

□

2. Besides the practical aspects, the joint utilisation of the two data sets was useful because none of the independent programmes covered the areas and the information which had been dealt in detail by the other. It is important to mention that both censuses were full-scale operations, as a consequence discrete data are available. The latter facilitates the aggregation of the data according to any territorial unit or other criteria. Today's highly developed IT infrastructure permits the aggregation of the data of the two data collections.²

3. By means of the databases of the two censuses, the first opportunity arose for matching the discrete data of the surveys. This is a new development not only in the history of the Hungarian censuses but also in international practices. Besides the impetus that was given by the access of the country into the European Community, an opportunity was given for the description of e.g. the demographic and housing characteristics of the population working in agricultural private holdings. The results of this exercise, with special regard to the territorial presentation, serve as a good platform for future surveys.

4. It was clarified already at the beginning that the precondition of the matching of the data of the two censuses is the conformity of the respective metadata of the two operations. The aim was to approximate the different concepts referring to similar groupings that have been applied in the two censuses. It was found that "private holding" and the "household" are the categories that can be used as the smallest unit for the matching. Consequently, the experts decided to utilise the following categories: the private holdings³ from the agricultural census and the households, more exactly the dwelling-households.⁴ There was a general agreement that, although the two censuses cover different aggregations, the categories used more or less refer to the same population groups. Moreover, thanks to the high frequency of the categories in the two censuses, there is a rather large common mass and only a limited amount is represented only in one of the censuses.

5. The precondition of creating a common database was the exact matching of the elements of the two surveys following the separation of the data with the identifiers defined above. The further analysis of the common database was based on the elements (private holdings and households respectively) marked with the same identifier. A suitable tool for the matching of the elements was the address of the enumeration.

6. The Law XLVI of 1993 on statistics serves as legal background of matching the two databases.

7. The aim of the project was to build up a common database allowing to restrict elementary information from both censuses. The content of the database had been fixed upon the requirements of the planned studies. By using the linked database the Hungarian Statistical Office issued a series of publications on the population living in agricultural private holdings in the countryside (i.e. outside of Budapest).

I. MATCHING METHOD

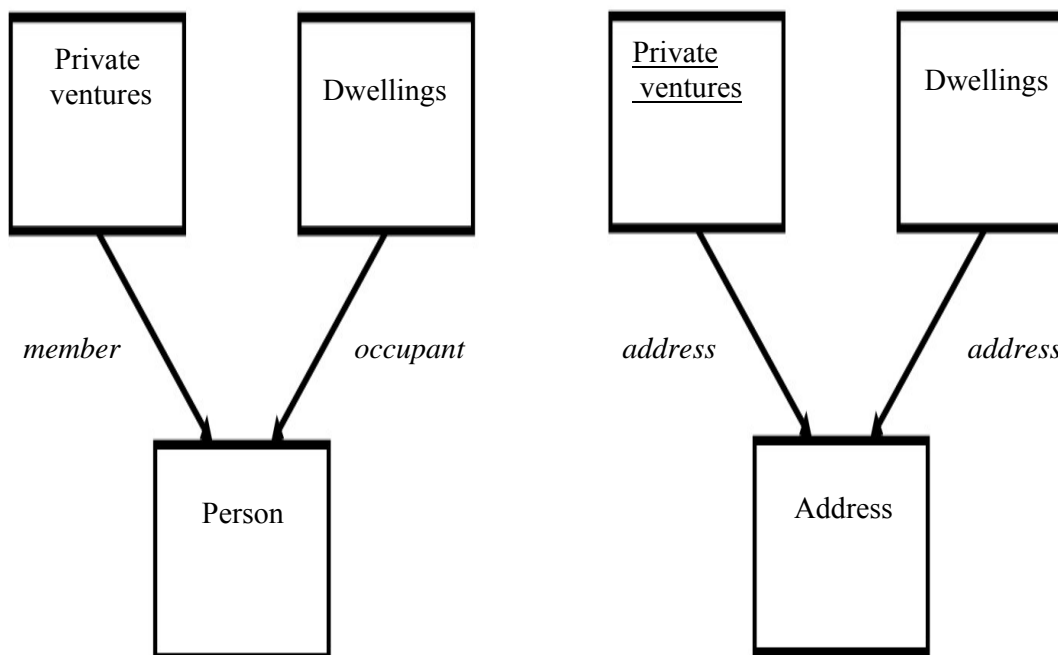
8. The aim of matching the two databases was to study the relationship between the categories "private holding" and "household" used by the agricultural census and the population census,

□

respectively. The correspondence is made via overlapping persons, partly via the same addresses of the given categories.

Matching based on the person

Matching based on the common address



9. The units of observation in the AC were agricultural private holdings, and dwellings and persons in the population and housing census. The AC recorded some information (e.g. gender, age) also on persons other than employees working in the private holding. Based on this, it was possible to estimate a link between members of the private holding and occupants of the dwelling. The only common base of connecting the given units was the address. The links between the private holding and the household can be based on the identity of the persons living in the dwelling. The matching of the databases requires the use of the individual identity codes. The individual identity code is a complex parameter which gives an exact match of the elements (characteristics) of a set. The links are represented by the couples formed by the identical codes of the occurrences.

10. Both censuses used individually defined complex codes for the identification of the occurrences (where, as due to the protection of personal data the PIN code could not be used, only the identity number of the locality could be the same). In defining the primary links the pairs

□

of the individual identifiers of the members of the private holdings and the occupants of the dwelling are used:

- (a) locality – AC enumeration district – serial number of the private holding – serial number of the person in the private holding;
- (b) locality – enumeration district – address number of the private holding – serial number of the person in the dwelling.

11. It has been anticipated that the persons belonging to the private holdings and the persons belonging to the population census might be identified on the basis of the common address and of some demographic characteristics.⁵

12. The identification code of the member of a private holding contains the identification number of the given private holding, consequently the common part of the identification code gives information about the relation. The situation is the same in case of the inhabitant and the inhabitant of dwelling. Taking into account that the two censuses applied unique identifiers the production of the relation table (incorporating the pairs defined by the method described) was a logical exercise. The fixing of the interrelations was based on the assumption that the persons being members of the private holdings enumerated by the AC, as well as the members of the households of the population census, will be clearly identified on the basis of the address and some demographic characteristics.

13. The steps of the matching process were the following:

- (a) nearly 90 percent of the 1 million agricultural private holdings have been identified by using a special programme for matching the elements of the two databases of the addresses;
- (b) the remaining, “unmatched” AC addresses were recognised with the help of an interactive programme which resulted in a further 7 percent of matched addresses; the process lasted for around 10 months;
- (c) the identification of the persons was performed later on, by a different programme: 78 percent of the members of agricultural private holdings could be “paired” with persons in the population census.

A. Problems with the identification of the addresses

14. All of the statistical data collections include the addresses of the data source. Consequently the data gatherings of the HCSO are linked by this method. Nevertheless, the identification and matching of the addresses is hampered by different factors such as: the variances in spelling of the addresses, incorrect indication (only the house number is marked), modifications of the addresses (renaming, application of different house numbers, etc.) during the two different data collections.

15. In matching the databases, a frequent problem is the difference in applying the names of the public areas i.e. using the full name Sándor Petőfi, only S. Petőfi, or simply Petőfi. A similar problem is that the buildings or dwellings are not marked uniformly, they might be labeled in different forms i.e. 1/a, 1/A, 1a, 1A, etc. Important precondition of the successful recognition process is the availability of a register of addresses with aligned common content. The addresses of the AC and the population census are constructed with a similar structure:

□

(a) locality – name of public place –character of the public place – address number – building – entrance – floor – door.

16. The process of recognition is rather difficult in case the real estates do not have proper physical identifiers, only further information on a person living on the address.

B. Identification of the agricultural private holdings and the members of the private holdings

17. In the process of the preliminary matching, as mentioned above, it was possible to identify 90 percent of the addresses with the help of a computer program. The high rate of coincidences is hiding a rather high variance. The aim of the interactive matching of the addresses was the verification of the barely defined areas. Finally, 97 percent of the addresses of the AC could be disclosed. The previous procedure was followed by clearing the persons; the identification of an address was regarded as successful provided the majority of the persons could be matched, i.e. identified.

18. The AC recorded only limited information on the persons other than employees taking part in the activity of the private holding. Nevertheless the respective information could be used for matching with the identical data of the population census. The most suitable personal data for the identification are the gender, age and highest education of a person.

19. Some remarks in regard to the above identifiers are the following: in case of the gender only a perfect match can be accepted. The AC recorded the age of the person while the population census enumerated the date of birth. Presumably the former is less correct. The age 60 recorded might refer the completed 60 years of age, or to a person aged over 59 years. Taking into account that the AC has been executed 10 months prior to the population census some deviations in the age had been accepted. The level of education was registered in the AC in a rather simplified way (without education, primary education attended, secondary education completed, higher education obtained) while the population census applied a detailed description convertible into the former categories. The deeper analysis has shown that due to the deficient data recording, the information of the AC are less appropriate for the recognition. This information was used only as supplementary criteria.

20. As mentioned above, after 97 percent matching of the addresses, only 78 percent of the agricultural private holdings could be declared as identical, i.e. the same in the two censuses. The difference is caused by effective changes (moving, deaths etc.) on the one hand, and by the use of different identifiers of the addresses, on the other. This shows that the address of a person is not as strong attribute of a person as gender and age.

II. SOME EXAMPLES OF SOCIO-DEMOGRAPHIC CHARACTERISTICS OF THE POPULATION LIVING IN PRIVATE AGRICULTURAL HOLDINGS

21. The aggregated database enables us to analyse the demographic characteristics, the family relations and housing conditions of the population living in private agricultural holdings, taking into account some elements of the agricultural census, such as the size of the cultivated land of a holding, the type of farming, the size of the agricultural output, the number of livestock, etc. Additionally, the data of the two censuses give a possibility for analysing the social stratification

□

of the population living in private agricultural holdings relying on the stratification scheme elaborated in the course of the processing of the 2001 census and improved after consultations with the experts of the subject. The following part of the paper describes some of the results of the analysis.

22. The population census 2001 enumerated 3 million 863 thousand households, while in the course of the AC 2000, 2.1 million households were visited. Only 4400 of the 770 thousand households in Budapest were subject to the AC. Consequently, in matching and analysing the two censuses, the data on Budapest have not been taken into account. As the enumerators had to visit only the households engaged in significant agricultural activity in the downtown areas of the cities, the evaluation covered only two third of the 3 million 93 thousand households living outside Budapest. 960 thousand of enumerated households proved to be agricultural holdings (its agricultural activity exceeding a threshold applied in Hungary), 835 thousand households performed agricultural activity but below the threshold, while 300 thousand households had no agricultural activity. 784 thousand (82 percent) of the 960 thousand private agricultural holdings could be linked and compared with the households enumerated by the population census. Excellent possibility for conducting a detailed analysis was given by the fact that there have been only slight deviations regarding the former coincidences by counties, sub-regions.

23. In 2001, the population living in the countryside (i.e. outside Budapest, the capital city) accounted to 8 million 420 thousand, of which the number of persons living in private agricultural holdings reached 27 percent (2 million 265 thousand persons).

24. It is valid for the country as a whole that the changes in the age structure show an ageing of a larger part of the population. The trend is similar in case of rural population: while the share of the population 60 years and over in 1960 was 14 percent, in 2001 it was 20 percent. The process of ageing was more rapid in the case of women; there was an 8-percentage point increase (from 15 to 23 percent) in the share of those 60 and older.

25. 23 percent of the population living in agricultural private holdings was at least 60 years old, the value exceeded by 3 percent the similar indicator characterising the population living in the countryside; the difference among those being 40 years and over shows the same tendency. The share of the population living in the countryside aged 40 years and over accounted to 47.7 percent, in the case of the agricultural population it was 53.4 percent. Obviously this relates to the fact that the majority of the persons in the agricultural private holdings live in the villages, where the share of the elderly population is higher than average. Furthermore the members of this age bracket are attached to the agriculture, where they can work – depending on the state of health - even when unemployed or retired.

26. The age structure of those living in agricultural private holdings producing for the market is more favourable. Within the latter group the share of those being 40-59 years of age, having key role in the production and sale of the agricultural products, is the highest. Taking the age structure by the types of agricultural private holdings it can be stated that it is the most favourable in case of the private holdings producing for the market. In the agricultural private holdings dealing with animal husbandry the younger population (15-39 years of age) is dominant. It is easy to understand that the share of the elderly population (60 years and over) is the highest in agricultural private holdings producing for self-consumption, while the share of

□

this age group in case of the agricultural private holdings producing crops was nearly 25 percent.

Population living in the countryside and population living in agricultural private holdings, by age groups, type and purpose of production

Percentages

Population	Total	-14	15-39	40-59	60-X
		years old, percentages			
Living in the countryside	100.0	17.4	34.9	27.8	19.9
Living in agricultural private holdings	100.0	15.6	31.0	30.1	23.3
Of which living in:					
Private agricultural holdings in crop production	100.0	14.9	30.2	30.0	24.9
Agricultural private holdings in animal husbandry	100.0	18.3	32.9	28.4	20.4
Agricultural private holdings in mixed (crops, animal husbandry) production	100.0	14.8	30.6	31.1	23.5
Agricultural private holdings producing for self-consumption	100.0	16.0	30.7	28.9	24.4
Agricultural private holdings selling the products exceeding self-consumption	100.0	14.8	30.6	31.5	23.1
Agricultural private holdings producing for the market	100.0	16.5	34.0	33.7	15.9

27. In the past decades the level of education of the population living in the countryside improved continuously: there was a decline of those passing the school-ages without having completed the 8th grade of primary education, while the level of education increased in every age group.

28. Although there was an improvement in the educational level of the population living in agricultural private holdings, due to the unfavorable age structure, their level of education is lower than that of the population living in the countryside on the average. This is most visible with those who have completed the 8th grade of primary education. The share of persons completing only the 8th grade of primary education was 50 percent in case of the population living in the countryside together, while the index had shown 56 percent in case of the agrarian population. In the latter group the share of those completing less than the 8th grade of primary education is higher (80 percent) than within the population living in the countryside on average (72 percent).

29. In evaluating the overall educational level of the persons aged 7 years and over in the different types of agricultural private holdings, it has been found that the educational level of the agrarian population is the highest in the agricultural private holdings producing crops, followed by those living in private holdings with mixed production, while the lowest educational level is

□

detected among those working/living in agricultural private holdings producing animal husbandry. The differences are most probably resulted from the privatisation of the agricultural land (in the last decade of the previous century), when a large portion of land had been privatised by the younger, more educated farmers. The latter group of persons used the land almost exclusively for crops production, while some of them were also involved in animal husbandry as a supplementary activity.

30. By approaching the educational level of the agrarian population from the angle of the purpose of production of the private holdings it was detected that the people living in agricultural private holdings producing for the market are more educated than those living in agricultural farms producing for their own consumption. The difference in educational level is especially significant in case of those completing the third level education: while 7 percent of the population aged 7 years and over living in farms producing for the market completed the university or higher educational institution, the same indicator in agricultural private holdings producing for self-consumption was only around 3 percent.

31. The economic activity structure of the population living in agricultural private holdings significantly diverges from that of the population living in the countryside in general: the share of persons in employment is lower, the share of inactive earners is higher. The differences in economic activity structure relate to fact that the elderly people living in agricultural private holdings have a higher percentage share in the total agrarian population than the elderly people's share in the population living in the countryside in general. Significant differences are also shown according to the purpose of production: the economic activity rate is far lower in the agricultural private holdings producing for self-consumption than in private holdings producing for the market, where 47 percent of the male population is in employment. (Another cross-tabulation on population living in the countryside and in agricultural private holdings by aggregated major groups of occupations, aggregated major groups of industry, type and purpose of production is presented in Annex 1).

32. There is a decrease in the difference between the level of equipment of urban and rural dwellings. Besides piped water networks, a wider spread of sewerage allow the use of modern, environment-friendly systems. In rural areas sewage and gas supply are more frequently provided individually, while in towns the construction and use of public utilities have a more important role. Accordingly, in dwellings occupied by agrarian population, the lack of community networks had to be compensated by individual domestic water conduit and sewer. This resulted in a large majority of housing units provided with flush toilet, equipped bathroom and hot water supply. The data relating to housing quality show that market-oriented holdings are in the best situation: here the share of dwellings provided with all modern conveniences (full comfort) is 61 percent, while that of dwellings of private holdings producing for self-consumption only is 39 percent.

□

Population living in the countryside and population living in agricultural private holdings by highest education, type and purpose of production

Percentages

Population	Total	Less than the 8th grade of primary school	8th grade of primary school	Secondary school with final exami- nation	Third- level (university , college etc.) education
		completed as a percentage of total population			
Living in the countryside (outside Budapest)	100.0	21.5	50.0	20.6	7.9
Living in agricultural private holdings	100.0	23.1	56.4	16.1	4.4
Of which:					
Agricultural private holdings in crops production	100.0	22.3	49.1	20.8	7.7
Agricultural private holdings in animal husbandry	100.0	25.7	60.0	12.1	2.2
Agricultural private holdings in mixed (crops, animal husbandry) production	100.0	22.3	60.5	14.3	2.9
Agricultural private holdings producing for self-consumption	100.0	24.9	57.7	14.1	3.3
Agricultural private holdings selling the products exceeding self-consumption	100.0	21.3	56.3	17.6	4.8
Agricultural private holdings producing for the market	100.0	17.5	48.0	24.2	10.2

33. The data of the 2001 population and housing census has also allowed for the elaboration of different models of social stratification. (In the model containing the total population, children under 15 were classified according to their parents' situation.)

Population living in agricultural private holdings by economic activity, purpose and type of production

Percentages

Economic activity	Total	Producing for			Producing crops	Producing animal husbandry	Mixed production
		self-consumption	self-consumption and market sale	market sale			
Economically active population							
Person in employment	32.9	31.1	34.0	41.1	34.8	30.7	32.5
Unemployed	4.3	4.7	3.9	3.1	3.7	5.4	4.2
Economically active population together	37.2	35.8	37.9	44.2	38.5	36.1	36.7
Economically non-active population							
Receiving childcare allowance	2.7	2.8	2.4	2.5	2.4	3.3	2.5
Pensioner, rentier on own right	21.9	23.0	21.7	15.5	23.2	19.7	22.0
Disability pensioner, rentier	7.6	7.5	8.1	6.0	6.4	7.9	8.5
Pensioner, rentier on derivative right	2.3	2.6	2.2	1.3	2.8	1.9	2.2
Other inactive earner	2.9	2.9	2.9	2.7	2.4	3.1	3.2
Inactive earner together	37.4	38.8	37.3	27.9	37.1	35.9	38.4
Dependent together	25.4	25.4	24.8	27.9	24.4	28.0	25.0
Economically non-active population together	62.8	64.2	62.1	55.8	61.5	63.9	63.3
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0

34. In this respect, the data show that Budapest has a privileged situation: while 10 percent of persons aged 15–74 living in the capital belong to the strata of large employers, higher managerial occupations and higher professionals, within the agrarian population the respective share is hardly two percent. On the other extreme, social strata with lower status are far more populous within the agrarian population, e.g. in this category the share of semi-routine occupations is three times as high as in the capital (30 and 10 percent, respectively). The social composition of the agrarian population is unfavorable as compared to those living in the countryside except the agrarians: the total share of persons belonging to the first three categories is 9 and 16 percent, respectively. Taking into account types of locality, the data show an unfavorable situation of the agrarian population living in villages (rural areas) as well, though there is no big difference in the share of persons seceding from the labour-market of the different groups within the population living in the countryside (see Annex 2.)

□

35. The related databases of the AC and the 2001 population census allow for analysing the social structure of the agricultural population by type and purpose of production of agricultural private holdings. Within the agrarian population aged 15–74, 35 percent were living in crops producing, 22 percent in animal keeping and 43 percent in mixed agricultural private holdings. As far as the purpose of the production is concerned, 58 percent of the agrarian population aged 15–74 belonged to agricultural private holdings producing for self-consumption, one third to private holdings producing for both self-consumption and market sale, and only 9 percent were living on farms producing purely for the market. Pending on the type of agricultural production, the social stratification of the population living in agricultural private holdings is highly differentiated.

36. Taking a couple of variables from the extremely rich information of the joint database of the two censuses, we may present some examples of the enlarged possibilities for studying the socio-demographic characteristics of the agrarian population. For instance, starting from the census data on knowledge of languages, it turns out that only 9.8 percent of the agrarian population aged 15–74 speaks some foreign language, against a national average of 21.3 percent, and a countryside indicator of 17.5. At the same time, the relevant data in the case of agricultural private holdings producing for the market – in line with the higher cultural level of this subgroup – show 15 percent.

37. The 2001 population census also observed the daily commuting of the population in employment. 43 percent of persons in employment living on private holdings commuted daily between their home and the locality of their workplace. (The respective national average was 30 percent, that of rural population 35 percent.) Here too, the type and purpose of production are the differentiating elements. For instance, within the employees belonging to the agrarian population of holdings producing for self-consumption, the share of commuters is 49 percent, in those selling the surplus this share is only 38 percent, and in market-oriented ones 29 percent.

38. According to the data on households of census 2001, the share of households containing lone parent families within private holdings, was 7.8 percent (as against 10.3 as rural average), while 23 percent is the share of households with employed members only. The respective values for the same indicators are 6.3 and 37 percent in the case of holdings with market-oriented production. As a result of the full-scope character of the two censuses, the aforementioned aspects may be analysed in any geographic breakdown.

ANNEX 1

Population living in the countryside and population living in agricultural private holdings by aggregated major groups of occupations, aggregated major groups of industry, type and purpose of production

Category	Total	Aggregated major groups of occupations						Industry		
		leading intellectuals	professionals, associate professionals, clerks	services workers	skilled agricultural and forestry workers	mining, manufacturing, construction workers	other	agriculture, forestry	mining, manufacturing, construction	services
Total population	100.0	17.8	18.5	15.8	3.8	34.7	9.4	6.8	35.8	57.4
Persons living on private holdings	100.0	11.9	14.9	14.6	8.9	39.4	10.3	15.4	35.8	48.8
Of which:										
Agricultural private holdings in crops production	100.0	18.3	18.4	14.9	5.1	35.1	8.2	10.3	34.7	54.9
Agricultural private holdings in animal husbandry	100.0	7.1	12.6	15.0	6.3	45.9	13.1	11.8	41.9	46.3
Agricultural private holdings in mixed (crops, animal husbandry) production	100.0	8.6	13.0	14.2	13.4	40.0	10.8	21.7	33.8	44.6
Agricultural private holdings producing for self-consumption	100.0	10.0	14.4	14.9	4.8	44.2	11.7	9.8	41.1	49.2
Agricultural private holdings selling the products exceeding self-consumption	100.0	12.6	15.4	14.7	12.0	35.8	9.4	20.4	30.8	48.8
Agricultural private holdings producing for the market	100.0	19.5	16.2	12.8	20.3	25.3	5.9	29.1	24.2	46.7

