group into another. In this case a choice of the optimal classification method is a challenge. Classification results of the file demonstrated that the most aggregative cluster structure results from classification by the first algorithm (the standardized data matrix, the Euclidean distances and the centroid method of joining). The use of other algorithms lead to a more fragmented structure of clusters obtained by the first algorithm.

This explains why we have selected the computation by the first algorithm as the strategic line of research into classification of the complete data file for 856 peasant households. Control computations were performed by the second algorithm, that is, without standardizing the input data.

The horizontal triangle tree diagram for cluster joining obtained through computations looks like a rectangular isoscales triangle with the vertex of the right angle being in the left upper corner of the printing. The triangle's vertical side is formed by 856 rows, the first of them corresponding to the household that is the first in the input data matrix. The rest of the rows correspond to the remaining households shifted to form the tree diagram. There are the household's ordinal number in the input data matrix, its name and tag (the name coincides with the number while the tag corresponds to any given household's territorial affiliation). Every row (except the first one) gives the ordinal number of the programme step at which the cluster with the given household was formed. The last row corresponds to the final, 855th, step of the programme. The horizontal side of the triangle of the tree diagram is printed as a line of 856 dashes. The hypotenuse is printed like a slanting line formed by strokes ascending from the printing's last row and corresponds to the merging of all households into a single whole.

The structure of cluster joining is represented inside the triangle with horizontal lines made of dashes and slanting lines made of strokes. All the households along the triangle's vertical side enclosed within the limits that corresponds to leftward movement away from any intersection of the horizontal and slanting lines belong to one and the same cluster.

In this way, the specific pattern of printing of the tree diagram for cluster joining is determined by the distribution of households on the input data matrix. It's easily allows to recognize the needed clusters in the set of structurally heterogeneous cases.

We do not show here the corresponding printing because of its size and because it should be compressed into a more integrated form. Below there are descriptions of the clusters that emerged as a result of integration of the full horizontal tree diagram.

Cluster I consists of 81 households. Joined with Cluster II at the 430th programme step.

Cluster II consists of 60 households.

Cluster III consists of 70 households. Joined with Clusters I and II at the 446th-560th programme steps.

Cluster IV consists of 46 cases; it joins the tree diagram at the 561st step.

Cluster V consists of 14 cases; it joins the tree diagram at the 563rd -619th steps.

Cluster VI consists of 61 cases; it joins the tree diagram at the 645th step.

Cluster VII consists of 45 cases; it joins the tree diagram at the 660th - 671st steps.

Cluster VIII consists of 107 cases; it joins the tree diagram at the 680th step.

Cluster IX consists of 128 cases; it joins the tree diagram at the 688th step. One individual household joins the tree diagram at the 705th step.

Cluster X consists of 26 cases; it joins the tree diagram at the 710th and 718th steps.

Cluster XI consists of 84 cases; it joins the tree diagram at the 734th - 799th steps.

Cluster XII consists of 28 cases; it joins the tree diagram at the 800th step.

Cluster XIII consists of 32 cases; it joins the tree diagram at the 801st - 823rd steps.

Cluster XIV consists of 69 cases; it joins the tree diagram at the 827th - 851st steps.

Four individual households join the tree diagram at the 852nd, 853rd, 854th and 855th steps.

Let us now look at how the results of cluster grouping can be interpreted. What are the clusters* economic and demographic descriptions and how do they correspond to the administrative division?

Clusters I and II consists of 81 and 60 poor households, respectively. Cluster I is made up of the families of one or two; some of them owned a vineyard of not more than one modij. In Kalamaria there were 50 households, in Ierisson 17 households, in Akros 5 households, in Hermeleia 3 and in Kassandra one household. They paid taxes that never went higher than 0.5 perpers (the Byzantine monetary unit). Cluster II includes the households with families of 3 to 4 members. Many of them had small vineyards and paid taxes not larger than 0.67 perpers. In Kalamaria there were 31 households, in Ierisson 15, in Akros and Revenikeia 6 in each, in Hermeleia and Kassandra 1 in each.

Cluster III consists of 70 peasant households of average wellbeing that concentrated on different things: some of them tilled vineyards, others raised cows, still other had orchards or just fruit trees. The families were of one to three members. There were 43 households in Kalamaria, 15 in Ierisson, 7 in Revenikeia, 4 in Akros and 1 in Hermeleia.

Cluster IV is formed by poor large families (from four to seven members). It consisted of 46 households who paid the tax of 0.33 to 1 perper. They tilled a small vineyard or, sometimes, an orchard, and had cows. There were 23 households in Kalamaria, 14 in Ierisson, 3 in Revenikeia and in Akros, Hermeleia and Kassandra 2 in each.

Cluster V consists of 14 households, 5 of them in Ierisson and Revenikeia, Akros and Kassandra 3 in each. These were the households who were battling at the brink of poverty and average welfare: they had small vineyards and some of them tilled a small patch of arable land, the families consisted of 3 to 4 members and they paid the tax of 0.33 to 1.33 perpers.

Cluster VI consists of 61 households of average richness with cows and vineyards, they paid no more than 1 perper of taxes. The family size ranged from one to five members; 38 out of 61 households were found in Kalamaria, 12 in Revenikeia, 10 in Ierisson and 1 in Hermeleia.

Cluster VII includes 45 households of average richness; it consisted of four economically homogeneous groups the structure-forming elements being vineyards or orchards or goats. Their taxes never rose higher than 2 perpers. There were 18 households in Ierisson, 15 in Revenikeia, 8 in Kalamaria and 4 in Akros.

Cluster VIII consists of 107 households of average richness with vineyards, bullocks, cows and orchards determining their economic structure. They paid up to 2 perpers, the families consisted of 8 members. There were 66 households in Kalamaria, 16 in Ierisson, 6 in Akros, 4 in Hermeleia and 2 in Kassandra.

Cluster IX consists of 128 households of average richness, their economic backbones being relays and vineyards, some of them also owned donkeys. All of them paid not more than 2.83 perpers of taxes. There were 124 households in Kalamaria, 2 in Ierisson and in Revenikeia and Kassandra 1 in each.

Cluster X includes 26 average households with families of 1 to 5, the economic structure being determined by vineyards, partial ownership of mills and orchards. They paid no more than 1.67 perpers. There were 11 households in Ierisson, 9 in Kalamaria and 6 in Revenikeia.

Cluster XI includes 84 rich households with the families of 1 to 8 members. They had polycultural economies with bullocks, donkeys, vineyards, orchards, goats, cows, fruit trees and arable land. Their taxes went as high as 3.5 perpers. There were 18 households in Kalamaria, 14 in Revenikeia, 7 in Akros and 4 in Kassandra.

Cluster XII contains 44 households, very rich and with a developed polyculture. The peasants owned relays,vineyards, cows, goats and arable land. Taxes were as high as 10.33 perpers. The family size ranged from 2 to 7. There were 23 households in Kalamaria and 5 in Ierisson.

Cluster XIII includes 32 households equally rich and polycultural. As distinct from the previous cluster there were practically no relays. The peasants owned donkeys, vineyards, orchards, sheep and fruit trees. The families were from 1 to 6, they paid not more than 3 perpers. Nineteen households were in Ierisson, 7 in Kalamaria, 5 in Revenikeia and 1 in Akros.

Cluster XIV embraces 69 richest landowners with polycultural households with the families of up to 10 and the taxes of up to 8 perpers. There were 34 households in Ierisson, 22 in Kalamaria, 8 in Revenikeia, 3 in Kassandra and in Akros and Hermeleia one in each.

Four households that came at the bottom of the classification diagram were in Revenikeia and belonged to very rich owners.

It is evident that in the majority of cases the households that formed the clusters were unevenly distributed in individual administrative units. The greater part of households belonged to one or two katepanykijs of Thessaloniki; in other words, there was a certain connection between the economic structure of peasant households and their administrative affiliation. In this way, administrative division was determined by economic factors together with territorial features - the economic and territorial factors were closely interwoven and should be regarded as an integral whole.

Notes

1. See X.V. Khvostova, »Kolichestvenny podkhod v srednevekovoi sotsialnoekonomicheskoi istorii«, Moscow, 1980; X.V. Khvostova, »Kolichestvennye metody pri izuchenii srednevekovykh sotsialno-ekonomicheskikh yavlenii« in: Ekonomicheskaya istoria. Problemy i issledovania, Moscow, Nauka Publishers, 1987; K.V. Chvostova, »Die Theorie der sozialökonomischen Differenzierung feudalabhän-

giger Bauern und das Problem ihrer teilweisen Formalisierung«, in: Wirtschaftsgeschichte und Mathematik, Akademie Verlag, Berlin, 1985.

2. »Actes de Lavra«, Vol.II, Ed. par A. Guillou, P. Lemerle, N. Svoronos, D. Papachryssnthou, Centre national de la Recherche Scientifique, Paris, 1979.
3. S.A.Aivazyan, V.M.Bukhshtaber, I.S.Enyukov, L.D.Meshalkin, »Prikladnaya statistika. Klassifikatsia i snizhenie razmernosti«, Finansy i statistika Publishers, Moscow, 1989.
4. The authors wish to express gratitude to research associates at the Institute of Sociological Studies of the USSR Academy of Sciences whose cooperation made our computations possible.
5. Dixon W.J. (ed.) »BMDP. Biomedical Computer Programs«, Berkeley, Los Angeles: Univ. Cal. Press, 1975.