

## On Family Budget Survey in Russia<sup>(1)</sup>

Kiyoshi Takeuchi

1. The present paper considers a historical development of family budget survey in Russia, focussing on the sampling methods and the way of interpreting survey results. It follows previous papers of the same author.<sup>(2)(3)</sup>

2. Let us first consider the sampling methods applied in the field of family budget survey in Russia, where Zemstvo statistics (Земская статистика)<sup>(4)</sup> played an important role, even though they had several serious drawbacks, as will be seen later. There had been several surveys before Zemstvo statistics were to apply certain sampling methods. For example, we have family budget surveys of peasant households carried out in the 1870's by E. N. Anuchin (Е. Н. АНУЧИН) in Samara (Самара) Prefecture, by T. I. Osadch (Т. И. ОСАДЧ) in Kherson (Херсон) Prefecture, by A. A. Rusov (А. А. РУСОВ) in Chernigov (Чернигов) Prefecture, by A. V. Karпов (А. В. КАРПОВ) in Nizhegorod (Нижегород) Prefecture, and by G. Manokhin (Г. МАНОХИН) in Perm (Пермь) Prefecture. Sizes of those surveys were not so large, but it is said that they were carried out rather well.<sup>(5)</sup>

---

(1) The author is much grateful to Miss H. Murakami (presently Mrs. Okamura) for her sincere assistance in completing his present paper.

(2) Kiyoshi Takeuchi, "On family budget surveys in Russia," *The Economic Review*, Vol. 19, No. 2, 1968, pp. 173-188.

(3) Kiyoshi Takeuchi, "On H. C. Четвериков's sampling survey theory," *The Keizai Gaku (Annual Report of the Economic Society, Tohoku University)*, Vol. 33, No. 1, 1971, pp. 111-116.

(4) Zemstvo statistics were statistics of the survey of economy and the other fields of social life, carried out by prefectural Zemstvo (elective district council) (sometimes by county Zemstvo) from the beginning of the 1870's to 1917. See, *Statisticheskii slovarj (Статистический словарь)*, 1965, pp. 157-158.

(5) I. Ya. Machukha, *Statistika byudzheta naseleniya (И. Я. Матюха, Статистика бюджет населения)*, 1967, p. 154.

Those surveys stemmed from the needs of detail investigation in case of available means and resources being limited, and the method of partial statistical observation came to be utilized. That is, in those surveys the idea and method of sampling came to be considered in a vague way.<sup>(6)</sup>

Only such family budget survey data as were carried out at the end of the 19th century should be used as the object of statistical analyses. In those surveys rather excellent sampling methods were applied as will be seen later, but almost of all them were subject to purposive sampling method. However, those trials of partial observation can be considered as a first step to scientific sampling survey developed later.

In this note we will consider the methods of typical sampling, proportional sampling, systematic sampling, and multistage sampling in Russia, connected with the methods of collecting family budget data.

3. In Voronezh (Воронеж) Prefecture family budget surveys of 24 peasant households, 67 households, and 230 households were carried out in 1885, in 1886-87, and in 1887-96, respectively. In those cases, statisticians were not subject to certain definite sampling procedures. Peasants were not classified by social form such as rich, middle, and poor peasant, but by the size of allotment.

Such a classification as a sampling criterion was not valid under the social conditions at that time.<sup>(7)</sup> Furthermore, in many cases such a peasant household was first sampled whose head could give informations on their own income and expenditures.<sup>(8)</sup> As a result those

---

(6) V. Krylov, "O primenenii vyborochnogo metoda v zemskoi statistike," *Vestnik statistiki*, (В. Крылов, "О применении выборочного метода в земской статистике," *Вестник статистики*), No. 6, 1955, pp. 54-56.

(7) Property of allotment belonged to landlord, and peasants were obliged to pay certain costs to landlords. Therefore, poor peasants who did not have capacity of production came to give up their allotments, and rich peasants rent them. In view of such a condition, the size of allotment was a very vague sampling criterion for «typical» peasant.

See, for example, M. Kikuchi, *Roshia Nodo Kaiho no Kenkyu*, 1964, p. 420.

(8) I. Ya. Machukha, *op. cit.*, p. 158.

samples were biased, and data obtained had higher proportion of rich peasants than whole data based on peasantry census.

In the survey carried out in Vologda (Вологда) Prefecture in 1903-1911, it was tried to sample «typical» peasant households based on peasantry census. It was due to the considerations that the sample be representative of various types of peasants in the whole group, that is, population, and sample result was tested with respect to important economic characteristics compared with the population. As an example, let us see the test data on 91 peasants sampled with respect to seeding index of the peasants in Totem (Тотем) County in 1907.<sup>(9)</sup>

**Table I**  
Comparison of seeding peasant groups between  
peasantry census and budget monograph

Percentage of the peasants falling into each seeding group (%)	Seeding peasant group <sup>(10)</sup>						Total of seeding peasants (%)
	~1.0	1.1~2.0	2.1~3.0	3.1~4.0	4.1~6.0	6.1~	
	dessiatina	dessiatina	dessiatina	dessiatina	dessiatina	dessiatina	
	II	III	IV	V	VI	VII	
By whole peasantsy census	3.4	18.7	26.0	17.4	18.5	16.0	100.0
	22.1		43.4		34.5		
By budget monograph	2.4	22.6	17.9	20.2	17.9	19.0	100.0
	25.0		38.1		36.9		

A statistician in Vologda Prefecture described, based on this result, that a difference shown was not so large as to be recognized significant. We can support his conclusion by way of, for example, the method of  $\chi^2$ -test. In those days an academic level of sampling theory and methods was not so advanced that the survey results were not analyzed in stochastic way.

Among statical surveys by typical sampling is very interesting the family budget survey of peasant households in Tambov (Тамбов) Pre-

(9) I. Ya. Machukha, *op. cit.*, p. 161.

(10) 1 dessiatina=2.7 acres.

fecture, carried out for two years from the spring of 1914 to the spring of 1916. They divided the prefecture into 3 districts, and within them homogeneous county, volostj (волостъ) (small rural district), and village were classified with respect to natural, historical and economic relations. And the sample was obtained in order that the index characterizing them could reflect the population. In each village 6 peasants were sampled, subject to checking, where 2 were peasants having average size of lands, 2 were ones having lands 1.5~2 times less than average size, and 2 were those having lands 1.5~2 times larger than average size. This result is interpreted as playing an important role to see typical characteristic.<sup>(11)</sup> However, in this case, it is much doubtful if random sampling was assured, and it is unavoidable to set a certain restriction in interpreting those results. It is necessary for us to interpret those results in view of something like case study.

In additon to above family budget surveys of peasant households, we have several others in this field. For example, we have family budget surveys by A. B. Karpov in the 1870's, by A. A. Rusov, and by G. I. Osadch in 1891. In this way Zemstvo statistics surveyed over 11,000 family budgets of peasant households from the beginning of the 1870's to 1917.<sup>(12)</sup>

4. Let us consider the survey carried out in Tula (Тула) Prefecture from 1911 to 1914 directed by statistician P. I. Popov (П. И. Попов). In this survey "typical" peasants were sampled. Sampling was carried out based on the peasantry census, and peasants were classified into various "typical" groups. Proportion of the sample to the population was calculated, and using this result the number of the peasants necessary in the survey was proportionally sampled. Furthermore, in this survey for the purpose to characterize the various conditions of production in agriculture, the sample was processed in order that various

---

(11) V. Krylov, "O primeneniі vyborochnogo metoda v zemskoy statistike," *Vestnik statistiki* (В. Крылов, "О применении выборочного метода в земской статистике," *Вестник статистики*), No. 6, 1955, p. 59.

(12) See, I. Ya. Machukha, *op. cit.*, p. 157.

types of peasants could be represented with respect to seeding and work by side work. Thus, index of the type of peasants came to assure to be representative by the type of seeding and work. "Typical" peasants were sampled with respect to such an economic condition as rich, middle, and poor one in 1911, and the indicators of working away from home, side job, and domestic industry were also considered in 1912-14.<sup>(13)</sup>

Among a small number of family budget surveys of the households of industrial workers in Russia, one of the households of workers in oil industry in Baku (Баку), carried out in 1910 by the Committee of Industrial Hygiene directed by A. M. Stopani (А. М. Стопани) could be interpreted as proportional sampling.

Principles of the type and proportionality of the households, and the uniformity of the distribution with respect to the employment of extraction and refining of oil were considered.<sup>(14)</sup> Characteristics of this survey in Baku were in the following points. That is, before households were sampled, quantitative relations of workers with respect to production, kinds of works, nationality, composition of household, and wage, were investigated. Then the type of surveyed household was determined, and the number of households was calculated.

Let us check the result of the sample (see, Table 2).

Total number of the workers in oil industry at Baku in 1909 could be estimated as about 39,000.

Using the  $\chi^2$ -test, the assertion that the sample is representative of the population will be rejected.<sup>(15)</sup>

Besides Baku, certain family budget surveys of industrial workers

---

(13) See, I. Ya. Machukha, S. V. Postnikov, V. A. Samoylov, "Statistika byudzhetrov naseleniya," *Istoriya sovetskoi gosudarstvennoi statistiki* (И. Я. Матюха, С. В. Постников, В. А. Самойлов, "Статистика бюджетов населения," *История советской государственной статистики*), 1960, pp. 297-316.

(14) I. Ya. Machukha, *op. cit.*, pp. 177-178.

(15) I. Ya. Machukha asserts this survey being representative of the population. However, in view of our analysis, his conclusion can not be accepted. See, I. Ya. Machukha, *op. cit.*, p. 173.

**Table 2**  
Number of workers in oil-industry by kinds of works in Baku

Kinds of works	By family budget survey, 1909		By mass data (%)	
	Absolute no.	%	1909	1913
Oil-extracting	1,612	71.9	67.3	64.4
Drilling by contract	258	11.5	17.5	18.6
Oil-refining factory	227	10.1	9.2	8.8
Mechanical production, etc.	147	6.5	6.0	8.2
Total	2,244	100.0	100.0	100.0

were carried out. However, though organizers took the selection of typical households into account, methodological problem of proportional sampling from population had not yet been solved. Moreover, in the above proportional sampling, the assurance and the necessity of randomness in determining sample size and in selecting sampling unit were not theoretically analyzed, which reflects the stage of development of statistics in those days.

5. We can recognize an important role played by Zemstvo statistics in the application of the principle of systematic sampling. The first survey which applied systematic sampling was carried out by A. V. Peshekhonov (А.В. Пешехонов) in several counties, Kaluga (Калуга) Prefecture, in 1896.<sup>(16)</sup> It was the family budget survey of peasant households, and played a significant role in the history of sampling survey.

It was pointed out that systematic sampling was the necessary condition to remove subjectivity from the sampling and to assure the type of the whole group.

The procedure of this survey is as follows. 2,417 peasants, which

(16) I. Ya. Machukha, *op. cit.*, p. 159.

I. Ya. Machukha, S. V. Postnikov, V. A. Samoylov, "Statistika byudzheto-  
tov naseleniya," *Istoriya sovetskoj gosudarstvennoj statistiki* (И.Я. Матюха,  
С.В. Постников, В.А. Самойлов, "Статистика бюджетов населения,"  
*История советской государственной статистики*), 1960, p. 300.

consisted of over 5% of whole peasants in the prefecture, were systematically sampled at the sampling interval equal to 10 at the village from the general public list of peasants. That is, 1st, 11th, 21st, 31st, peasants, and so on were sampled. It is regrettable that since sufficient checking was not carried out with respect to this survey, we can not see how this systematic sampling was representative of the population. In the selection of peasants, some peasants first selected systematically were replaced by the one who voluntarily wanted to be surveyed, and since the understanding of people regarding systematic sampling was not sufficient, they doubted the procedure that every 10th peasant should be surveyed, and rejected to be investigated, so that the consistency of systematic sampling up to the last stage could not be maintained. As a result, the necessity for individual group to be representative socially and economically could not always be recognized.

As an example, let us show the proportion surveyed by each county (see, Table 3).<sup>(17)</sup>

**Table 3**  
Number of peasants by seeding area  
(per 100 peasants)

Group by seeding area	per 100 peasants described in detail			
	Kozelj County	Likhvin County	Peremyshlj County	Kaluga County
No seeding	0.9	0.2	0.5	0.2
Seeding up to 3 dessiatina	6.9	3.4	3.1	1.6
Seeding, 3-6 dessiatina	8.6	5.1	6.0	1.0
Seeding, 6-9 dessiatina	10.3	6.9	6.0	0.9
Seeding, 9-12 dessiatina	10.3	5.1	7.0	1.2
Seeding, 12- dessiatina	13.6	4.5	5.2	0.0
Average	8.3	4.7	5.2	0.7

It is interesting to see how the repeated census carried out in Voronezh Prefecture in 1900 and the repeated survey in Vyatka (Вятка)

(17) V. Krylov, *op. cit.* p. 61.

Prefecture in 1901-2 played an important role in the development of sampling survey methods.

In the former census 46,000 peasants in Zadonsk (Задонск), Zemlyansk (Землянский), Nizhne-Devitsk (Нижне-Девичь), Korotoyask (Коротояк), and Ostrogozhsk (Острогожск) counties were contained, and a third of all counties were sampled in which 40% peasant households were contained.<sup>(18)</sup>

Checking of the type of the sample was carried out in each surveyed village by comparing the percentage and mean between the sample and the characteristic data of previous whole survey. However, this did not characterize the whole prefecture.

In the latter survey selection of typical village for the survey was carried out in 954 areas which were set up by the survey on the districts most important with respect to the unification of natural and historical conditions.<sup>(19)</sup> Setting of typical village in each area was done as follows. A number of the villages for survey was determined to be not less than 1/5 of the villages in the area. Then all villages in the area were divided into groups whose numbers were equal to the ones of the villages for survey. Division of the villages into each group was first carried out by the size of land-ownership of the community, and then, if the number of the villages set for selection was larger than the one of groups by land ownership, these groups in their turn were divided by the size of land-ownership in peasant households. But if this indication did not give sharp distinction, they were divided by the proportion of farm land and other economic indications. In the groups formed in such a way one village was sampled which was more common regarding such indications that were not taken into account under the division of the groups of communities.

Apparently above systematic sampling method looks like the present one which we usually utilize now. However, there is a significant difference between them regarding stochastic background in determining

---

(18) V. Krylov, *op. cit.* pp. 56-57.

(19) V. Krylov, *op. cit.* pp. 57-58.



sample size, start number, etc., and in interpreting survey results. Those days statistical theory and methods could not offer such a background in sample survey. Theoretically, such a systematic sampling method in Russia had a drawback, but we cannot but highly evaluate its historical role in practical sample surveys and its contribution to the development of sampling theory afterwards.

6. We could find the application of multi-stage sampling method in budget survey of peasant households in Penza (Пенза) Prefecture in 1909-1911.<sup>(2)</sup> This survey was a part of general survey on the economic conditions of agricultural production, carried out by the method of five stage sampling. This family budget survey consisted of the combination of peasantry census with systematic sampling and detail budget descriptions of 25 peasants in each county, which meant about one peasant per 1,000 peasants.

Sampling consisted of five stages. First stage consisted of conducting census of all peasants by the condensed cards of peasants. In the second stage every third peasant was surveyed by more complete, and concise cards. In the third stage every ninth peasant was surveyed by still more complete, and detail cards. In the fourth stage every 27th peasant was surveyed by special cards. In the fifth stage family budgets of 25 peasant households in the county were surveyed in detail.

In the above stages, ages of household members, draught animals being provided with, agricultural implements, cost of buildings, yield and income, expenditure on hiring of manpower, selling and buying of agricultural products and turnover of stock-raising were considered in sampling related with the peasantry census. In each sampling stage except the last one systematic sampling method was applied, being considered the relation with the whole group.

Those days they did not have any established sampling theory, and even if statisticians did not intend to introduce arbitrariness in the

---

(2) See, I. Ya. Machukha, *op. cit.*, pp. 162-163.  
V. Krylov, *op. cit.*, pp. 61-62.

survey, living conditions and illiteracy of sampled people, degree of their recognition to statistical survey brought about arbitrariness in the result of survey.

7. In this section we consider the method of collecting data in family budget surveys in Russia. In the surveys carried out from the 1870's to 1917 they applied mainly expeditionary method in gathering budget data, that is, investigators interviewed and asked the members of sampled households on their budget in certain period of time. The most fundamental reason for applying such a method was the high percentage of illiteracy in Russia of those days. For example, rather roughly we can estimate the percentage of illiteracy by age in 1906, using the census data in 1926<sup>(21)</sup> (see, Table 4).

**Table 4**  
Estimate of the percentage of illiteracy by age  
in 1906 in Russia

Age	percentage of illiteracy (%)	Age	percentage of illiteracy (%)
15	15	40	50
20	20	45	57
25	26	50	64
30	34	55	70
35	41	60	76

From our rough estimate, we could derive the conclusion that the percentage of illiteracy of those, who were directly asked by investigator, was very high. In addition to this, the way how investigators themselves were trained, and the way how the survey was conducted could influence the result of the survey. Furthermore, under the expeditionary method, in general the members of the sampled households answered calling to remembrance the questions on their budget in certain period of time, so that the data obtained by investigators

(21) A. Boyarskii, "K voprosu o metodologii gramotnosti naseleniya," *Vestnik statistiki* (А. Боярский, "К вопросу о методологии грамотности населения," *Вестник статистики*), No. 3, 1928, p. 68.

could not be free from inaccuracy.

The survey by questionnaire method could hardly be found in Zemstvo statistics. We have only two cases: That is, we have temporary survey of industrial workers in Peterburg (Петербург) in 1907-8 and the survey of industrial workers and craftsmen in Kiev (Киев) in 1913. In the former survey, of 1016 questionnaires collected only 632 were effective to be processed, and there were found many cases of unbalance between income and expenditure in describing. This survey method was confined to those households whose members were literate, and it could not assure the accuracy of the survey result.

The method of collecting data from housekeeping account books, which is now used in family budget survey in the U.S.S.R. and many other countries, was not applied in Russia.

As is seen from the above, the method of collecting family budget data in Russia was much restricted by the social, cultural, and economic conditions of those days.

8. In the previous sections, we considered some aspects of family budget surveys of peasants and industrial workers in Russia. From the view point of present sampling survey methods, most interesting is the systematic sampling method first applied in Kaluga Prefecture in 1896 among those applied in family budget surveys in Russia. It is an excellent procedure to apply an idea of systematic sampling to practical field to insure the sampled group to be representative of the whole group.

A. A. Gurjev (А. А. Гурьев) described that there did not exist any literature or documental data on the popularization of sampling survey by A. Kiaer in the middle of the 1890's or even in the beginning of 1898, When the systematic sampling survey in Kaluga Prefecture was carried out by A. V. Peshekhonov (А. В. Пешехонов).<sup>(22)</sup> Furthermore,

---

(22) A. A. Gurjev, "Proiskhozhdenie vyborochnogo issledovaniya i pervye ego opytu v Rossii," *Vestnik statistiki* (А. А. Гурьев, "Происхождение выборочного исследования и первые его опыты в России," *Вестник статистики*), No. 1-4, 1921, pp. 45-46.

he develops his interpretation as follows. That is, A. Kiaer's idea of typical sampling survey, which was applied in Norway in 1895, might not have any direct influence on the systematic sampling survey conducted by A. V. Peshekhonov in Kaluga Prefecture in 1890; A. I. Chuprov's (А. И. Чупров) reference to A. Kiaer, which was recorded in the statistical literature for the first time, will be found in connection with the report of Zemstvo statistics in 1898; if A. Kiaer's ideas were known to Zemstvo statisticians in the middle of the 1890's, they could be utilized in the project of repeated investigation in Samara Prefecture, but it really was not; if there existed any influence on the project by anyone, it is A. I. Chuprov's ideas which were described in the report of 1894.

In view of the above A. A. Guriev's statements, we would be able to derive our conclusion that the systematic sampling survey in Kaluga Prefecture in 1896 was not influenced by A. Kiaer, and they were really independent of each other. Both of them should be interpreted as the first example of systematic or typical sampling survey in the history of sampling survey.<sup>(23)</sup>

In Zemstvo statistics they recognized the necessity to insure the sampled group to be representative, applied an idea of proportional sampling in practice, compared the survey results with the whole peasantry census regarding various characteristics, and paid attention to the validity of analyzed results. Those are points which furnish us with much information even in the present time. The method of multi-stage sampling applied in Penza Prefecture in 1901-1911 was a novel one, but this idea was not utilized afterwards in the field of family budget survey in Russia.

For the purpose to insure the sampled group to be representative of the whole group, Zemstvo statistics intended to sample the peasant households to represent the distribution in the population with respect to certain indications. However, in view of the present theory of

---

(23) See, for example, Kiyoshi Takeuchi, *op. cit.*

sample survey, they did not apply stratified random sampling to sample typical peasant households, so that in case of extending the survey results to whole group, it might be permissive to analyze and interpret them from the view point of case study, but we should not forget that there is a serious restriction in interpreting survey results. It is especially so when sample size was so small.

Under the expeditionary method mentioned above, the investigators took a certain technical training, which insured the higher quality of primary data than questionnaire method in those days in Russia, but there existed the following serious drawbacks inherent in it. In the first, it did not insure the accuracy of the indicators on family budget, for it was much hard for the peasants to give accurate answers to all questions asked by investigator putting in remembrance regarding certain period of time—sometimes all the year. In the second, it took so long a time as the government could not use family budget data for administration to conduct survey and sum up them.

At any rate, family budget surveys in Russia had relatively advanced characteristics, compared with those in the European countries at that time. Even though Zemstvo statistics had some drawbacks, we cannot but evaluate their positive contributions and roles in the history of the development of family budget survey in Russia as well as in the world.