Graduate Research 2010

Description of Methodology

The basic population of the research

The population of the research: the survey included those who graduated from state recognised higher education institutions in 2007 from either undergraduate, supplementary undergraduate or post graduate programmes.

The survey covered ten educational fields
- agriculture
- humanities
- economics
- information technology
- law and administration
- technical
- medical and health care
- teacher training
- social science
- natural science

The four educational fields (sports science, defence and military, arts and arts management) with the smallest number of graduates were not included due to the small population (the graduates of these fields are being interviewed parallel with the questionnaire survey).

The basic population is 4,382 people (Source: Statistics from the Ministry of Education and Culture)

The data sources are the address lists of the study database of higher education institutions containing partial data cycles provided by the institutions. The address list of the sample was upgraded with the 2010 data through the data provision of KEK-KH.

Method of data collection

This was a standard questionnaire survey with notebook data recording based on personal interviews. Data collection was facilitated through a card leaflet given to the interviewee. Data collection was performed by interviewers from the Medián Közvélemény- és Piackutató Intézet and the GFK Hungária Piackutató intézet. In the case of graduates from the Budapest Business School, data was collected with student participation in May and June 2010 with Medián providing the professional background.

Managing address lists

The address list that served as the basis for sampling included 4,879 main addresses, and under each main address ideally three, hierarchical, supplementary addresses, with one clearly supplementary address. The full sample included 17,793 records, since four appropriate (equivalent in terms of gender, educational field and institution) supplementary addresses were organised to each main address.

After data collection had commenced, and during the first few hundred interview attempts, the planned hierarchic management of the address list was modified at the suggestion of the company performing the data collection. This resulted from feedback on the field work of the survey that in this structure, i.e. that visiting main addresses and the appropriate main addresses* sequentially cannot be performed effectively. (This was in addition to the territorial division and the difficulties of establishing contact, which required multiple visits, the fact that many subjects lived in settlements different to those included in the sample, consequently—after a new address had been determined—new visits were required on the main address and the original supplementary addresses became inapplicable. This required complex logistics and consequently an amount of additional time that threatened the completion of the survey within the planned time frame.)

Having recognised the problem, the company commissioned to control the data collection worked out a system within which it undertook to attempt to reach the whole sample independently of the rank of the address and would organise these into the original system afterwards. In order to reduce the distortion of the data collection to the minimum, reaching the subject at the main address was still preferred in terms of multiple visits and following the change of address. We received some information about 14,284 (80 percent) of the 17,793 addresses, or 4,051 (83 percent) of the 4,879 main addresses. Interviews were conducted at 1,647 (41 percent) of the 4,051 main addresses.

Reaching addresses, number and rate of successful interviews

Address	reached sample	actual sample	rate of interviews
main address	4051	1647	41
first supplementary address	3725	1066	29
second supplementary address	3525	968	27
third supplementary address	2983	826	28
Total	14 284	4507	31

The assumption that the modified data collection process did not damage the quality of data collection is supported by the fact that the mean three dimensional weighting variable does not depend on whether the interview was conducted at the main or (any) supplementary address.

Characteristics of the weighing variable in the sample not weighed according to main address and supplementary addresses

Address	Mean	N	Standard Deviation
main address	,9799	1647	,48675
first supplementary address	,9974	1066	,53372
second supplementary address	,9910	968	,48875
third supplementary address	1,0543	826	,54503
Total	1,0000	4508	,51009

Sampling

The method of sampling is simple random sampling. The sample is formed by institutional sub-samples to which the list of graduates was provided by the institutions participating in the survey. (In the case of the Budapest Business School and the University of Szeged the address lists were not delivered, and the sample-rated address lists of those two institutions produced on the basis of detailed sampling instructions were directly managed by the companies conducting the data collection without Educatio Nonprofit LLC.)

Almost all institutions working on the given educational field were included in the survey sample. Since the almost complete list of graduates was available, the sample formed by simple random method does not represent the national division of graduates in individual educational fields by the place (faculty) of graduation.

The principle for forming the sample was based on institutional sub-samples that were not rated among themselves. This was in order to provide a sufficient number of elements for the later analysis of educational fields for educational fields with fewer graduates. Consequently, the individual educational fields are not represented in the sample according to their rate of graduates. This intentional distortion was corrected by weighting after the database was formed.

The number of elements in the planned sample: 5500 people.

The division of the planned sample according to educational fields:

	Number of graduates	Female	Sample size
Agriculture educational field	3879	2318	500
Humanities educational field	7592	6092	700
Economics educational field	11504	8595	800
Information technology educational field	3293	666	475
Law and administration educational field	3714	2412	500
Technical educational field	6016	1787	700
Medical and health care educational field	3195	2527	475
Teacher training	3732	3552	500
Social science educational field	4499	3631	500
Natural science educational field	1958	1184	350
Total	49 382	32 764	5500

Weighing

Size of planned sample: 5500 people
Number of elements in the database established: 4511 people

Weighting was based on the division of the total sample frame according to gender, institution and educational field. (The extent of distortion arising from the difference of the number of supplementary addresses can not be specified precisely; nevertheless this cannot have much effect due to the fact that the address lists had to reflect the division of educational fields according to gender by institution, and in order to do so their numbers are primarily significant, which was ensured [36 percent of the complete basic population]. Where there were not enough supplementary addresses, it resulted from the fact that there were not sufficient numbers of cases in the basic population in that educational field.) Since the graduates from Budapest Business School and the University of Szeged were not included in the sample frame, the graduates of these two institutions were considered according to the ratio of the number of graduates and the weighting factors.

We applied multiple factor cell weighting. The weighting variable is three dimensional: it includes theoretically 480 and practically (since not all institutions provide education in all educational fields) 244 cells according to gender (2 categories), educational field (10 categories) and institution (24 categories). Cell weighting is a significantly more reliable procedure than edge-weighting because the 'edge variables' used for weighting with the

latter become almost independent from one another, thus precisely the same number of men and women are represented in the sample in each institution and in each educational field as in reality.

The value of the weighting variable ensuring representation within the educational fields is between .2 and 7.1, its average is 1, and its dispersion is .507. The closer the average of the weighting variable is to one the closer the given group is represented in the full sample according to its actual rate; the value is above one in the case of groups that were reached in a lower rate than their true proportion (these are underrepresented in the unweighted real sample), and it is below one in the case of groups that were reached at higher rate than their true rate. There is hardly any difference in terms of reaching men and women; graduates from Zsigmond Király College, Pázmány Péter University and the Budapest Business School (in a different structure, using the database of the institution, interviewed by the students of the institution) are significantly underrepresented, while graduates of the University of Kaposvár and the University of West Hungary are overrepresented. In terms of reaching educational fields, the differences are significantly smaller than in terms of institutions: reaching graduates of teacher training and social sciences majors was the easiest, while graduates of the law and administration educational fields were the most difficult to reach.

Characteristics of the educational field (suly1_1) weighting variable in the sample not weighted according to weighting factors

gender	Mean	N	Standard Deviation
male	1.0922	1401	.61148
female	.9585	3110	.45083

institution	Mean	N	Standard Deviation
BCE	1.2022	176	.38403
BGF	1.9163	253	.06853
BME	1.4265	208	.36174
DE	.8666	482	.22245
DF	.6783	65	.11405
EKF	.7165	110	.24582
ELTE	1.2887	377	.31220
KE	.5796	151	.18459
KF	.6369	176	.18999
KJF	1.0125	43	.30617
KRE	.9868	55	.34646
KRF	.9865	144	.12217
ME	1.2092	181	.39478
NYF	.8142	143	.28866

NYME	.5839	389	.12914
PE	.7576	145	.26515
PPKE	2.1014	33	.57450
PTE	.7714	485	.20790
SE	1.2436	131	.05625
SZE	.6960	231	.19425
SZF	.8482	65	.32575
SZIE	.8223	213	.44250
ZSKF	2.1616	32	.53744
SZTE	1.4036	223	1.03652

educational field	Mean	N	Standard Deviation
agricultural	.9600	383	.37465
humanities	1.0516	536	.46357
economics	1.3179	733	.55822
information technology	1.1464	278	.62166
law and administration	1.2536	308	.40308
technical	.9303	569	.33096
medical and health care	1.0327	355	.33107
teacher training	.6446	581	.31877
social science	.7627	516	.62369
natural science	.9709	249	.45378
Total	1.0000	4508	.51009

Since individual educational fields are not represented in the frame sample according to their actual rate in order to be dependably analysable, the weighting variable established by cell weighting was corrected by edge weighting (on the basis of the OKM statistics serving as the basis for developing the sample). The correction hardly affected the division of the variable used for the other two weightings, there is only considerable difference (1 percentage point) in terms of the rate for the Budapest Business School.

Division of the data file according to the variables used for weighting (percentage

	cell weight	corrected weight
gender		
male	33.9	33.7
female	66.1	66.3

institution		
BCE	4.7	4.6
BGF	10.7	11.7
BME	6.6	6.7
DE	9.3	8.7
DF	1.0	1.0
EKF	1.7	1.7
ELTE	10.8	10.6
KE	1.9	1.8
KF	2.5	2.4
KJF	1.0	1.1
KRE	1.2	1.2
KRF	3.1	3.2
ME	4.9	4.8
NYF	2.6	2.6
NYME	5.0	5.1
PE	2.4	2.5
PPKE	1.5	1.5
PTE	8.3	8.4
SE	3.6	2.9
SZE	3.6	3.7
SZF	1.2	1.3
SZIE	3.9	3.9
ZSKF	1.5	1.7
SZTE	6.9	7.0
educational field		
agricultural	8.2	7.9
humanities	12.5	15.4
economics	21.4	23.3
information technology	7.1	6.7
law and administration	8.6	7.5
technical	11.7	12.2
medical and health care	8.1	6.5
teacher training	8.3	7.6
social science	8.7	9.1
natural science	5.4	4.0

In order to explore the characteristics of individual educational fields and to compare the data of educational fields we recommend the use of cell weight (weight1_1), while for the analysis of the full basic population of the 10 educational fields we recommend the use of the weighing variable corrected by edge weighing (final weight sample).