

The Switch to Telephone Survey Mode in the CIS Barometer. A Comparison with ICC Time Series

*El cambio al modo de administración telefónico en el barómetro del CIS.
Una comparación con las series temporales del ICC*

Jaime Balaguer

Key words

Centre for Sociological Research

- Face-to-Face Surveys
- Telephone Surveys
- Data Collection Methodology

Palabras clave

Centro de Investigaciones Sociológicas

- Encuestas presenciales
- Encuestas telefónicas
- Metodología de recolección de datos

Abstract

In April 2020, Spain's Centre for Sociological Research (CIS) changed its survey mode from face-to-face (PAPI/CAPI) to telephone-based (CATI). In this study we examine the impact of this change on two *CIS barometer* variables: marital status and education level. We describe their evolution from 2013 to 2023 using the Consumer Confidence Index (ICC), a continuous survey that has always been telephone-based, as a comparative reference. As we will see, the change in mode produced discontinuities in the measurement of education level, with implications for the quality and interpretation of results. In addition, the evidence suggests that the private research companies the CIS contracted to carry out telephone fieldwork had their own biases that must be taken into account for series comparison and monitoring.

Resumen

En abril de 2020 el Centro de Investigaciones Sociológicas (CIS) cambió el modo de administración de sus encuestas, pasando del personal (PAPI/CAPI) al telefónico (CATI). En el presente trabajo examinamos su impacto en dos variables del barómetro: el estado civil y el nivel de estudios. Describimos su evolución mensual desde 2013 a 2023 utilizando el Índice de Confianza del Consumidor (ICC), un estudio continuo que siempre ha sido telefónico, como referencia comparativa. Como veremos, el cambio de modo produce discontinuidades en la medición del nivel educativo, con implicaciones para la calidad e interpretación de resultados. Además, la evidencia apunta a que las empresas privadas que el CIS ha contratado para realizar los trabajos de campo telefónicos del ICC presentan sesgos propios que deben tomarse en cuenta para la comparación y seguimiento de las series.

Citation

Balaguer, Jaime (2024). "The Switch to Telephone Survey Mode in the CIS Barometer. A Comparison with ICC Time Series". *Revista Española de Investigaciones Sociológicas*, 185: 145-162. (doi: 10.5477/cis/reis.185.145-162)

Jaime Balaguer: Ministerio de Hacienda y Función Pública | jtbalaguer@sepg.hacienda.gob.es



INTRODUCTION

In April 2020, Spain's Centre for Sociological Research (CIS) changed its usual way of carrying out surveys, shifting from face-to-face interviews (PAPI/CAPI) to telephone surveys (CATI). In a context marked by the COVID-19 pandemic, in which the use of the telephone was a reasonable substitution, the CIS outsourced the survey process for some months to the firms Intercampo and IMOP. In September of 2020, its commitment to CATI was solidified through the establishment of an infrastructure for telephone-based fieldwork through TRAGSATEC (a state-owned enterprise providing professional services to public administrations).

The shift in the mode of survey administration involved an organizational and methodological transformation that, taking into account the widespread use of CIS surveys, is of interest to the research community. Telephone-based surveys have the advantage of speed, lower cost, high professionalization of the interviewers, a wider geographic sample and the possibility of increasing contact attempts. In terms of disadvantages, they involve addressing a significant volume of telephone-related issues, the limitation in the interview to the use of auditory stimuli, the possibility of increased distractions for interviewees and a decline in the overall quality of responses (Bosch and Torrente, 1993: 10-22; Díaz-de-Rada, 2019; Dillman, Smyth and Christian, 2014; Holbrook, Green and Krosnick, 2003; Tourangeau and Smith, 1996).

Faced with a transition of this magnitude, it would have been useful to carry out an exhaustive methodological study to assess its impact on two key elements: sample representativeness and the presence of measurement errors. However, the COVID-19 crisis hastened the adoption of CATI, which subsequently became permanent, without first examining these issues.

Despite this shortcoming, we have certain useful tools for assessing this change. In this study, we focus on the CIS barometer, one of the CIS's most well-known studies, which, since 1996, has gathered monthly information on the state of public opinion in Spain. We compare it with the Consumer Confidence Index (ICC), another monthly project, which looks at the expected spending of the resident population and which the CIS has been responsible for since 2011. The fact that ICC data has always been gathered through the telephone makes it a unique reference for evaluating the effects of this change. However, to establish an adequate comparison we must first establish a robust analytical approach.

METHODOLOGICAL APPROACH

Two important factors for our object of study coincide in time: the change to telephone-based surveying for the barometer, and home confinement due to the COVID-19 crisis. This means that when we compare the barometers for March and April of 2020 and look at the variables that repeat in both, the changes that we observe may have three different origins: the impact of the pandemic on public opinion, the change in the administration of the barometer (which may have affected both the representativeness of the sample as well as response errors), and home confinement. The latter specifically led to temporary unemployment for a significant part of the population, which in turn may have affected the probability of contact and acceptance to participate in the survey, as the persons selected were more available (consequently, temporarily impacting the representativeness of the sampling).

To separate these factors, we used the following strategy. First, we work with a wide observation period. The change in the mode of administration (in contrast to

the confinement) is a permanent phenomenon. Therefore, we look for lasting changes, that is, differences between the information gathered before and after 2020. To do this, we can identify variables that are used monthly, observe their evolution and discriminate between long-term trends, irregular variations and truncation of the constant and *jumps* in level that mark a discontinuity in the data series. In concrete, the presence of a jump in level, coinciding with the shift to the use of the telephone, which has subsequently been maintained, would indicate a change in the structure of the data series that permits us to refer to differentiated periods (Peña, 2005: 377-382; Arnaud, 2001: 97 and ff).

Secondly, we excluded attitudinal and opinion variables, on which the COVID-19 crisis (as well as individuals' changed employment situation) could have had a significant impact. Including these would require a differentiated analysis that goes beyond the aims of this research. We focus on variables that describe relatively stable parameters within the population. We discard sex and age as quotas for selecting interviewees. Nor do we use vote recall, as although it refers to a fixed characteristic of the participant (that should only change with the celebration of new elections), as has been shown by Balaguer (2022), there is an important response error in which the political conjuncture plays a significant role. As a result of this process, two variables remain: marital status and education level. Both describe characteristics that change slowly in the population and that should not have undergone important variations during the COVID-19 crisis.

Lastly, we use the ICC as a comparative reference. The ICC has always been a telephone-based survey, so that it has not been impacted by any change in mode. If the confinement involved changes in the probability of being contacted for or accepting participation in the survey, resulting in ef-

fects on marital status and education level, these would be reflected in the ICC series. If the ICC did not experience changes in April 2020, or they were limited, then any jumps in level that we find in the barometer could be associated with the change to CATI mode. As a result, our main hypothesis (H1) proposes that *in April 2020, jumps in level in the series for marital status and education level for the barometer will be of greater magnitude than those observed for the ICC.*

DESCRIPTION OF SERIES

We have data since 1996 for the barometer and since 2011 for the ICC. However, in May of 2013, the ICC began to include individuals that only have cell phones in its samples. This was a substantive change, as it affected coverage of the student population¹. Therefore, we establish this as the initiation date, concluding in March 2023, with a total of 119 monthly observations, 83 prior to April 2020, and 36 after².

In Table 1 we see the averages and standard deviation for the two series of the barometer. We can see that, with the change to CATI, marital status remains relatively stable, while the distribution for education level varies substantially. In particular, the proportion of persons with lower education levels is reduced and the proportion of persons with university education has increased notably (almost doubling).

¹ According to the Survey on Equipment and Use of Information and Communication Technologies of Spain's INE, in 2013, 21.4% of households did not have a landline telephone, a proportion which grew to 37.3 % in 2022.

² The barometer considers the Spanish population of 18 years of age and above, and the ICC, the resident population of 16 years of age and above. This difference can affect series averages but not general trends. In the barometer, the observations missing from the months of August and macro-barometers for 2019, have been replaced by the averages for adjacent scores.

TABLE 1. Averages and standard deviations for the series by marital status and education level in the barometer, by mode of administration (2013-2022)

	PAPI/CAPI (05/2013-03/2020)		CATI (04/2020-03/2023)	
	Average	SD	Average	SD
Marital status				
Married	53.5	1.14	52.4	2.02
Single	31.6	0.90	31.8	1.41
Separated	2.5	0.28	2.0	0.28
Divorced	4.6	0.62	6.6	0.73
Widowed	7.7	0.67	6.9	0.62
Education level				
Without formal ed	5.6	0.60	2.8	0.70
Primary	17.5	1.55	7.2	1.28
Secondary (obligatory)	23.8	1.18	14.6	1.16
Secondary (2nd stage)	13.6	0.68	14.4	1.28
Vocational	18.0	0.89	19.2	1.01
University	21.4	1.58	41.3	2.08
N	83		36	

Source: Own elaboration based on the CIS Data Bank.

It must be taken into account that the first four observations of the barometer using CATI were carried out by private firms contracted by the CIS. Although these companies comply with the criteria on human and material resources required for their contracting (the proposals of the companies are homogeneous in a series of characteristics, such as the minimum number of contact attempts, the percentage of interviews that are supervised or the amount of own resources employed), in practice, they are generic proposals. Variations in the telephone load, in the number and time distribution of contact attempts or in the experience of the assigned staff can lead to differences in the data collected. We refer here to a “House effect”, defined as the bias of the company or entity that carries out the survey, a “black box” that can involve variations in representativeness and information quality.

This factor is even more relevant for the ICC. Of the 119 observations that we analyse, 88 were made by private firms, the rest were

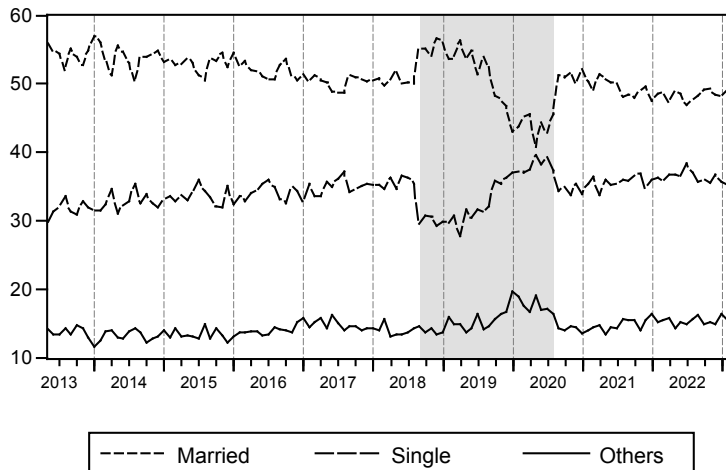
by the CIS-TRAGSATEC. In concrete, 39 were made by GfK (between May 2013 and August 2016), 25 by IMOP (until August 2018) and 24 by Sigma Dos (until August 2020). In Table 2, we see the descriptives for the ICC disaggregated by the firm responsible for the field work. Regarding education level, the increase in the proportion with primary level educations stands out under Sigma Dos, as do the fall in the proportion with obligatory secondary education and the growing proportion with university education, in particular in the last phase. Regarding marital status, we see relative stability, although the measurements from Sigma Dos have a higher standard deviation for the married and single categories.

This anomaly in the variance for marital status during the Sigma Dos period requires a more detailed examination. In Graph 1, we represent its evolution. We can distinguish two stages in the measurements by this firm: a first, beginning in September 2018, and a second from October 2019 to August 2020, with a reverse trend.

TABLE 2. Averages and standard deviations for the series by marital status and education level for the ICC, by the firm that carried out the field work (2013-2022)

	GFK (05/2013-07/2016)		IMOP (08/2016-08/2018)		Sigma Dos (09/2018-08/2020)		CIS-TRAGSATEC (09/2020-03/2023)	
	Average	SD	Average	SD	Average	SD	Average	SD
Marital status								
Married	53.4	1.61	50.6	1.12	50.0	5.18	49.2	1.40
Single	33.0	1.45	34.9	1.15	33.6	3.72	35.8	1.04
Separated	2.4	0.42	2.3	0.39	2.8	0.42	2.2	0.25
Divorced	3.9	0.37	4.6	0.45	5.8	1.39	6.2	0.69
Widowed	7.1	0.49	7.5	0.60	7.2	0.55	6.5	0.50
Education level								
Without formal ed	3.3	0.64	5.2	0.71	4.96	0.89	2.9	0.39
Primary	7.7	1.13	6.6	1.15	12.1	1.81	7.2	1.27
Secondary (1st stage)	27.2	1.83	26.0	1.73	17.9	2.66	16.2	0.95
Secondary (2nd stage)	16.0	1.09	15.5	0.99	16.2	0.98	15.6	1.27
Vocational	19.1	1.19	19.0	1.11	16.4	1.79	18.7	0.73
University	26.5	1.46	27.7	1.27	31.6	1.98	38.9	1.83
	39		25		24		31	

Source: Own elaboration based on the CIS Data Bank.

GRAPH 1. Time series for the ICC for marital status (2013-2023). The Sigma Dos stage in shaded area

Source: Own elaboration based on the CIS Data Bank.

One possible explanation might be found in the contacting strategies. In September 2018, there was a modification made in technical specifications. At that moment the following clarification was added: “persons contacted by mobile

phone that also had a landline were incorporated into the sample of landline + cell phone” (CIS, 2018). In practice, this meant that the number of persons interviewed by cell phone increased from approximately 20 % to approximately 70 %.

As a result, we have examined in detail the telephone-based field work for the ICC from August 2016 to March 2020³. As can be seen in table 3, with the arrival of Sigma Dos, we see a notable increase in the telephones employed, particularly mobile phones. The ratio of calls made/telephones used declines. However, this strategy does not remain stable, but changes. Beginning in November 2019, the use of telephones moderates and the calls increase. So, the ratio rises again.

These two stages could be linked to the variance in the marital status series. Specifically, the ratio calls/telephones has a Pearson correlation coefficient of -0.786 with the proportion of married people and $+0.870$ with single people.

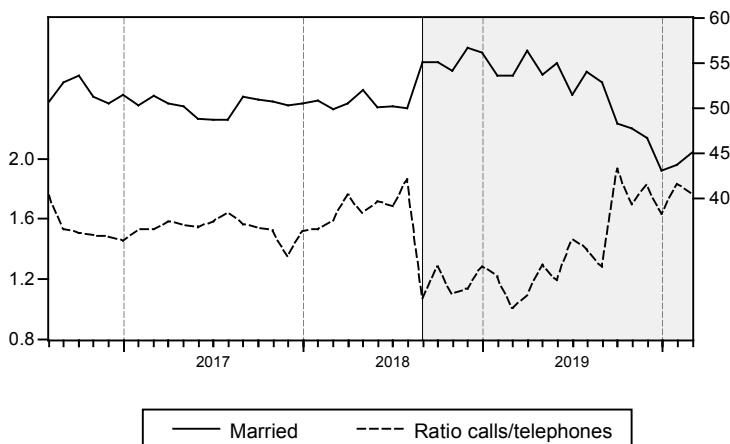
As can be seen in Graph 2, we can talk about a relevant covariation between both phenomena. Although we cannot be certain, nor much less establish a causal relationship, we can take into account the existence of two differentiated stages in the field work of Sigma Dos in the analysis that follows.

TABLE 3. Monthly average of telephones used and calls made in field work for the ICC (2016-2020)

	IMOP (08/2016-08/2018)	Sigma Dos – 1st stage (09/2018-09/2019)	Sigma Dos – 2nd stage (10/2019-03/2020)
Telephones used			
Landlines	35,475	47,381	41,520
Mobiles	14,331	51,585	44,486
Total	49,806	98,966	86,006
Calls made			
Landlines	51,283	49,627	65,301
Mobiles	27,542	71,643	88,033
Total	78,826	121,269	153,334
Ratio calls/telephones	1.58	1.23	1.78

Source: Own elaboration based on ICC field reports.

GRAPH 2. Proportion of married and ratio calls/telephones in the ICC field work (2016-2020). Shaded area, the Sigma Dos stage



Source: Own elaboration based on ICC field reports.

³ We would like to thank Valentín Martínez, responsible for sampling for the CIS, for his help with this part.

RESULTS

So that we can establish the slopes and possible jumps in the series, we use a deterministic model, taking account that our objective is to describe a past trend and not to predict its behaviour (Peña, 2005: 44-45). We use linear regression, estimated through least squares and carried out with EViews 12 software. To simplify, we combine the categories with the least frequency for the variables marital status (separated + divorced + widowed) and education level (no formal education + primary education). As parameters, we introduce a constant, the slope and dummy variables that identify the periods in which the firms mentioned carried out the field work (with two segments for Sigma Dos). Our main independent var-

iable is a dummy that identifies the period that begins in April 2020. We also examine if modifications in the slopes occur.

In Table 4 we show the results of the model for marital status. It must be considered that in the barometer the jump in level of April 2020 coincides with the field work carried out by Intercampo, so that a direct reading of the coefficients can be confusing. We identify their subsequent magnitude, in view of the fit of the regression line. The corrected r-squared is relatively low in several of the response categories, which, along with the normality and homoscedasticity observed in the residuals (Jarque-Bera test and White test)⁴, suggests that, once the effects of the equation are discounted, there can be an important component from the error term (Peña, 2005: 319-321).

TABLE 4. Regression model for the evolution of marital status in the barometer and the ICC (2013-2022)

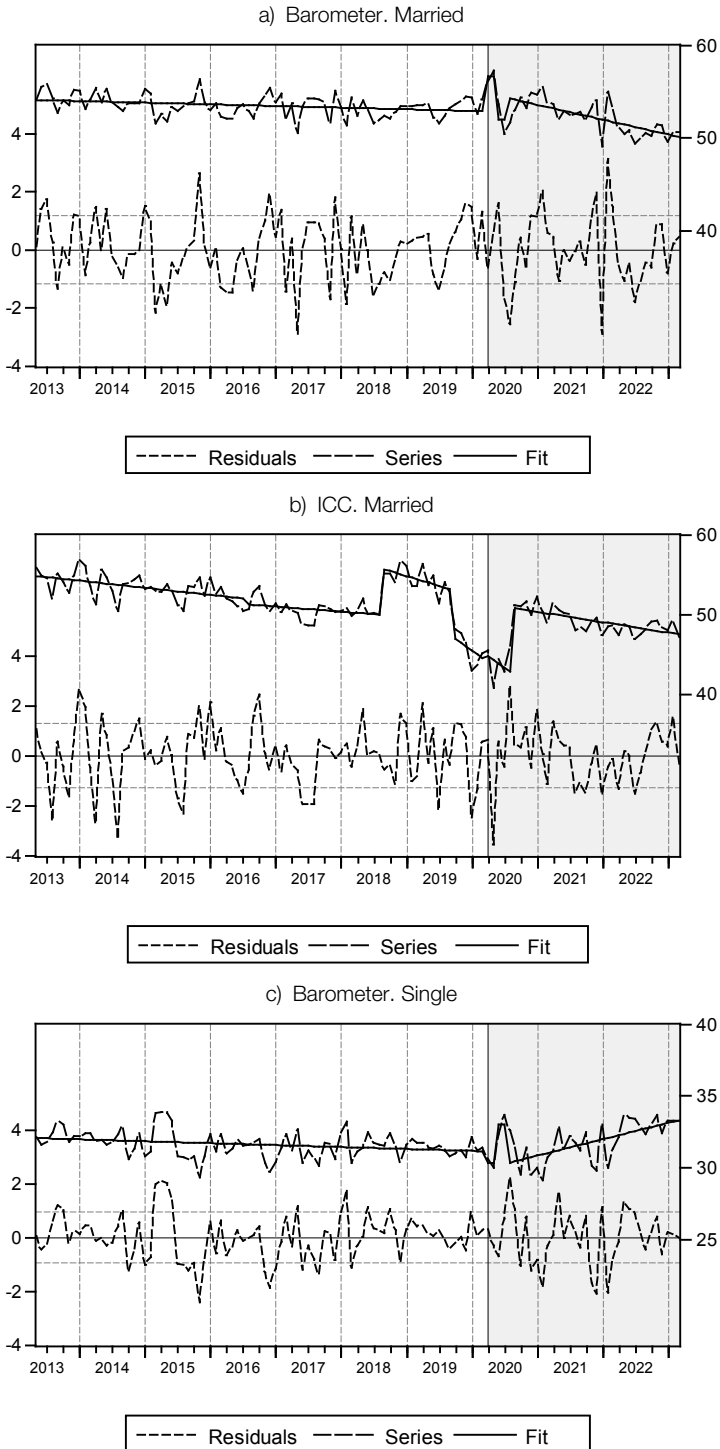
	Married		Single		Others (separated, divorced and widowed)	
	Barometer	ICC	Barometer	ICC	Barometer	ICC
Constant	57.09**	71.87**	34.33**	10.69*	8.49**	16.72**
Slope	-0.01**	-0.07**	-0.01*	0.08**	0.03**	0.00
GfK (03/2015)		-1.88		4.56**		-2.49†
IMOP (08/2016)		-6.74		2.29		6.07
Slope during IMOP		0.02		0.01		-0.03
Sigma Dos-1st phase (09/2018)		30.87		-16.56		-17.56
Slope during Sigma Dos-1st phase		-0.13		0.07		0.08
Sigma Dos-2nd phase (10/2019)		77.69		-21.21		-53.23
Slope during Sigma Dos-2nd phase		-0.40		0.12		0.26†
CATI/COVID-19 (April 2020)	26.18**	7.89	-22.93**	3.90	-5.85	-12.97**
Slope after CATI/COVID-19	-0.12**	-0.03	0.11**	-0.01	0.02	0.05**
Intercampo (04/2020-05/2020)	1.93*		0.28		-2.25**	
IMOP (06/2020-07/2020)	-2.46**		2.79**		0.06	
Corrected R-squared	0.424	0.832	0.252	0.801	0.417	0.690
Durbin Watson	1.731	1.762	1.499	2.072	1.597	1.884
Avg. of dependent variable	53.2	51.0	31.7	34.2	15.0	14.5
Standard deviation	1.54	3.14	1.08	2.28	1.16	1.37
F	18.3**	59.6**	9.0**	48.4**	17.9**	24.0**
N	119	119	119	119	119	119

**p 0.01; *p 0.05 and † p < 0.1.

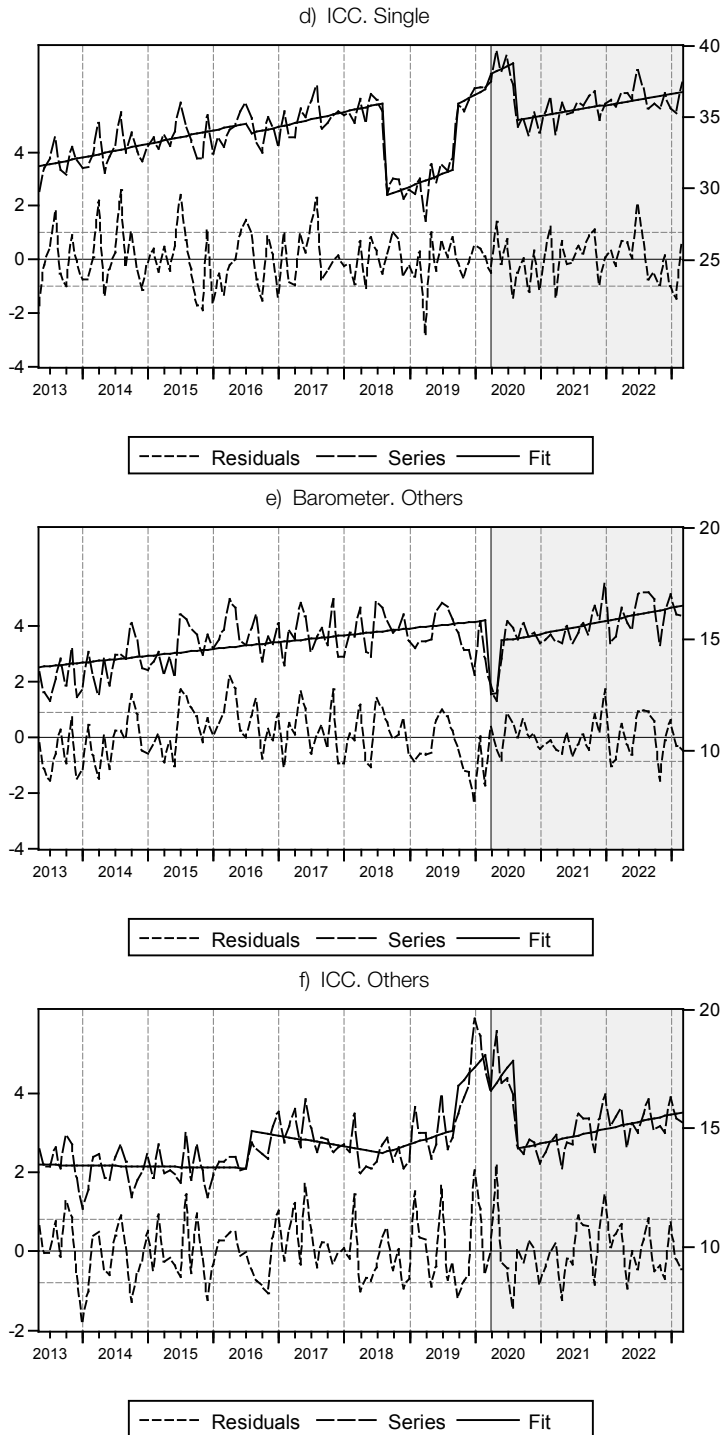
Source: Own elaboration.

⁴ In the marital status category "single" in the barometer we found differences in the variance of the residuals.

GRAPH 3. Time series for the barometer and ICC for marital status, regression model fit and residuals (2013-2022)



GRAPH 3. Time series for the barometer and ICC for marital status, regression model fit and residuals (2013-2022) (Continuation)



Source: Own elaboration.

In Graph 3 we show its representation. We introduce a shaded part starting in April 2020 to facilitate its visualization. As can be seen, the barometer reveals jumps in April 2020 in the categories for married and others, but this only affects the period of 2/4 months in which Intercampo and IMOP carried out the field work (see graph 3a and 3e), returning after to, approximately, their prior level. In the ICC either there are no effects, or they are of limited magnitude (see Graph 3f). The truncations observed correspond to the two Sigma Dos periods. In short, we do not find effects linked to the change in the mode of administration when examining marital status, but we do when looking at the firms carrying out the field work.

For education level the situation is different. In Table 5 we can see that the corrected R-squared is close to one in the lowest and highest education categories. The results are intermediate in the other categories, except in the ICC series for the second stage of secondary education, in which it is practically stationary. The residuals fit normality, except in the two categories of the ICC, without formal education/primary education and the first stage of secondary education (see Graphs 4b and 4c), where we see two peaks of opposite sign. These are homoscedastic, except in the case of university education, where we find a greater dispersion of high values.

As can be seen in Graph 4, at the lowest levels and in higher education there are rectifications in the position of the barometer series that we do not find in the ICC. These jumps, although more intense in the two months of Intercampo, are maintained until the end. That is, they are sustained during the whole period of telephone field-work. In the rest of the categories and in the ICC series, in contrast, we do not find these effects and we only find jumps associated with changes in the firm carrying out the field work (see Graph 4b, 4d and 4h).

In concrete, based on the fit of the regression line, the introduction of CATI in the barometer meant a decline in the groups without formal education-primary education and with obligatory secondary education of 8.2 points and 10 points respectively, as well as an increase in university education of 15.5 points (discounting the temporary effect of Intercampo). If we look at Spain's Labour Force Survey, the proportion of persons with university educations in this same time period grew from 26 % to 31 %, so that if we approach it as a problem of representativeness, we could say that they were under-represented in the PAPI/CAPI period (21.4 %) and over-represented in the CATI period (41.3 %).

This is not a new phenomenon. We find variations of a similar type in methodological study no. 2676 (CIS, 2007), or when we compare telephone surveys commissioned by the CIS for debates on the state of the nation and the barometers for those same months. This effect has also been described on several occasions in national (Díaz-de-Rada, 2010; Galán, Rodríguez and Zorrilla, 2004; Salinas and Aguilar, 2004) and international literature (Fessler, Kasy and Lindner, 2018; Jäckle, Roberts and Lynn, 2006, Lau *et al.*, 2019, St-Pierre and Béland 2004). Groves and Kahn (1979: 94-95) mention three possible explanations: A greater probability that persons with higher educations accept being interviewed by telephone, a greater tendency of the face-to-face field work to not cover these people, and a greater response error in telephone surveys that leads to reporting higher educational achievements than actually reached. Regarding the last, a greater susceptibility to social desirability bias, acquiescence, extreme scores and a recency effect have all been ascribed to CATI (Leeuw and Zouwen, 1988: 283 and ff; Fessler, Kasy and Lindner, 2018, Heerwegh and Loosveldt, 2011; Sloan, Wright and Barrett, 2006; Krosnick *et al.*, 2002; Narayan and Krosnick, 1996), making it complex to determine the origin of this overrepresentation.

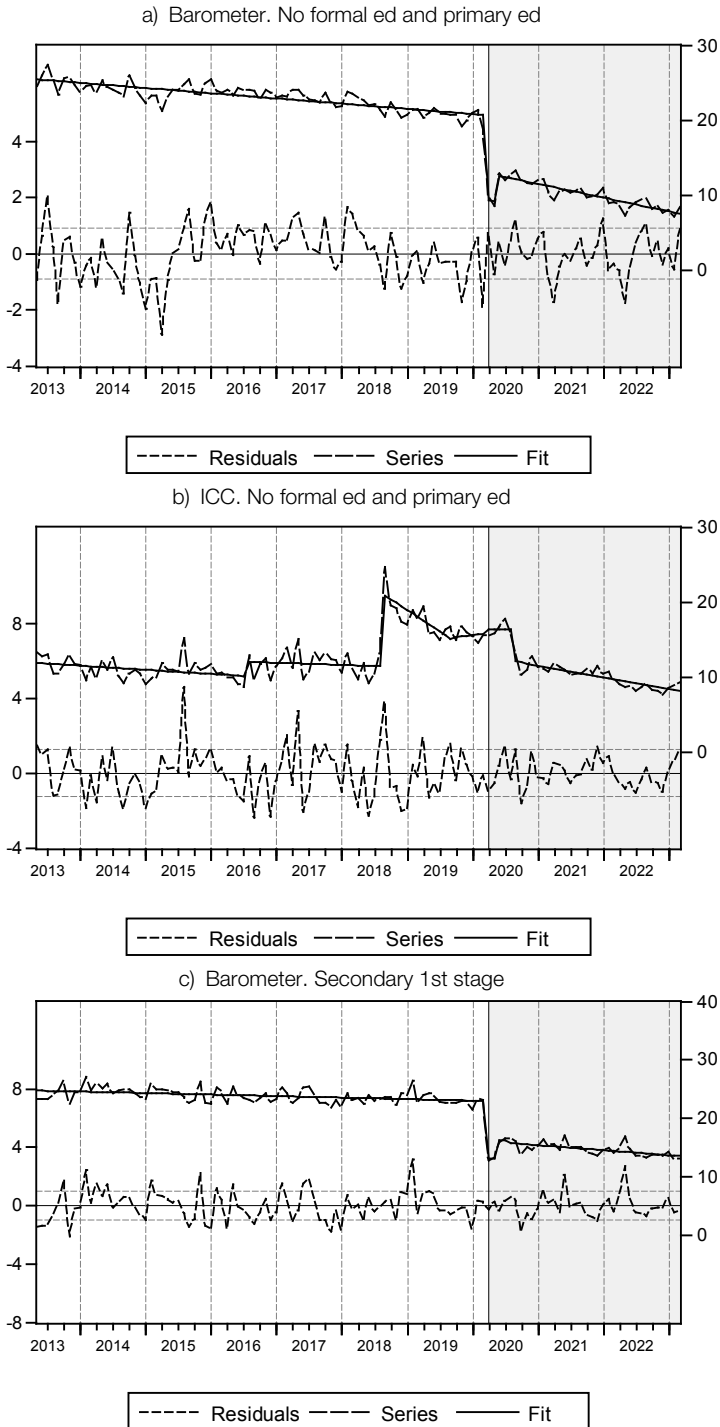
TABLE 5. Regression model for the evolution of education level for the barometer and the ICC (2013-2022)

	No formal ed/primary			Secondary (1st stage)			Secondary (2nd stage)			VT			University		
	Barometer	ICC		Barometer	ICC		Barometer	ICC		Barometer	ICC		Barometer	ICC	
Constant	37.30**	25.36**		29.13**	38.98**		10.63**	12.23*		13.22**	5.60		10.01**	16.53**	
Slope	-0.06**	-0.05*		-0.02**	-0.07**		0.01**	0.01		0.02**	0.04**		0.05**	0.07**	
GfK (03/2015)		-4.13*			3.16			0.95			4.92**			-4.62*	
IMOP (08/2016)		-6.59			5.68			14.37†			-1.47			-11.36	
Slope during IMOP		0.02			-0.01			-0.07†			0.02			0.03	
Sigma Dos-1st phase (09/2018)		96.53**			-112.37**			-15.57			27.12†			-13.26	
Slope during Sigma Dos-1st phase		-0.43**			0.52**			0.08			-0.13†			0.05	
Sigma Dos-2nd phase (10/2019)		-24.63			0.10			-35.69			98.20*			-38.80	
Slope during Sigma Dos-2nd phase		0.13			0.00			0.17			-0.47*			0.16	
CATI/COVID-19 (April 2020)		11.17**			3.05			-9.84**			6.28†			-18.83**	
Slope after CATI/COVID-19		-0.09**			-0.05*			0.06			-0.03*			0.13**	
Intercampo (04/2020-05/2020)		-22.52**			-12.94**			8.68**			-3.58			30.55**	
IMOP (06/2020-07/2020)		-19.21**			-9.74*			7.10*			-4.07			26.60**	
Corrected R-squared	0.979	0.821		0.947	0.909		0.519	0.034		0.523	0.594		0.984	0.938	
Durbin Watson	1.391	1.886		1.783	1.658		1.873	1.692		1.740	1.859		2.072	1.847	
Avg. of dependent variable	19.1	12.1		21.1	22.2		13.8	15.8		18.3	18.4		27.4	31.0	
Standard deviation	6.26	2.99		4.41	5.25		0.98	1.12		1.07	1.58		9.36	5.28	
F	1111.8**	54.9*		425.0**	118.8**		26.5**	1.4		26.9**	18.2**		1488.2**	180.9**	
N	119	119		119	119		119	119		119	119		119	119	

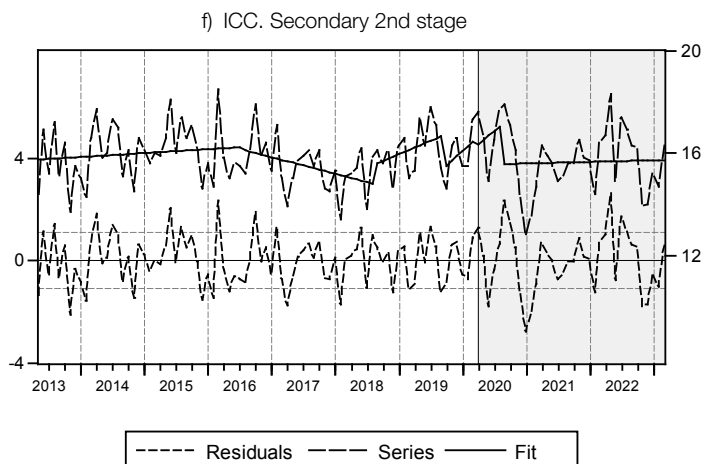
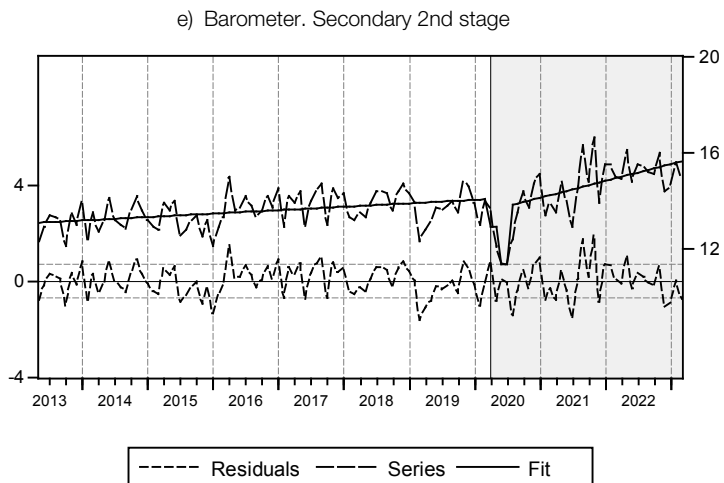
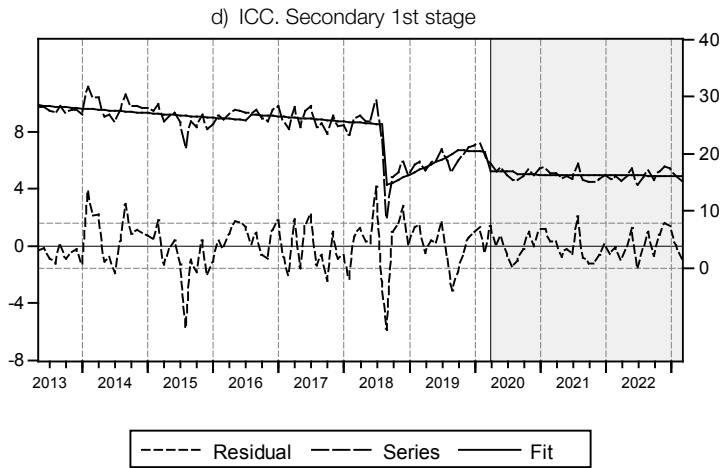
**p 0.01; *p 0.05 and † p < 0.1.

Source: Own elaboration.

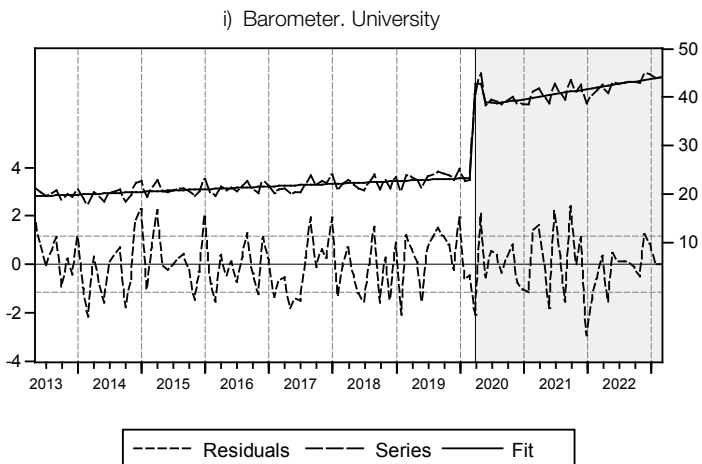
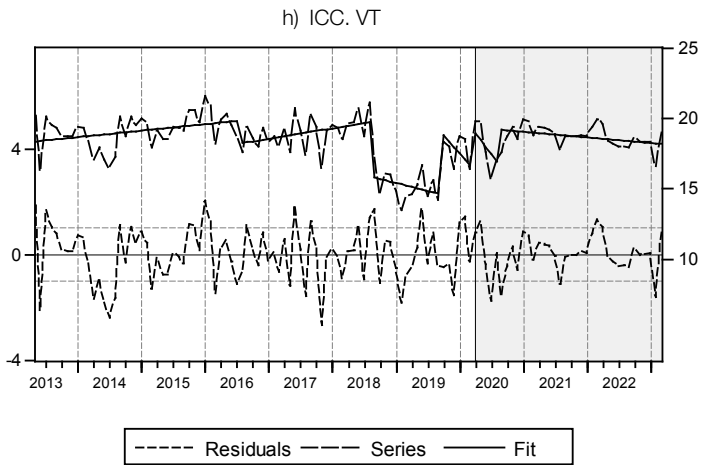
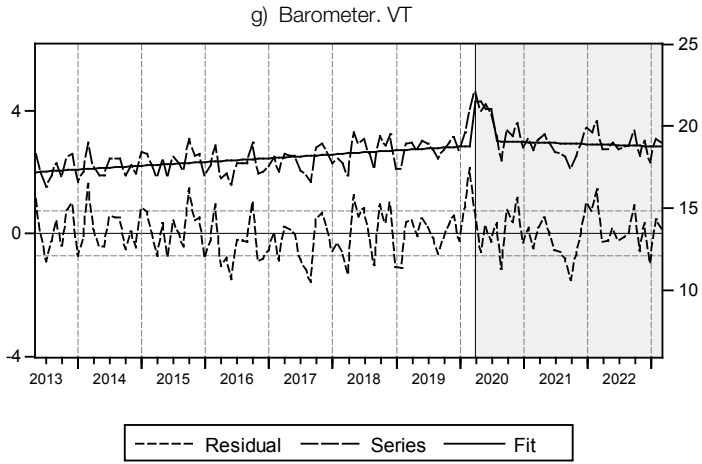
GRAPH 4. Time series for the barometer and the ICC for education level, regression model fit and residuals (2013-2022)



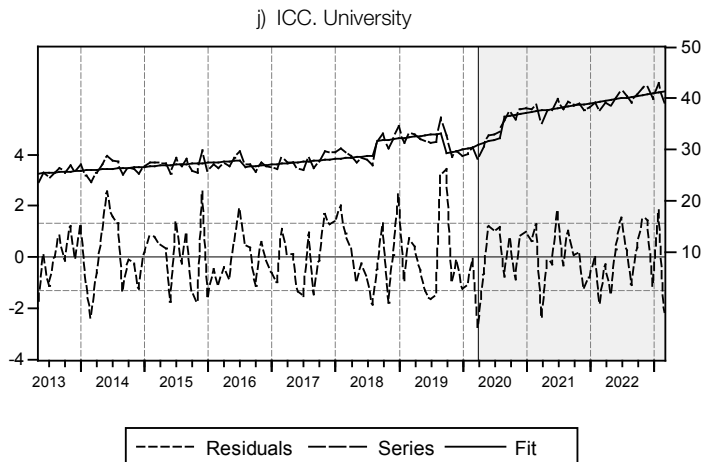
GRAPH 4. Time series for the barometer and the ICC for education level, regression model fit and residuals (2013-2022) (Continuation)



GRAPH 4. Time series for the barometer and the ICC for education level, regression model fit and residuals (2013-2022) (Continuation)



GRAPH 4. Time series for the barometer and the ICC for education level, regression model fit and residuals (2013-2022) (Continuation)



Source: Own elaboration.

CONCLUSIONS

The change in the mode of survey administration breaks the continuity of the data. This is not something new. The modes have substantial differences that prevent their integration or make it more difficult, particularly in terms of mixed designs (Hox, Leeuw and Klausch, 2017). The CIS distributes these series separately on its website, indicating that their direct comparison is not possible.

What we have done, with the aid of time series analysis, is offer a more precise measurement of the impact of the mode change in the particular context of April 2020, with a comparative reference from the ICC. However, we do not find clear evidence that home confinement *per se* had an effect on the series. The results suggest that changing the mode of administration was the main event and that it had clear effects on the values for education level. The question that remains is if these effects are due to modifications in the composition of the samples or to response errors. This study cannot, due to its nature, answer this question, although we

must note the stability found for intermediate education categories.

In addition, the reiteration of a “House effect” for the ICC when the firm carrying out the field work changes, despite using the same sample design, the same questionnaire and the same contract procedures with similar technical prescriptions, suggests that we are, at least partially, dealing with problems related to representativeness. In other words, the effects found for the ICC are most likely to have their cause in the execution of the field work: the provision and use of telephone numbers, the actual volume of calls made, schedules, dealing with postponements and the training of interviewers, are all factors that could have affected the continuity of some of the ICC series.

In short, this study points to a potential vulnerability that needs to be addressed. Until 2020, the CIS used a face to face interview fieldwork network that, given the few variations in procedures, assured a certain stability in the data gathered. Now that CATI is the hegemonic mode, the comparability of the data gathered must be strengthened. The CIS and

TRAGSATEC must maintain a certain stability in their field work procedures, especially regarding those factors that can affect the process of contacting participants and coverage, and that, when it is necessary to introduce changes, that these should be monitored and evaluated through adequate designs.

BIBLIOGRAPHY

- Arnau, Jaume (2001). *Diseños de series temporales. Técnicas de Análisis*. Barcelona. Edicions Universitat de Barcelona.
- Balaguer, Jaime (2022). *El recuerdo de voto en las encuestas. Un análisis exhaustivo a partir del caso español*. Madrid: Consejo Superior de Investigaciones Científicas.
- Bosch, Josep L. and Torrente, Diego (1993). *Encuestas telefónicas y por correo*. Madrid: Centro de Investigaciones Sociológicas.
- CIS (2007). *La percepción social de las encuestas*. Estudio 2676. Banco de Datos del CIS.
- CIS (2018). *Índice de Confianza del Consumidor-Septiembre 2018*. Estudio 3225. Banco de Datos del CIS.
- Díaz-de-Rada, Vidal (2010). *Comparación entre los resultados proporcionados por encuestas telefónicas y personales. El caso de un estudio electoral*. Madrid: Centro de Investigaciones Sociológicas.
- Díaz-de-Rada, Vidal (2019). "Calidad de respuesta y modo de administración del cuestionario". *Revista Internacional de Sociología*, 77: 1-18.
- Dillman, Don; Smyth, Jolene and Christian, Dean (2014). *Internet, Phone, Mail and Mixed Mode Surveys. The Tailored Design Method*. New York: Wiley.
- Fessler, Pirmin; Kasy, Maximilian and Lindner, Peter (2018). "Survey Mode Effects on Measured Income Inequality". *The Journal of Economic Inequality*, 16: 487-505.
- Galán, Iñaki; Rodríguez, Fernando and Zorrilla, Belén (2004). "Comparación entre encuestas telefónicas y encuestas 'cara a cara' domiciliarias en la estimación de hábitos de salud y prácticas preventivas". *Gaceta Sanitaria*, 18: 440-450.
- Groves, Robert and Kahn, Robert (1979). *Surveys by Telephone. A National Comparison with Personal Interviews*. London: Academic Press.
- Heerwegh, Dirk and Loosveldt, Geert (2011). "Assessing Mode Effects in a National Crime Victimization Survey using Structural Equation Models. Social Desirability Bias and Acquiescence". *Journal of Official Statistics*, 27: 49-63.
- Holbrook, Allyson; Green, Melanie and Krosnick, Jon (2003). "Telephone versus Face-to-face Interviewing of National Probability Samples with Long Questionnaires. Comparisons of Respondent Satisficing and Social Desirability Response Bias". *Public Opinion Quarterly*, 67: 79-125.
- Hox, Joop; Leeuw, Edith de and Klausch, Thomas (2017). Mixed-Mode Research. Issues in Design and Analysis. In: P. Biemer; E. D. de Leeuw; S. Eckman; B. Edwards; F. Kreuter; L. E. Lyberg; N. C. Tucker and B. T. West (eds.). *Total Survey Error in Practice*. New Jersey: John Wiley & Sons Inc.
- Jäckle, Annette; Roberts, Caroline and Lynn, Peter (2006). "Telephone versus Face-to-face Interviewing. Mode Effects on Data Quality and Likely Causes". *ISER Working Paper Series*, 41. Colchester: University of Essex.
- Krosnick, Jon (2002). "The Impact of 'No Opinion' Response Options on Data Quality. Non-attitude Reduction or an Invitation to Satisfice?". *Public Opinion Quarterly*, 66: 371-403.
- Lau, Charles; Cronberg, Alexandra; Marks, Leenisha and Amaya, Ashley (2019). "In Search of the Optimal Mode for Mobile Phone Surveys in Developing Countries. A Comparison of IVR, SMS, and CATI in Nigeria". *Survey Research Methods*, 13: 305-318.
- Leeuw, Edith de and Zouwen, Johannes van der (1988). Data Quality in Telephone and Face to Face Surveys. A Comparative Meta-Analysis. In: R. Groves; W. L. Nicholls; J. T. Massey; J. Waksberg; P. P. Biemer and L. E. Lyberg (eds.). *Telephone Survey Methodology*. New York: Wiley.
- Narayan, Sowmya and Krosnick, Jon (1996). "Education Moderates some Response Effects in Attitude Measurement". *Public Opinion Quarterly*, 60: 58-88.
- Peña, Daniel (2005). *Análisis de Series Temporales*. Madrid: Alianza Editorial.
- Salinas, José M. and Aguilar, María del C. (2004). "Un análisis comparativo entre la entrevista telefónica y la entrevista presencial en la determinación de la prevalencia de los juegos de azar". *Metodología de Encuestas*, 6: 119-132.
- St-Pierre, Martin and Béland, Yves (2004). "Mode effects in the Canadian Community Health Survey. A Comparison of CAPI and CATI". *2004 Pro-*

ceedings of the American Statistical Association Meeting. Survey Research Methods. Toronto: American Statistical Association.

Sloan, Matt; Wright, Debra and Barrett, Kirsten (2006). "Data Comparability in a Mixed Mode Telephone and Face to Face Survey of Persons

with Disabilities". *Mathematica Policy Research Reports*.

Tourangeau, Roger and Smith, Tom (1996). "Asking Sensitive Questions. The Impact of Data Collection Mode. Question Format and Question Context". *Public Opinion Quarterly*, 60: 275-304.

RECEPTION: February 27, 2023

REVIEW: April 27, 2023

ACCEPTANCE: June 07, 2023

El cambio al modo de administración telefónico en el barómetro del CIS. Una comparación con las series temporales del ICC

*The Switch to Telephone Survey Mode in the CIS Barometer.
A Comparison with ICC Time Series*

Jaime Balaguer

Palabras clave

Centro de Investigaciones Sociológicas
 • Encuestas presenciales
 • Encuestas telefónicas
 • Metodología de recolección de datos

Key words

Centre for Sociological Research
 • Face-to-Face Surveys
 • Telephone Surveys
 • Data Collection Methodology

Resumen

En abril de 2020 el Centro de Investigaciones Sociológicas (CIS) cambió el modo de administración de sus encuestas, pasando del personal (PAPI/CAPI) al telefónico (CATI). En el presente trabajo examinamos su impacto en dos variables del barómetro: el estado civil y el nivel de estudios. Describimos su evolución mensual desde 2013 a 2023 utilizando el Índice de Confianza del Consumidor (ICC), un estudio continuo que siempre ha sido telefónico, como referencia comparativa. Como veremos, el cambio de modo produce discontinuidades en la medición del nivel educativo, con implicaciones para la calidad e interpretación de resultados. Además, la evidencia apunta a que las empresas privadas que el CIS ha contratado para realizar los trabajos de campo telefónicos del ICC presentan sesgos propios que deben tomarse en cuenta para la comparación y seguimiento de las series.

Abstract

In April 2020, Spain's Centre for Sociological Research (CIS) changed its survey mode from face-to-face (PAPI/CAPI) to telephone-based (CATI). In this study we examine the impact of this change on two *CIS barometer* variables: marital status and education level. We describe their evolution from 2013 to 2023 using the Consumer Confidence Index (ICC), a continuous survey that has always been telephone-based, as a comparative reference. As we will see, the change in mode produced discontinuities in the measurement of education level, with implications for the quality and interpretation of results. In addition, the evidence suggests that the private research companies the CIS contracted to carry out telephone fieldwork had their own biases that must be taken into account for series comparison and monitoring.

Cómo citar

Balaguer, Jaime (2024). «El cambio al modo de administración telefónico en el barómetro del CIS. Una comparación con las series temporales del ICC». *Revista Española de Investigaciones Sociológicas*, 185: 145-162. (doi: 10.5477/cis/reis.185.145-162)

La versión en inglés de esta nota puede consultarse en <http://reis.cis.es>

Jaime Balaguer: Ministerio de Hacienda y Función Pública | jtbalaguer@sepg.hacienda.gob.es



INTRODUCCIÓN

En abril de 2020 el CIS cambia su modo habitual de administración de las encuestas, pasando de las entrevistas cara a cara (PAPI/CAPI) a las telefónicas (CATI). En un contexto marcado por la pandemia de la COVID-19, en el que la vía telefónica resulta una sustitución razonable, externaliza durante algunos meses la realización de encuestas a las empresas Intercampo e IMOP. En septiembre de 2020 formaliza su apuesta por el CATI estableciendo una infraestructura de campo telefónico permanente a través de TRAGSATEC (sociedad mercantil estatal considerada medio propio de las Administraciones públicas).

El cambio de modo de administración supone una transformación organizativa y metodológica que, tomando en cuenta la difusión de las encuestas del CIS, es de interés para la comunidad investigadora. La vía telefónica presenta las ventajas de la rapidez, el menor coste, la facilidad para supervisar entrevistas, la alta profesionalización del personal entrevistador, la elevada dispersión de la muestra y la posibilidad de incrementar los intentos de contactación. Por contra, implica afrontar el importante volumen de incidencias telefónicas, la limitación en la entrevista al uso de estímulos auditivos, la posible mayor distracción de las personas entrevistadas y una merma en la calidad general de las respuestas (Bosch y Torrente, 1993: 10-22; Díaz-de-Rada, 2019; Dillman, Smyth y Christian, 2014; Holbrook, Green y Krosnick, 2003; Tourangeau y Smith, 1996).

Ante una transición de este calado, habría sido de interés realizar un estudio metodológico exhaustivo que permitiera evaluar su impacto en dos elementos clave: la representatividad de las muestras y la presencia de errores de medición. Sin embargo, la crisis de la COVID-19 precipitó la adopción del CATI —que después devino en permanente— sin que pudieran aplicarse estas cautelas. Pese a esta carencia, contamos con

elementos útiles para describir sus efectos. En este trabajo nos centramos en el barómetro, uno de los estudios más notorios del CIS que, desde 1996, recoge mensualmente información sobre el estado de la opinión pública española. Lo comparamos con el Índice de Confianza del Consumidor (ICC), otro proyecto mensual que estudia las intenciones de gasto de la población residente y del cual el CIS es responsable desde 2011. El hecho de que el ICC siempre haya sido telefónico lo convierte en una referencia singular para valorar los efectos del cambio de modo. Pero para poder establecer una adecuada comparación debemos, en primer lugar, plantear una propuesta de análisis consistente.

PROPUESTA METODOLÓGICA

Dos eventos importantes para nuestro objeto de estudio coinciden en el tiempo: el cambio al modo telefónico del barómetro y el confinamiento domiciliario por la crisis de la COVID-19. Esto supone que, cuando comparamos los barómetros de marzo y de abril de 2020 y atendemos a las variables que se repiten en ambos, los cambios que observamos pueden tener tres orígenes distintos: el impacto de la pandemia en la opinión pública; el cambio en el modo de administración (que pudo afectar tanto a la representatividad de la muestra como al error de respuesta); y el confinamiento domiciliario (y la desocupación temporal de una parte importante de la población), que pudo afectar a la probabilidad de contactación o de aceptar la entrevista (al estar las personas seleccionadas más disponibles en los hogares) y, en consecuencia, incidir temporalmente en la representatividad de las muestras.

Para separar estos elementos empleamos la siguiente estrategia. Por un lado, trabajamos con un período de observación amplio. El cambio de modo de administración (al contrario que el confinamiento) es un fenómeno permanente. Por tanto, bus-

camos cambios duraderos, esto es, diferencias entre la información recogida antes y después de 2020. Para ello, podemos identificar variables que se pregunten mensualmente, observar su evolución y discriminar entre tendencias a largo plazo (pendientes), variaciones irregulares y truncamientos de la constante o saltos de nivel (*jumps*) que marquen una discontinuidad en las series de datos. En concreto, la presencia de un salto de nivel, coincidente con el paso al modo telefónico, que se mantenga en años posteriores, indicaría un cambio en la estructura de la serie de datos que nos permitiría hablar de etapas diferenciadas (Peña, 2005: 377-382; Arnau, 2001: 97 y ss.).

Por otro lado, descartamos utilizar las variables actitudinales y de opinión, en las que la crisis de la COVID-19 pudo tener un impacto relevante (así como la situación laboral). Requieren de un análisis diferenciado que excede los propósitos de esta investigación. Nos centramos en variables que describen parámetros relativamente estables de la población. Descartamos el sexo y la edad por conformar las cuotas para la selección de las personas entrevistadas. Tampoco empleamos el recuerdo de voto, pues si bien refiere una característica fija de las personas entrevistadas (que solo debería cambiar con la celebración de nuevas elecciones), como ha puesto de manifiesto de manera exhaustiva Balaguer (2022), presenta un importante error de respuesta en el que la coyuntura política tiene un peso significativo. Como resultado de esta criba, nos quedan dos variables disponibles: el estado civil y el nivel de estudios alcanzado. Ambas describen características que cambian lentamente en la población y que no deberían haber experimentado variaciones relevantes durante la crisis de la COVID-19.

Finalmente, utilizamos el ICC como referencia comparativa. El ICC ha sido siempre un estudio telefónico, por lo que no está afectado por el cambio de modo. Si el confinamiento implicó variaciones en la proba-

bilidad de contactación o de aceptar la encuesta, con efectos en el estado civil o en el nivel de estudios, estas deberían reflejarse en las series del ICC. Si el ICC no experimentó cambios en abril de 2020, o estos fueron reducidos, entonces los saltos de nivel que encontremos en el barómetro podríamos asociarlos al cambio de modo de administración. De esta manera, nuestra hipótesis principal (H1) propone que *en abril de 2020 los saltos de nivel en las series del estado civil y el nivel de estudios del barómetro serán de mayor magnitud que los observados en el ICC.*

DESCRIPCIÓN DE LAS SERIES

Del barómetro contamos con datos desde 1996; del ICC desde 2011. Al respecto, en mayo de 2013 el ICC empezó a incluir en sus muestras a personas que solo tienen teléfono móvil. Esto supone un cambio sustantivo, ya que afecta a la cobertura de la población estudiada¹. Por tanto, la establecemos como fecha de inicio, concluyendo en marzo de 2023, con un total de 119 observaciones mensuales, 83 anteriores a abril de 2020 y 36 posteriores².

En la tabla 1 tenemos la media y la desviación típica de las dos series en el barómetro. Podemos observar que, con el cambio al CATI, el estado civil permanece relativamente estable, mientras que la distribución del nivel de estudios varía sustancialmente. En particular, se reducen los niveles más bajos y se incrementan notablemente los estudios superiores.

¹ Según la Encuesta sobre Equipamiento y Uso de Tecnologías de la Información y Comunicación del INE, en 2013 el 21,4 % de los hogares no disponía de teléfono fijo, alcanzándose el 37,3 % en 2022.

² El barómetro se dirige a población española de 18 y más años, y el ICC a población residente de 16 o más. Esta diferencia puede afectar a los promedios de las series, pero no a la tendencia general. En el barómetro las observaciones faltantes de los meses de agosto y los macrobarómetros de 2019 han sido reemplazadas por el promedio de las puntuaciones adyacentes.

TABLA 1. *Medias y desviación estándar de las series del estado civil y del nivel de estudios en el barómetro, según modo de administración (2013-2022)*

	PAPI/CAPI (05/2013-03/2020)		CATI (04/2020-03/2023)	
	Media	Desv.est.	Media	Desv.est.
Estado civil				
Casado/a	53,5	1,14	52,4	2,02
Soltero/a	31,6	0,90	31,8	1,41
Separado/a	2,5	0,28	2,0	0,28
Divorciado/a	4,6	0,62	6,6	0,73
Viudo/a	7,7	0,67	6,9	0,62
Nivel de estudios				
Sin estudios	5,6	0,60	2,8	0,70
Primarios	17,5	1,55	7,2	1,28
Secundarios (1.ª etapa)	23,8	1,18	14,6	1,16
Secundarios (2.ª etapa)	13,6	0,68	14,4	1,28
FP	18,0	0,89	19,2	1,01
Superiores	21,4	1,58	41,3	2,08
N	83		36	

Fuente: Elaboración propia a partir del Banco de Datos del CIS.

Hay que tener en cuenta que las primeras cuatro observaciones del barómetro empleando CATI fueron realizadas por empresas privadas contratadas por el CIS. Si bien estas empresas cumplen con unos criterios sobre medios humanos y materiales exigidos para su contratación (las ofertas de las empresas resultan homogéneas en una serie de características, como el número mínimo de intentos de contactación, el porcentaje de supervisión o la cantidad de medios propios empleados), en la práctica, se trata de propuestas genéricas. Las variaciones en la carga de teléfonos, en el número y distribución horaria de los intentos de contactación o en la experiencia del personal asignado, pueden conducir a diferencias en los datos recolectados. Hablamos aquí de un «efecto casa», definido como el sesgo propio de la empresa o entidad que realiza la encuesta, una «caja negra» que puede implicar variaciones en

la representatividad y la calidad de la información.

Este factor resulta todavía más relevante en el ICC. De las 119 observaciones que analizamos, 88 fueron realizadas por empresas privadas, correspondiendo las restantes al CIS-TRAGSATEC. En concreto, 39 las realizó la empresa GfK (entre mayo de 2013 y agosto de 2016), 25 IMOP (hasta agosto de 2018) y 24 Sigma Dos (hasta agosto de 2020). En la tabla 2 tenemos los descriptivos para el ICC desagregados según la empresa adjudicataria del trabajo de campo. En el nivel de estudios resulta destacable el incremento de los primarios con Sigma Dos, la caída de los estudios secundarios y la evolución creciente de los superiores, en particular en la última etapa. En el estado civil se observa una relativa estabilidad, si bien las mediciones de Sigma Dos presentan una desviación típica más elevada en las categorías de casados/as y solteros/as.

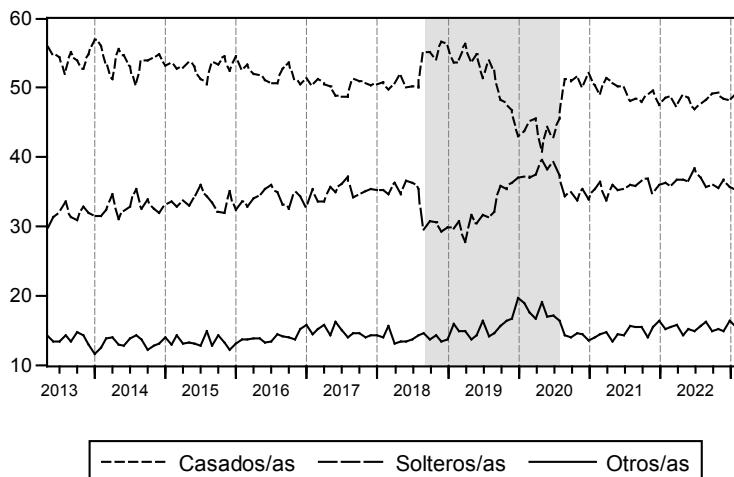
TABLA 2. *Medias y desviación estándar de las series del estado civil y nivel de estudios en el ICC, según la empresa que realizó el trabajo de campo (2013-2022)*

	GFK (05/2013-07/2016)		IMOP (08/2016-08/2018)		Sigma Dos (09/2018-08/2020)		CIS-TRAGSATEC (09/2020-03/2023)	
	Media	Desv. est.	Media	Desv. est.	Media	Desv. est.	Media	Desv. est.
Estado civil								
Casado/a	53,4	1,61	50,6	1,12	50,0	5,18	49,2	1,40
Soltero/a	33,0	1,45	34,9	1,15	33,6	3,72	35,8	1,04
Separado/a	2,4	0,42	2,3	0,39	2,8	0,42	2,2	0,25
Divorciado/a	3,9	0,37	4,6	0,45	5,8	1,39	6,2	0,69
Viudo/a	7,1	0,49	7,5	0,60	7,2	0,55	6,5	0,50
Nivel de estudios								
Sin estudios	3,3	0,64	5,2	0,71	4,96	0,89	2,9	0,39
Primarios	7,7	1,13	6,6	1,15	12,1	1,81	7,2	1,27
Secundarios (1.ª etapa)	27,2	1,83	26,0	1,73	17,9	2,66	16,2	0,95
Secundarios (2.ª etapa)	16,0	1,09	15,5	0,99	16,2	0,98	15,6	1,27
FP	19,1	1,19	19,0	1,11	16,4	1,79	18,7	0,73
Superiores	26,5	1,46	27,7	1,27	31,6	1,98	38,9	1,83
	39		25		24		31	

Fuente: Elaboración propia a partir del Banco de Datos del CIS.

Esta anomalía en la varianza del estado civil durante el período de Sigma Dos requiere de una mirada en detalle. En el gráfico 1 representamos su evolución. Podemos distinguir dos etapas en la medición de esta em-

presa: una primera, de septiembre de 2018 a septiembre de 2019, en la que la proporción de casados asciende y la de solteros desciende, y una segunda, de octubre de 2019 a agosto de 2020, con la oscilación inversa.

GRÁFICO 1. *Series temporales del ICC para el estado civil (2013-2023). En sombreado, la etapa de Sigma Dos*

Fuente: Elaboración propia a partir del Banco de Datos del CIS.

Una posible explicación está en las estrategias de contactación. En septiembre de 2018 se produce una modificación en la ficha técnica. A partir de ese momento se añade la siguiente aclaración: «Las personas contactadas por teléfono móvil que también tengan teléfono fijo se consideran incorporadas a la muestra de fijo + móvil» (CIS, 2018). Esto supone, en la práctica, que el número de personas entrevistadas por móvil sube de aproximadamente el 20 %, a una cifra que varía en torno al 70 %.

Al respecto, hemos podido examinar en detalle las incidencias del campo telefónico del ICC de agosto de 2016 a marzo de 2020³. Como podemos observar en la tabla 3, con la llegada de Sigma Dos se produce un incremento notable de los teléfonos empleados, particularmente de los móviles. La ratio de llamadas realizadas/

téfonos usados desciende. Sin embargo, esa estrategia no se mantiene fija, cambia. A partir de noviembre de 2019 se modera el uso de teléfonos y se incrementa el volumen de llamadas, con lo que la ratio se vuelve a elevar.

Estas dos etapas podrían estar vinculadas a la varianza de la serie del estado civil. En concreto, la ratio llamadas/téfonos presenta un coeficiente de correlación de Pearson de $-0,786$ con la proporción de personas casadas y de $+0,870$ con las solteras. Como puede apreciarse en el gráfico 2, podemos hablar de una covariación relevante entre ambos fenómenos. Aunque no podemos ser resolutivos ni mucho menos establecer una relación de causalidad, lo que sí haremos es tomar en cuenta, en el análisis que sigue a continuación, la existencia de dos etapas diferenciadas en el trabajo de campo de Sigma Dos.

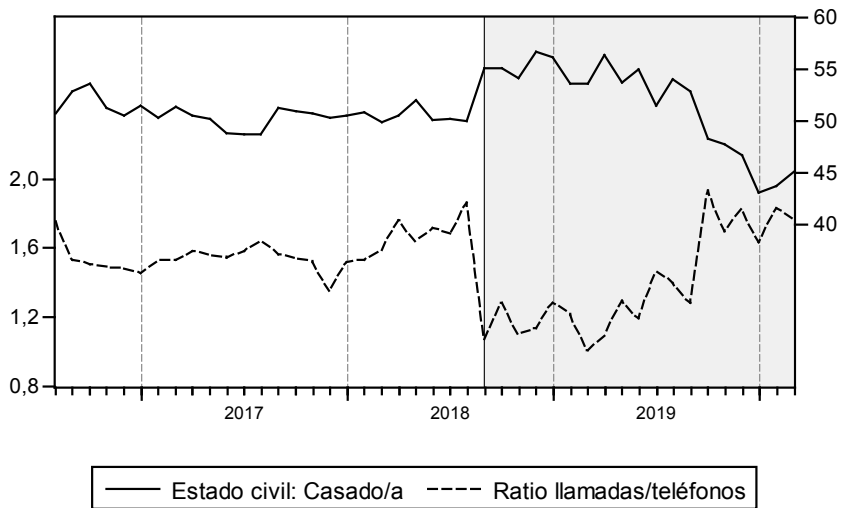
TABLA 3. Promedio mensual de teléfonos usados y de llamadas realizadas en los trabajos de campo del ICC (2016-2020)

	IMOP (08/2016-08/2018)	Sigma Dos – 1.ª etapa (09/2018-09/2019)	Sigma Dos – 2.ª etapa (10/2019-03/2020)
Teléfonos usados			
Fijos	35.475	47.381	41.520
Móviles	14.331	51.585	44.486
Total	49.806	98.966	86.006
Llamadas realizadas			
Fijos	51.283	49.627	65.301
Móviles	27.542	71.643	88.033
Total	78.826	121.269	153.334
Ratio llamadas/téfonos	1,58	1,23	1,78

Fuente: Elaboración propia a partir de los Informes de campo del ICC.

³ Queremos agradecer a Valentín Martínez, responsable de muestras del CIS, su ayuda en esta parte.

GRÁFICO 2. Proporción de estado civil casado/a y ratio llamadas/teléfonos en el trabajo de campo del ICC (2016-2020). En sombreado, la etapa de Sigma Dos



Fuente: Elaboración propia a partir de los Informes de campo del ICC.

RESULTADOS

Para poder establecer las pendientes y los posibles saltos de nivel (*jumps*) en las series implementamos un modelo determinista, habida cuenta que de nuestra finalidad es describir una tendencia pasada y no predecir su comportamiento (Peña, 2005: 44-45). Utilizamos la regresión lineal, estimada mediante mínimos cuadrados e implementada con el software EViews 12. Por razones de simplificación fusionamos las categorías con las menores frecuencias del estado civil (separado/a + divorciado/a + viudo/a) y del nivel de estudios (sin estudios + estudios primarios). Como parámetros introducimos una constante, la pendiente y variables *dummy* que identifican los períodos en que las empresas mencionadas realizaron los trabajos de campo (con dos segmentos para Sigma Dos). Nuestra variable independiente principal es una *dummy* que identifica el período que comienza en abril de 2020. Establecemos

también si se producen modificaciones en las pendientes.

En la tabla 4 tenemos los resultados de los modelos para el estado civil. Hay que tener en cuenta que, en el barómetro, el salto de nivel de abril de 2020 coincide con los encargos a Intercampo, con lo que una lectura directa de los coeficientes puede resultar confusa. Identificamos su magnitud posteriormente, a la vista de la recta de ajuste de la regresión. La R-cuadrado corregida es relativamente baja en varias de las categorías de respuesta, lo que unido a la normalidad y homocedasticidad observada en los residuos (test Jarque-Bera y test de White)⁴ apunta a que, una vez descontados los efectos de la ecuación, puede haber un importante componente del término error (Peña, 2005: 319-321).

⁴ En la categoría soltero/a del barómetro sí se observaron diferencias en la varianza de los residuos.

TABLA 4. Modelo de regresión para la evolución del estado civil en el barómetro y el ICC (2013-2022)

	Casados/as		Solteros/as		Otros/as (separados/as, divorciados/as y viudos/as)	
	Barómetro	ICC	Barómetro	ICC	Barómetro	ICC
Constante	57,09**	71,87**	34,33**	10,69*	8,49**	16,72**
Pendiente	-0,01**	-0,07**	-0,01*	0,08**	0,03**	0,00
GfK (03/2015)		-1,88		4,56**		-2,49†
IMOP (08/2016)		-6,74		2,29		6,07
Pendiente durante IMOP		0,02		0,01		-0,03
Sigma Dos-1. ^a etapa (09/2018)		30,87		-16,56		-17,56
Pendiente durante Sigma Dos-1. ^a etapa		-0,13		0,07		0,08
Sigma Dos-2. ^a etapa (10/2019)		77,69		-21,21		-53,23
Pendiente durante Sigma Dos-2. ^a etapa		-0,40		0,12		0,26†
CATI/COVID-19 (abril 2020)	26,18**	7,89	-22,93**	3,90	-5,85	-12,97**
Pendiente tras CATI/COVID-19	-0,12**	-0,03	0,11**	-0,01	0,02	0,05**
Intercampo (04/2020-05/2020)	1,93*		0,28		-2,25**	
IMOP (06/2020-07/2020)	-2,46**		2,79**		0,06	
R cuadrado Corregido	0,424	0,832	0,252	0,801	0,417	0,690
Durbin Watson	1,731	1,762	1,499	2,072	1,597	1,884
Media de la V. dependiente	53,2	51,0	31,7	34,2	15,0	14,5
Desviación típica	1,54	3,14	1,08	2,28	1,16	1,37
F	18,3**	59,6**	9,0**	48,4**	17,9**	24,0**
N	119	119	119	119	119	119

**p 0,01; *p 0,05 y † p < 0,1.

Fuente: Elaboración propia.

En el gráfico 3 mostramos su representación. Introducimos un sombreado a partir de abril de 2020 para facilitar su visualización. Como puede observarse, el barómetro registra saltos de nivel en abril de 2020 en la categoría de casado/a y en otros/as, pero solo afectan al período de dos/cuatro meses en que Intercampo e IMOP realizan el campo (véanse gráficos 3a y 3e), reto-

mando después de manera aproximada el nivel anterior. En el ICC, o bien no hay efectos o bien son de reducida magnitud (véase gráfico 3f). Los truncamientos observables corresponden a las dos etapas de Sigma Dos. En definitiva, en el estado civil no encontramos efectos vinculados al cambio de modo, pero sí a las empresas ejecutoras del campo.

GRÁFICO 3. Series temporales del barómetro e ICC del estado civil. Ajuste del modelo de regresión y residuos (2013-2022)

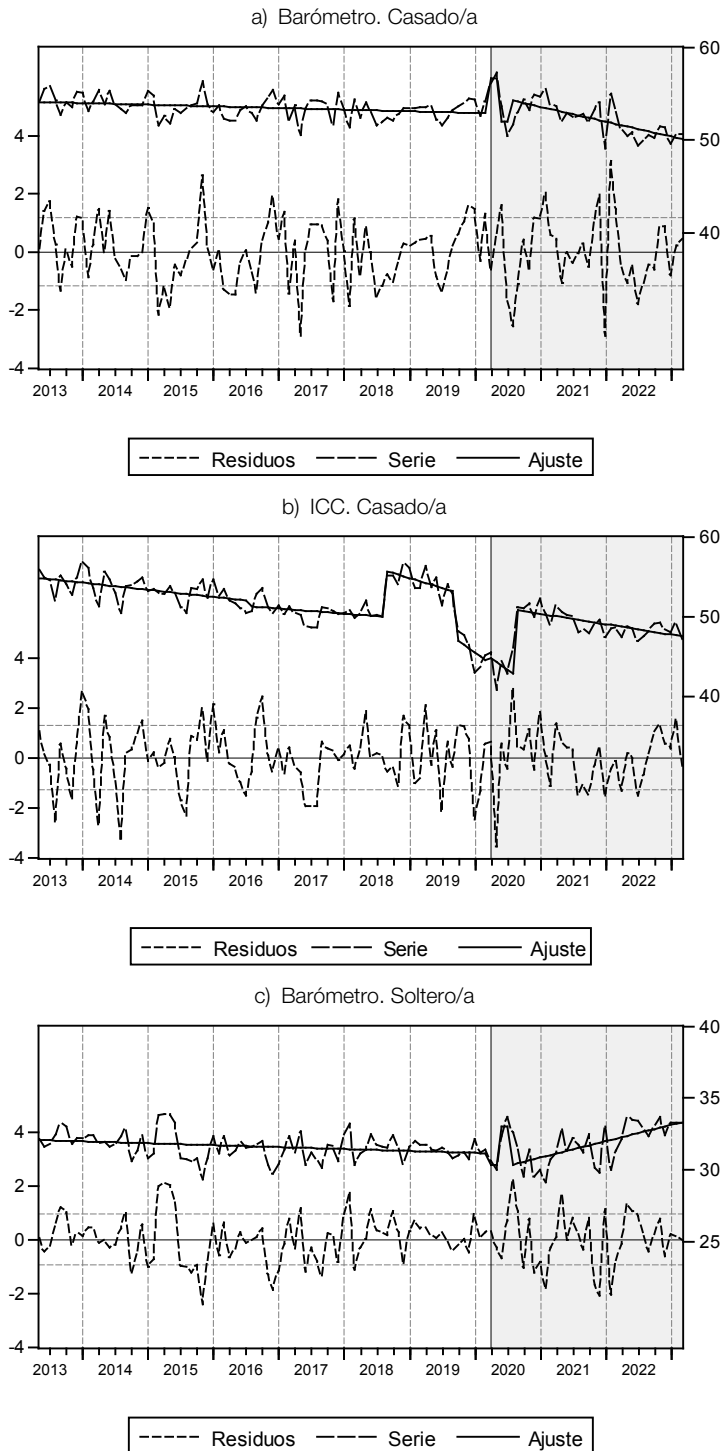
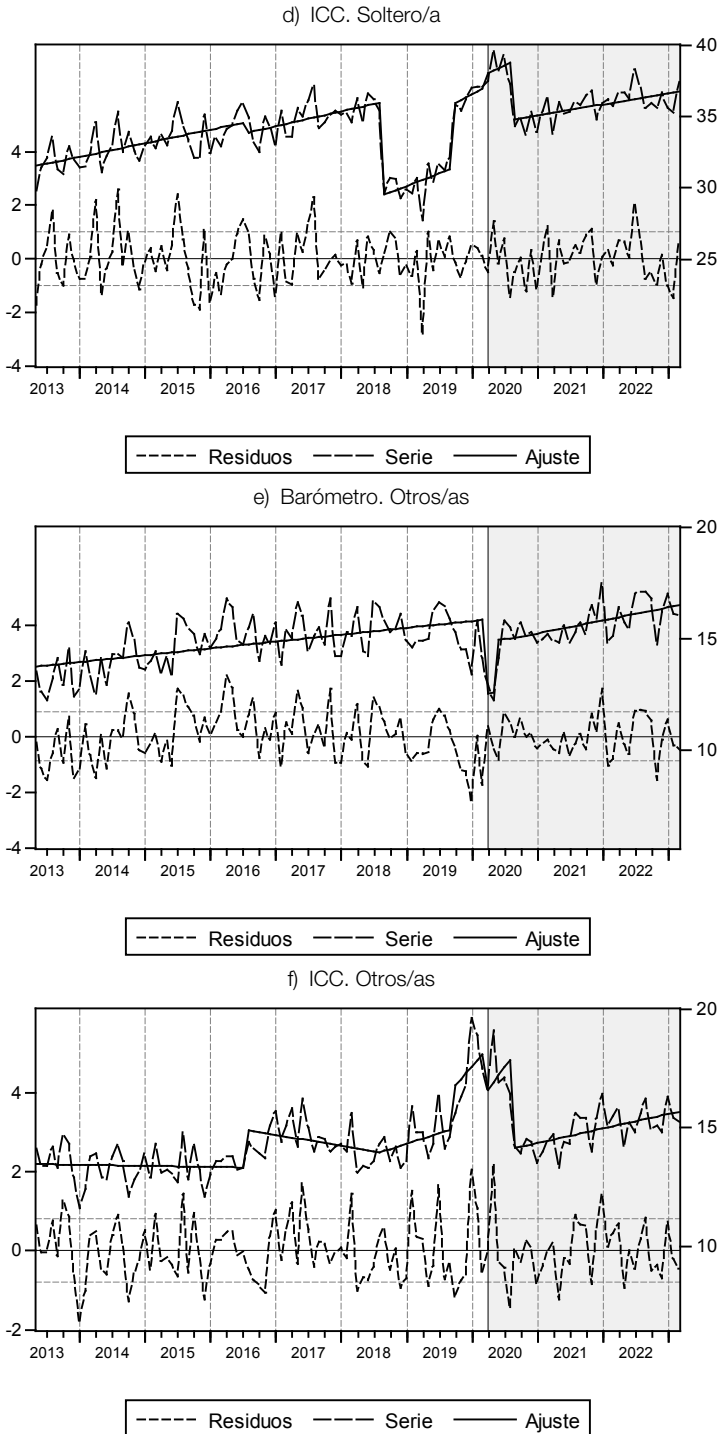


GRÁFICO 3. Series temporales del barómetro e ICC del estado civil. Ajuste del modelo de regresión y residuos (2013-2022) (continuación)



Fuente: Elaboración propia.

TABLA 4. Modelo de regresión para la evolución del estado civil en el barómetro y el ICC (2013-2022)

	Sin estudios y primarios		Secundaria (1.ª etapa)		Secundaria (2.ª etapa)		FP		Superiores	
	Barómetro	ICC	Barómetro	ICC	Barómetro	ICC	Barómetro	ICC	Barómetro	ICC
Constante	37,30**	25,36**	29,13**	38,98**	10,63**	12,23*	13,22**	5,60	10,01**	16,53**
Pendiente	-0,06**	-0,05*	-0,02**	-0,07**	0,01**	0,01	0,02**	0,04**	0,05**	0,07**
GfK (03/2015)		-4,13*		3,16		0,95		4,92**		-4,62*
IMOP (08/2016)		-6,59		5,68		14,37†		-1,47		-11,36
Pendiente durante IMOP		0,02		-0,01		-0,07†		0,02		0,03
Sigma Dos-1.ª etapa (09/2018)		96,53**		-112,37**		-15,57		27,12†		-13,26
Pendiente durante Sigma Dos-1.ª etapa		-0,43**		0,52**		0,08		-0,13†		0,05
Sigma Dos-2.ª etapa (10/2019)		-24,63		0,10		-35,69		98,20*		-38,80
Pendiente durante Sigma Dos-2.ª etapa		0,13		0,00		0,17		-0,47*		0,16
CATI/COVID-19 (abril 2020)		11,17**		3,05		-15,31†		6,28†		-11,19*
Pendiente tras CATI/COVID-19		-0,09**		-0,05*		0,06		-0,03*		0,13**
Intercampo (04/2020-05/2020)		-22,52**		-12,94**		8,68**		-3,58		30,55**
IMOP (06/2020-07/2020)		-19,21**		-9,74*		7,10*		-4,07		26,60**
R cuadrado		0,979		0,947		0,909		0,523		0,984
Corregido		1,391		1,783		1,658		1,740		2,072
Durbin Watson		19,1		21,1		22,2		18,3		27,4
Media de la V. dependiente		6,26		4,41		5,25		1,07		9,36
Desviación típica		1111,8**		425,0**		118,8**		26,9**		1488,2**
F		119		119		119		119		119
N		119		119		119		119		119

**p 0,01; *p 0,05 y † p < 0,1.

Fuente: Elaboración propia.

En el nivel de estudios la situación es diferente. En la tabla 5 podemos observar que el R-cuadrado corregido se aproxima a la unidad en las categorías de estudios más bajos y superiores. Su rendimiento es intermedio en las demás categorías, salvo en la serie del ICC de estudios secundarios de 2.^a etapa, en que es prácticamente estacionaria. Los residuos se ajustan a la normalidad, salvo en dos categorías del ICC: sin estudios-primarios y secundarios de 1.^a etapa (véanse gráficos 4b y 4c), donde observamos dos picos de signo contrario en ambas. Estos son homocedásticos, salvo en el caso de los estudios superiores,

donde encontramos mayor dispersión en los valores altos.

Como puede observarse en el gráfico 4, en los niveles más bajos y en los estudios superiores se producen rectificaciones de la posición de la serie del barómetro que no encontramos en el ICC. Estos truncamientos, aunque más intensos en los dos meses de Intercampo, se mantienen hasta el final. Es decir, se sostienen durante toda la etapa de campo telefónico. En el resto de categorías y en la serie del ICC, en cambio, no encontramos estos efectos o solo observamos saltos asociados a los cambios de empresa (véanse gráficos 4b, 4d y 4h).

GRÁFICO 4. Series temporales del barómetro e ICC del nivel de estudios, ajuste del modelo de regresión y residuos (2013-2022)

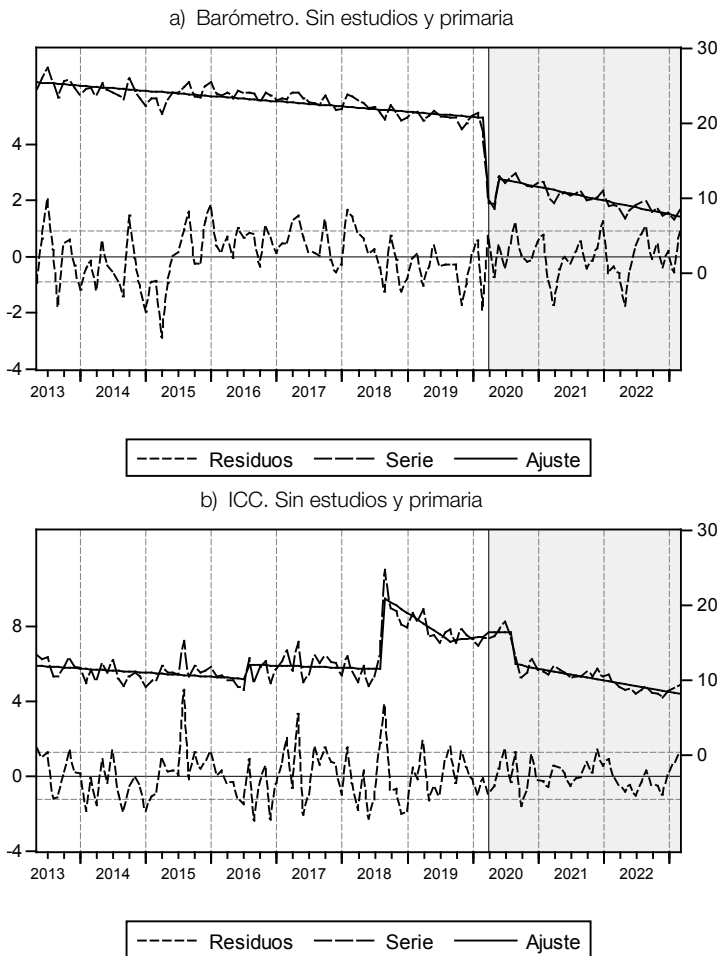


GRÁFICO 4. Series temporales del barómetro e ICC del nivel de estudios, ajuste del modelo de regresión y residuos (2013-2022) (continuación)

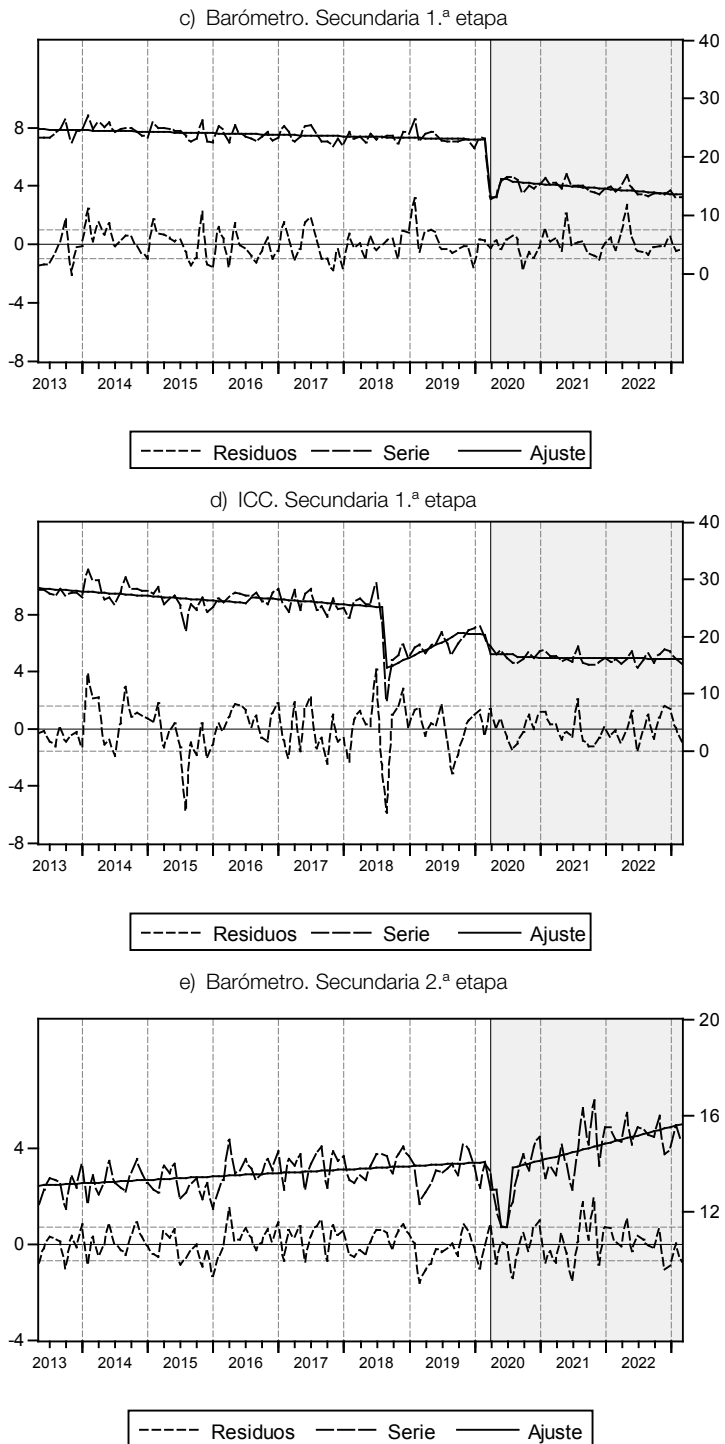


GRÁFICO 4. Series temporales del barómetro e ICC del nivel de estudios, ajuste del modelo de regresión y residuos (2013-2022) (continuación)

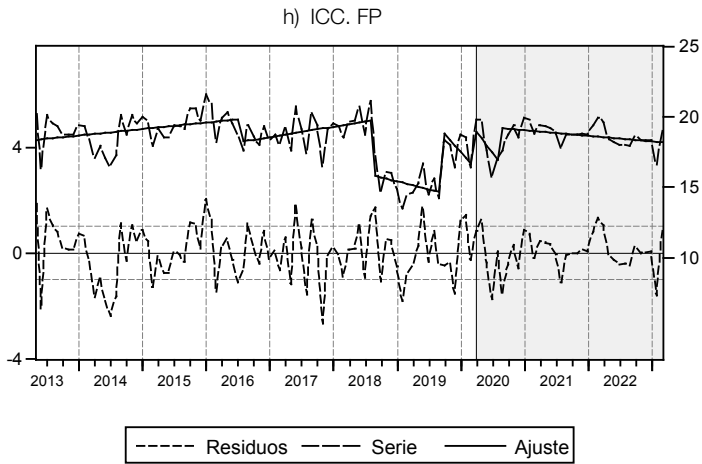
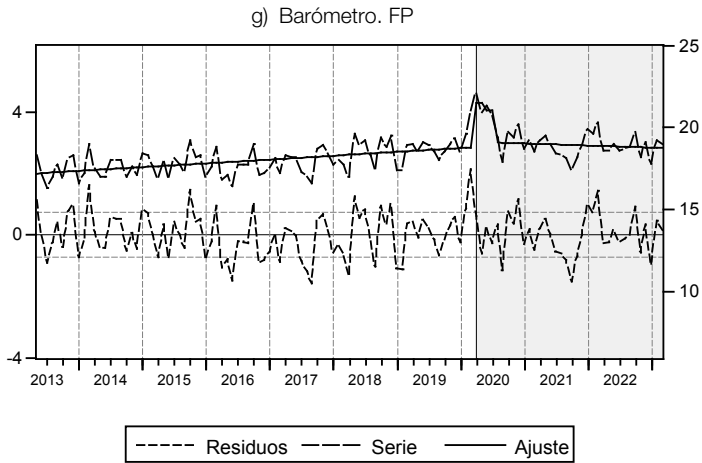
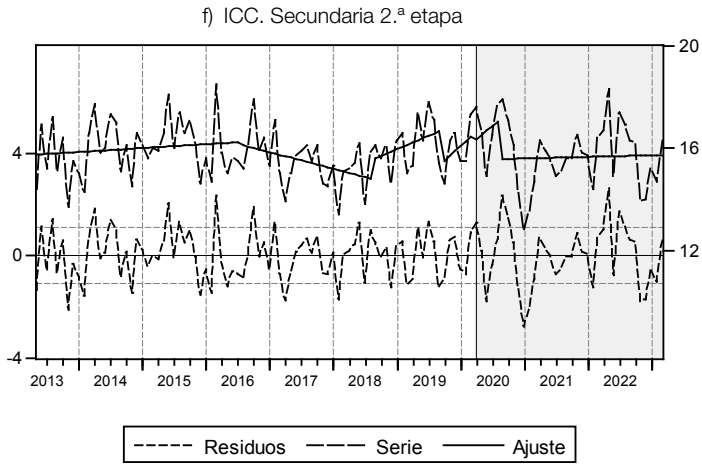
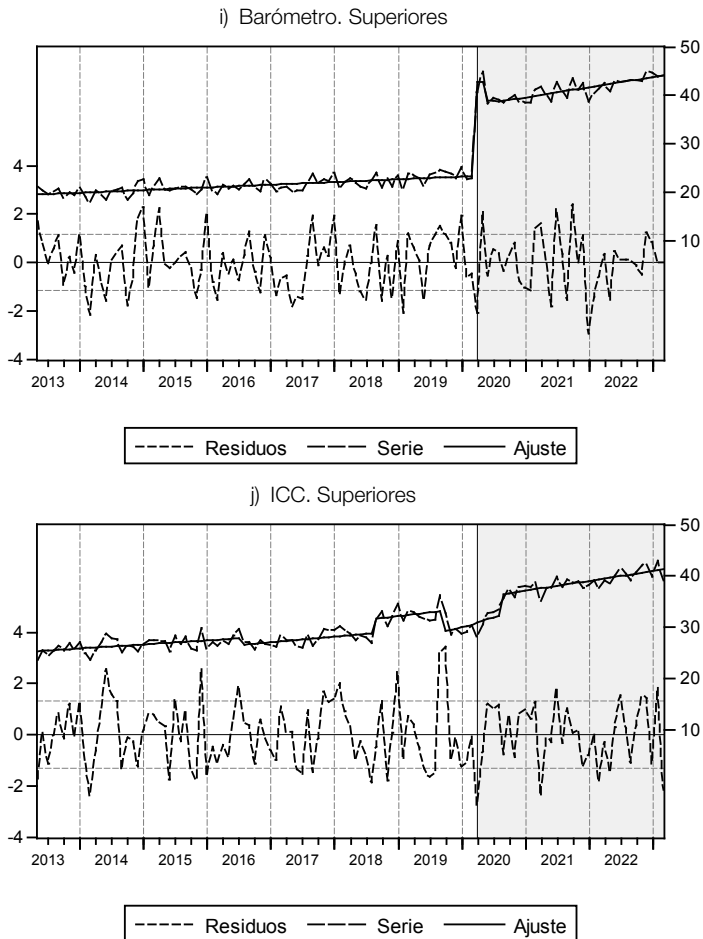


GRÁFICO 4. Series temporales del barómetro e ICC del nivel de estudios, ajuste del modelo de regresión y residuos (2013-2022) (continuación)



Fuente: Elaboración propia.

En concreto, según la línea de ajuste de la regresión, la introducción del CATI en el barómetro supuso un retroceso de los grupos sin estudios-primarios y de los secundarios de 1.^a etapa de 8,2 y 10 puntos respectivamente, así como un incremento de los estudios superiores de 15,5 puntos (descontado el efecto temporal de Intercampo). Si atendemos a la Encuesta de Población Activa, la proporción con estudios superiores en este mismo período creció del 26 % al 31 %, por lo que, si lo enfocásemos como un problema de representatividad, podríamos decir que estaban infrarrepresentados

en la etapa PAPI/CAPI (21,4 %) y sobrerrepresentados en el CATI (41,3 %).

Este fenómeno no es una novedad. Encontramos variaciones del mismo signo en el estudio metodológico n.º 2676 (CIS, 2007), o cuando comparamos las encuestas telefónicas encargadas por el CIS para los debates sobre el estado de la nación y los barómetros de esos mismos meses. Este efecto también se ha descrito en varias ocasiones en la literatura nacional (Díaz-de-Rada, 2010; Galán, Rodríguez y Zorrilla, 2004; Salinas y Aguilar, 2004) e internacional (Fessler, Kasy y Lindner, 2018; Jäckle, Roberts y Lynn, 2006;

Lau *et al.*, 2019; St-Pierre y Béland, 2004). Al respecto, Groves y Kahn (1979: 94-95) mencionan tres posibles explicaciones: una mayor probabilidad de las personas de estudios altos de aceptar la entrevista telefónica, una mayor tendencia del campo personal a no cubrir a estas personas, o un mayor error de respuesta en el teléfono que lleva a reportar logros educativos superiores a los alcanzados. Respecto a esto último, se ha descrito para el CATI (no de manera unívoca) una mayor exposición al sesgo de deseabilidad social, a la aquiescencia, a puntuaciones extremas o al efecto recencia (Leeuw y Zouwen, 1988: 283 y ss.; Fessler, Kasy y Lindner, 2018; Heerwegh y Loosveldt, 2011; Sloan, Wright y Barrett, 2006; Krosnick, 2002; Narayan y Krosnick 1996), por lo que resulta complejo determinar su origen.

CONCLUSIONES

El cambio de modo de administración interrumpe la continuidad de los datos. Esto no es ninguna novedad. Los modos presentan diferencias sustanciales que impiden su integración o la dificultan notablemente, en particular en el marco de diseños mixtos (Hox, Leeuw y Klausch, 2017). El propio CIS distribuye en su web las series separadamente, evidenciando que no es posible su comparación directa.

Lo que hemos hecho, con la ayuda del análisis de series temporales, es ofrecer una medición más precisa del impacto del cambio de modo en el particular contexto de abril de 2020, con una referencia comparativa en el ICC. Al respecto, no encontramos evidencias claras de que el confinamiento domiciliario *per se* tuviera un efecto en las series. Los resultados apuntan a que el cambio de modo fue el evento principal y que afectó de manera clara a los valores del nivel de estudios. La pregunta que queda pendiente es si estos efectos se deben a modificaciones en la composición de las muestras

o a errores de respuesta. Este trabajo no puede, debido a su naturaleza, contestarla, aunque llama la atención la estabilidad de las categorías de estudios intermedios.

Por otro lado, la reiteración de un «efecto casa» en el ICC cuando cambia la empresa adjudicataria, pese a emplearse el mismo diseño muestral, el mismo cuestionario y procedimientos de contratación con prescripciones técnicas similares, apunta a que sí que tratamos, al menos parcialmente, con problemas de representatividad. Es decir, los efectos encontrados en el ICC tendrían su causa más probable en la ejecución del campo: la provisión y uso de los números de teléfono, el volumen real de llamadas realizadas, los horarios, la ejecución de los aplazamientos o el entrenamiento del personal entrevistador/a son elementos que podrían afectar a la continuidad de algunas de sus series.

En definitiva, este trabajo apunta a una vulnerabilidad potencial a la que es necesario atender. Hasta 2020 el CIS venía utilizando una red de campo personal que, por las escasas variaciones de sus procedimientos, le aseguraba una cierta estabilidad a sus datos. Ahora que el CATI es el modo hegemónico, se necesita reforzar la comparabilidad de la información recogida en las distintas mediciones. Para ello es clave que la labor del CIS y de TRAGSATEC mantenga cierta estabilidad en sus formas de actuación, especialmente en aquellos elementos que puedan afectar a la contactación y a la cobertura, y que, cuando sea necesario introducir cambios, estos se monitoricen y se evalúen con los diseños adecuados.

BIBLIOGRAFÍA

- Arnau, Jaume (2001). *Diseños de series temporales. Técnicas de Análisis*. Barcelona: Edicions Universitat de Barcelona.
- Balaguer, Jaime (2022). *El recuerdo de voto en las encuestas. Un análisis exhaustivo a partir del caso español*. Madrid: Consejo Superior de Investigaciones Científicas.

- Bosch, Josep L. y Torrente, Diego (1993). *Encuestas telefónicas y por correo*. Madrid: Centro de Investigaciones Sociológicas.
- CIS (2007). *La percepción social de las encuestas*. Estudio 2676. Banco de Datos del CIS.
- CIS (2018). *Índice de Confianza del Consumidor-Septiembre 2018*. Estudio 3225. Banco de Datos del CIS.
- Díaz-de-Rada, Vidal (2010). *Comparación entre los resultados proporcionados por encuestas telefónicas y personales. el caso de un estudio electoral*. Madrid: Centro de Investigaciones Sociológicas.
- Díaz-de-Rada, Vidal (2019). «Calidad de respuesta y modo de administración del cuestionario». *Revista Internacional de Sociología*, 77: 1-18.
- Dillman, Don; Smyth, Jolene y Christian, Dean (2014). *Internet, Phone, Mail and Mixed Mode Surveys. The Tailored Design Method*. New York: Wiley.
- Fessler, Pirmin; Kasy, Maximilian y Lindner, Peter (2018). «Survey Mode Effects on Measured Income Inequality». *The Journal of Economic Inequality*, 16: 487-505.
- Galán, Iñaki; Rodríguez, Fernando y Zorrilla, Belén (2004). «Comparación entre encuestas telefónicas y encuestas “cara a cara” domiciliarias en la estimación de hábitos de salud y prácticas preventivas». *Gaceta Sanitaria*, 18: 440-450.
- Groves, Robert y Kahn, Robert (1979). *Surveys by Telephone. A National Comparison with Personal Interviews*. London: Academic Press.
- Heerwegh, Dirk y Loosveldt, Geert (2011). «Assessing Mode Effects in a National Crime Victimization Survey Using Structural Equation Models. Social Desirability Bias and Acquiescence». *Journal of Official Statistics*, 27: 49-63.
- Holbrook, Allyson; Green, Melanie y Krosnick, Jon (2003). «Telephone versus Face-to-face Interviewing of National Probability Samples with Long Questionnaires. Comparisons of Respondent Satisficing and Social Desirability Response Bias». *Public Opinion Quarterly*, 67: 79-125.
- Hox, Joop; Leeuw, Edith de y Klausch, Thomas (2017). Mixed-Mode Research. Issues in Design and Analysis. En: P. Biemer; E. D. de Leeuw; S. Eckman; B. Edwards; F. Kreuter; L. E. Lyberg; N. C. Tucker y B. T. West (eds.). *Total Survey Error in Practice*. New Jersey: John Wiley & Sons Inc.
- Jäckle, Annette; Roberts, Caroline y Lynn, Peter (2006). «Telephone versus Face-to-face Interviewing. Mode Effects on Data Quality and Likely Causes». *ISER Working Paper Series*, 41. Colchester: University of Essex.
- Krosnick, Jon (2002). «The Impact of “No Opinion” Response Options on Data Quality. Non-attitude Reduction or an Invitation to Satisfice?». *Public Opinion Quarterly*, 66: 371-403.
- Lau, Charles; Cronberg, Alexandra; Marks, Leenisha y Amaya, Ashley (2019). «In Search of the Optimal Mode for Mobile Phone Surveys in Developing Countries. A Comparison of IVR, SMS, and CATI in Nigeria». *Survey Research Methods*, 13: 305-318.
- Leeuw, Edith de y Zouwen, Johannes van der (1988). Data Quality in Telephone and Face To Face Surveys. A Comparative Meta-Analysis. En: R. Groves; W. L. Nicholls; J. T. Massey; J. Waksberg; P. P. Biemer y L. E. Lyberg (eds.). *Telephone Survey Methodology*. New York: Wiley.
- Narayan, Sowmya y Krosnick, Jon (1996). «Education Moderates Some Response Effects in Attitude Measurement». *Public Opinion Quarterly*, 60: 58-88.
- Peña, Daniel (2005). *Análisis de Series Temporales*. Madrid: Alianza Editorial.
- Salinas, José M. y Aguilar, María del C. (2004). «Un análisis comparativo entre la entrevista telefónica y la entrevista presencial en la determinación de la prevalencia de los juegos de azar». *Metodología de Encuestas*, 6: 119-132.
- St-Pierre, Martin y Béland, Yves (2004). «Mode effects in the Canadian Community Health Survey. A Comparison of CAPI and CATI». *2004 Proceedings of the American Statistical Association Meeting*. Survey Research Methods. Toronto: American Statistical Association.
- Sloan, Matt; Wright, Debra y Barrett, Kirsten (2006). «Data Comparability in a Mixed Mode Telephone and Face to Face Survey of Persons with Disabilities». *Mathematica Policy Research Reports*.
- Tourangeau, Roger y Smith, Tom (1996). «Asking Sensitive Questions. The Impact of Data Collection Mode. Question Format and Question Context». *Public Opinion Quarterly*, 60: 275-304.

RECEPCIÓN: 27/02/2023

REVISIÓN: 27/04/2023

APROBACIÓN: 07/06/2023

