

**Effects of rotation groups, interviewing modes
and interviewers on the LFS estimates**

Florentina Álvarez

Juana Porras

The views expressed in this working paper are those of the authors and do not necessarily reflect the views of the Instituto Nacional de Estadística of Spain

First draft: December 2010

This draft: December 2010

Effects of rotation groups, interviewing modes and interviewers on the LFS estimates

Abstract

This paper examines the influence of three factors (rotation groups, method of interview and interviewer effect) on the main estimations in the Labour Force Survey (LFS), by performing probit and homogeneity analysis. It also shows that the influence of the interview method is partially due to the different representation of foreign people in the CATI and CAPI samples. Finally, it highlights the importance of a correct identification of the bias sources and outlines the future plans to improve the standardization in the Spanish LFS.

Keywords

Bias Sources, LFS, Rotation Groups, CAPI y CATI, Interviewer effects

Authors and Affiliations

Florentina Álvarez and Juana Porras

National Statistics Institute of Spain

Workshop on Labour Force Survey (LFS) methodology Paris, 15-16th April 2010

Effects of rotation groups, interviewing modes and interviewers on the LFS estimates

Florentina Alvarez
Juana Porras
SPAIN, National Statistics Institute

I Introduction

The Labour Force Survey (LFS) is included in Spain in the Active Population Survey, conducted since 1964.

The Spanish LFS sample is divided into six waves (rotation groups), each of them being a representative sample of the population. The households of the sample are interviewed once per quarter, during six consecutive quarters. All households of a specific rotation group are renewed every quarter. In this way, two consecutive quarters have in common five-sixths of the sample, and the whole sample has been completely renewed after one year and a half.

The CAPI system is used for data collection in the first interview and the CATI system in the others. Families without telephone or unwilling to be interviewed using this system are always interviewed face to face. Therefore, approximately one quarter of the sample provides information by means of CAPI and the rest of the sample by telephone.

Traditionally, the existence of a rotation bias in the estimates has been assumed. Similarly, it has been found that different methods of data collection are not neutral regarding the estimation of certain variables. Finally, another important bias source is the interviewer effect.

This paper examines the influence of these three factors (rotation groups, method of interview and interviewer effect) on the main estimations, by performing probit and homogeneity analysis.

On the other hand, it shows that the influence of the interview method is partially due to the different representation of foreign people in the CATI and CAPI samples.

In conclusion, the paper highlights the importance of a correct identification of the bias sources. It also outlines the future plans to improve such standardization in the Spanish LFS.

II. Influence of the rotation groups and the interview methods over the main variables of the surveys. Other influences.

The rotation groups and the interview methods have been considered as the most important sources of bias in the Spanish LFS, regarding the classification of people as in employment, unemployed or inactive.

In order to test the influence of these sources, we have performed two probit analyses. The results, in table 1, refer to the second quarter of 2009. They are similar for the other quarters of the same year.

TABLE 1

ESTIMATION OF THE PROBABILITY OF BEING IN EMPLOYMENT EXPLICATIVE VARIABLES: ROTATION GROUP, INTERVIEW MODE

Dependent variable: In employment=1, not in employment=0

Rotation group: TR= 1,2,...,6

Interview mode: METODO = 1(CATI), 6(CAPI)

Type III Analisis of effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
TR	5	62,1909	<.0001
METODO	1	87,6655	<.0001

Analysis of Parameter Estimates								
Parameter	DF	Estimate	Standard error	95% Confidence limits		Chi-Square	Pr > ChiSq	
Intercept		1	-0.2519	0.0289	-0.3085	-0.1953	76.01	<.0001
TR	1	1	0.0185	0.0219	-0.0245	0.0614	0.71	0.3989
TR	2	1	0.0144	0.0221	-0.0288	0.0576	0.43	0.5141
TR	3	1	-0.0111	0.0223	-0.0548	0.0326	0.25	0.6179
TR	4	1	0.2356	0.0329	0.1711	0.3001	51.26	<.0001
TR	5	1	-0.0065	0.0220	-0.0496	0.0366	0.09	0.7670
TR	6	0	0	.	.	.	.	.
METODO	1	1	0.2496	0.0267	0.1973	0.3018	87.67	<.0001
METODO	6	0	0	.	.	.	.	.

ESTIMATION OF THE PROBABILITY OF BEING UNEMPLOYED EXPLICATIVE VARIABLES: ROTATION GROUP, INTERVIEW MODE

Dependent variable: Unemployed=1, not unemployed=0

Rotation group: TR= 1,2,...,6

Interview mode: METODO = 1(CATI), 6(CAPI)

Type III Analisis of effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
TR	5	17,7089	0.0033
METODO	1	12,2771	0.0005

Analysis of Parameter Estimates								
Parameter	DF	Estimate	Standard error	95% Confidence limits		Chi-Square	Pr > ChiSq	
Intercept		1	-11.504	0.0369	-12.227	-10.781	972.35	<.0001
TR	1	1	0.0593	0.0291	0.0023	0.1164	4.15	0.0415
TR	2	1	-0.0094	0.0298	-0.0678	0.0490	0.10	0.7522
TR	3	1	0.0084	0.0300	-0.0503	0.0672	0.08	0.7788
TR	4	1	-0.0968	0.0425	-0.1802	-0.0135	5.19	0.0227
TR	5	1	0.0443	0.0293	-0.0132	0.1017	2.28	0.1308
TR	6	0	0	.	.	.	.	.
METODO	1	1	-0.1179	0.0336	-0.1838	-0.0519	12.28	0.0005
METODO	6	0	0	.	.	.	.	.

The dependent variables are EMPLOYMENT in the first test and UNEMPLOYMENT in the second. The explicative variables are TR=rotation group and METODO=interview method.

The models give the estimated probability of being in employment or unemployment.

Regarding employment, there is a significant influence of the rotation group number 4 (TR 4), the group that is in the first interview in the second quarter. The parameter of the CATI method is also significant. The results are in the same line for the other three quarters (always for the group in first interview and the CATI mode).

The two parameters (TR=4 and CAPI) could also be considered significant for the unemployment, given that the p-values are small, although not in the same degree as in the case of the employment.

The fact that only the group in first interview, all done with CAPI, seems to be different, could indicate that all the groups could be equivalent and only the method is the cause of the differences regarding the probability of being in employment.

To analyse this hypothesis, a test of homogeneity has been applied between the sample of TR 4 with the sample of the other TR. In this way, the difference of means test between each TR has been used.

The null hypothesis that is tested is that the two samples come from the same population. Under the null hypothesis, the confidence interval $\hat{X}_1 - \hat{X}_2 \pm 1,96\hat{\sigma}_{(\hat{X}_1 - \hat{X}_2)}$ contains the value 0, at a confidence level of 95%.

The results are in table 2

TABLE 2

TR	C.V.TRA(%)	C.V. TRB(%)	OCU TRA	OCU TRB	VAR(TRA-TRB)	LONGINTER	OCU(TRA-TRB)	I1	I2	P-Valor
1-4	0,89	0,86	18951787	18771144	54510112943	915218	180642	-276967	638251	0,4391
2-4	0,90	0,86	19237327	18771144	56036293417	927942	466182	2211	930153	0,0489
3-4	0,91	0,86	19100719	18771144	56272429857	929895	329574	-135373	794522	0,1647
5-4	0,75	0,86	18736483	18771144	45807127591	838982	-34662	-454153	384829	0,8713
6-4	0,82	0,86	19067891	18771144	50507659637	880977	296747	-143742	737235	0,1867

It is noted that, except in the case of rotation group 2, in all the other cases the samples are homogeneous. However, in the above case the p-value, although less than 0,05 it is very closed to it, i.e. the probability of not accepting the homogeneity is very small.

Same results are observed in the other quarters

As the above results do not justify the difference in the behaviour of this TR, a new test of homogeneity has been applied to see the effects of the interview method on the estimates.

The difference of means test between the samples interviewed with CAPI and CATI method respectively, has the following result:

TABLE 3

C.V.CATI(%)	C.V.CAPI(%)	OCUCATI	OCUCAPI	VAR(CATI-CAPI)	LONGINTER	OCU(CATI-CAPI)	I1	I2	P-Valor
0,43	1,06	19072690	18550743	45392513673	835176	521948	104359	939536	0,0143

In this case the confidence interval does not include the value 0, so we do not accept the hypothesis that the two subsamples, CATI and CAPI are homogeneous, i.e. it exists different behaviour between both samples.

Regarding unemployment, the results of the difference of means test indicates that there is not effect of the interview methods. The samples CATI and CAPI are homogeneous for this variable.

C.V.CATI(%)	C.V.CAPI(%)	PAROCATI	PAROCAPI	VAR(CATI-CAPI)	LONGINTER	PARO(CATI-CAPI)	I1	I2	P-Valor
1,40	3,01	4110995	4211999	19385918213	545795	-101004	-373902	171893	0,4682

There is another factor to take into account in the analysis. In the periodical feedbacks done with the interviewers and supervisors in the CATI and CAPI centres, it was found that there is a greater proportion of foreigners in the CAPI method, because it is more difficult to conduct the interview by telephone for them (they do not have landline telephone, change very frequently the number of the mobile phone, etc.) and in the second and subsequent interviews the foreigners are passed from the CATI to the CAPI centres. This issue was confirmed with a cross-tabulation of variables 'nationality' and 'interview mode'.

TABLE 4

LFS. Interview mode and citizenship (%)

2-2009	CITIZENSHIP		
	ALL	SPANIARDS	FOREIGNERS
All	100,0	94,6	5,4
CATI	100,0	94,9	5,1
CAPI	100,0	93,7	6,3

It seems probable that all the variables correlated with the nationality have different proportion in the sample of CAPI and CATI, and this fact is not a consequence of the rotation group and of the method itself. That could be the case for the classification as employed or unemployed.

In order to analyse this, we have performed two probit analyses again, to test the influence of the fieldwork method, the rotation group and the nationality, over the classification regarding economic activity (employment, unemployment). The results are in table 5, referred also to the second quarter of 2009. They are similar to the other quarters of the same year.

Regarding employment, there is a significant influence of the rotation group number 4, the CATI method and the Spanish nationality (in this case, negative). The results are in the same line for the other three quarters.

Regarding unemployment, there is a significant negative influence of the Spanish nationality, in the four quarters. The CATI method has also an influence, although less remarkable.

TABLE 5

ESTIMATION OF THE PROBABILITY OF BEING IN EMPLOYMENT
EXPLICATIVE VARIABLES: ROTATION GROUP, INTERVIEW METHOD AND CITIZENSHIP

Dependent variable: In employment=1, not in employment=0
 Rotation group: TR= 1,2,...,6
 Interview method: METODO = 1(CATI), 6(CAPI)
 Citizenship: NAC = 1 (Spaniards and double citizenship), 6 (foreigners)

Analysis of Parameter Estimates								
	Parameter	DF	Estimate	Standard error	95% Confidence limits		Chi-Square	Pr > ChiSq
2-2009	Intercept	1	-0,0929	0,0329	-0,1574	-0,0284	7,97	0,0047
	TR	1 1	0,0181	0,0219	-0,0249	0,0611	0,68	0,4093
	TR	2 1	0,0139	0,0221	-0,0293	0,0572	0,4	0,5277
	TR	3 1	-0,0085	0,0223	-0,0522	0,0352	0,15	0,7032
	TR	4 1	0,249	0,033	0,1844	0,3136	57,03	<,0001
	TR	5 1	-0,0072	0,022	-0,0504	0,0359	0,11	0,7421
	TR	6 0	0					
	METODO	1 1	0,2646	0,0267	0,2122	0,317	97,94	<,0001
	METODO	6 0	0					
	NAC	1 1	-0,1976	0,0195	-0,2357	-0,1594	102,9	<,0001
	NAC	6 0	0					

ESTIMATION OF THE PROBABILITY OF BEING UNEMPLOYED
EXPLICATIVE VARIABLES: ROTATION GROUP, INTERVIEW METHOD AND CITIZENSHIP

Dependent variable: Unemployed=1, not unemployed=0
 Rotation group: TR= 1,2,...,6
 Interview method: METODO = 1(CATI), 6(CAPI)
 Citizenship: NAC = 1 (Spaniards and double citizenship), 6 (foreigners)

Analysis of Parameter Estimates								
	Parameter	DF	Estimate	Standard error	95% Confidence limits		Chi-Square	Pr > ChiSq
2-2009	Intercept	1	-0,7394	0,0409	-0,8195	-0,6593	327,17	<,0001
	TR	1 1	0,059	0,0295	0,0012	0,1167	4,01	0,0453
	TR	2 1	-0,0088	0,0301	-0,0678	0,0502	0,09	0,7702
	TR	3 1	0,0191	0,0303	-0,0403	0,0786	0,4	0,5283
	TR	4 1	-0,0557	0,0431	-0,1402	0,0288	1,67	0,1965
	TR	5 1	0,0451	0,0296	-0,0129	0,1032	2,32	0,1277
	TR	6 0	0					
	METODO	1 1	-0,0727	0,0342	-0,1397	-0,0057	4,52	0,0335
	METODO	6 0	0					
	NAC	1 1	-0,5427	0,0225	-0,5868	-0,4987	583,72	<,0001
	NAC	6 0	0					

To test the results obtained above, a new test of homogeneity has been done between CATI and CAPI, but considering only the Spanish people. The results in this case are in table 6:

TABLE 6

C.V.CATI(%)	C.V.CAPI(%)	OCUCATI	OCUCAPI	VAR(CATI-CAPI)	LONGINTER	OCU(CATI-CAPI)	I1	I2	P-Valor
0,42	0,98	19484604	18907751	41031614833	794045	576853	179831	973876	0,0044

From this result we can not conclude that the effect is due exclusively to the different representation of the nationality in each subsample.

III. The interviewer effect

One of the most important sources of error in the estimate is the interviewer effect. The problem in this case is the difficulty to measure it.

An approach to this measure has been done, taking advantage of the increase in the sample in an autonomous community. The increased sample has exactly the same design of the original one and the field work methods are also identical. We have compared the original sample with the sample added. They have been interviewed by different CATI teams and we have considered the two CATI teams as two different interviewers.

In the same way that it is explained in the previous paragraph, a test of homogeneity has been applied.

TABLE 7

C.V.INE(%)	C.V.IGE(%)	OCUINE	OCUIGE	VAR(INE-IGE)	LONGINTER	OCU(INE-IGE)	I1	I2	P-Valor
2,18	2,32	1134091	1111912	1276688357	140065	22179	-47854	92211	0,5348

According to this test both samples are homogeneous, so we can not conclude that there is an interviewer effect.

Similar results have been obtained in the case of the sample interviewed by CAPI.

TABLE 8

C.V.INE(%)	C.V.IGE(%)	OCUINE	OCUIGE	VAR(INE-IGE)	LONGINTER	OCU(INE-IGE)	I1	I2	P-Valor
3,57	3,35	1112753	1121890	2990602769	214371	-9137	-116322	98049	0,8673

The result is not surprising, given that the increase of the sample was designed with the aim of having the same process than in the original sample, in particular preserving the control of the non-sampling errors.

With the CATI analysis we can conclude that as far as data collection is standardized, the interviewer effects can be controlled.

IV. Conclusions and future plans

Different probit analyses have been performed in the Spanish Labour Force Survey, to test the influence of the rotation group, the interview method and the proportion of foreigners over the classification of the population (as in employment or unemployed). The conclusion is that the three factors have significant effects in the case of the employment and, in minor degree, over the unemployment.

Homogeneity tests have also been performed with the same objectives and the conclusion, in this case, is that only the interviewing mode is significant.

In the same way, the homogeneity tests show that there is not an interviewer effect.

Therefore, from the analyses done of some possible sources of bias in the main LFS results, it is clear that the interview method has a significant effect over the estimates. The modifications in the Spanish LFS in 2005 were made with the first objective of standardising the interviews in all the waves and in all the regions of the country. Five years after, it seems that the process has not finished. The interviews performed with CATI are fully controlled and harmonized, but it is not the case for those done with CAPI (that essentially correspond to the first interview).

The next step in the process will be to introduce a mixed method in the first interview. It is planned to have a personal interview to get: a) the demographic characteristics of the persons living in the household, and b) a telephone number of the family. The rest of the questions of the LFS questionnaire would be collected by telephone, in a complementary CATI first interview.

The new process has to be tested and, in case of positive results and not budgetary restrictions, applied in the LFS.