



Reconstrucción y consistencia factorial: la regla del codo aplicada al RMSEA, análisis paralelo y otras pruebas confirmatorias

RABADÁN-PÉREZ, FRANCISCO

Universidad Rey Juan Carlos (España)

Correo electrónico: francisco.rabadan@urjc.es

BERUMEN, SERGIO A.

Universidad Rey Juan Carlos (España)

Correo electrónico: sergio.berumen@urjc.es

GUIANCE-LAPIDO, JAIME

PricewaterhouseCoopers Auditores

Correo electrónico: jaime.guance.lapido@pwc.com

HERNÁNDEZ MORA, CRISTÓBAL

Universidad Rey Juan Carlos (España)

Correo electrónico: c.hernandezmora@alumnos.urjc.es

RESUMEN

La comparación de encuestas presenta problemas metodológicos recurrentes, dentro de los que destacan particularmente: i) las dificultades asociadas a la falta de continuidad, bien por el añadido de nuevas preguntas o por la eliminación de otras; ii) las dificultades propias en la selección de las variables a comparar; iii) la necesidad de recodificación entre cuestionarios no coincidentes, y iv) la discrecionalidad en las unidades de medida de las variables. Así, para la investigación en métodos cuantitativos se precisan procedimientos que permitan recuperar la estructura factorial, realizar comparaciones entre variables y recodificar las unidades de medida y en la escala, todo ello en distintos periodos. Para evaluar en qué medida se mantiene la estructura factorial, en el presente artículo se proponen medidas de bondad del ajuste y contrastes propios del análisis factorial confirmatorio, a lo que se añade una aplicación novedosa del principio de parsimonia a los estadísticos RMSEA, FIT y BIC, a través de la conocida regla del codo en el gráfico de sedimentación. La propuesta metodológica se valida a través de dos publicaciones del Centro de Investigaciones Sociológicas (CIS) de distintos años y centradas en el fenómeno del fraude fiscal.

Palabras clave: análisis multivariante, análisis factorial exploratorio, máxima-verosimilitud, comparación de estructuras factoriales, gráfico de sedimentación.

Clasificación JEL: C19; C38; H26.

MSC2010: 62H25; 62P25.

Reconstruction and Factorial Consistence: the Elbow Rule Applied to RMSEA, Parallel Analysis and other Confirmatory Techniques

ABSTRACT

The comparison of surveys presents recurring methodological problems, among which the following stand out particularly: i) the difficulties associated with the change of the questions, either due to the addition of new ones, changing in the mode that are asked, or the elimination of others; ii) the own difficulties selecting the variables to be compared; iii) the need for recoding between mismatched surveys, and iv) the different measurement scales of the variables. Thus, in order to carry on research in quantitative methods, procedures are required to allow for the factorial structure, to compare variables, and to recode the measurement and the scale, all of them in different periods. To evaluate the extent to which the factorial structure is maintained, this work proposes measures of goodness of fit, attached to the confirmatory factorial analysis, where we can find also an original interpretation of the parsimony principle with the RMSEA, FIT and BIC, through the well-known elbow rule in the scree plot. The methodological proposal is validated through two publications, from different years, of the Sociological Research Center (CIS, by its Spanish acronym) and focused on the phenomenon of fiscal fraud.

Keywords: multivariate analysis, exploratory factor analysis, maximum Likelihood, comparison of factor structures, RMSEA scree plot.

JEL classification: C19; C38; H26.

MSC2010: 62H25; 62P25.



1. Introducción

La comparación de encuestas presenta problemas metodológicos recurrentes, dentro de los que destacan particularmente: i) las dificultades asociadas a la falta de continuidad en las preguntas, bien por exceso o por defecto, es decir, por el añadido de nuevas preguntas o por la eliminación de otras; ii) las dificultades propias en la selección de las variables a comparar; iii) la necesidad de recodificación entre cuestionarios no coincidentes, y iv) la discrecionalidad en las unidades de medida de las variables, lo que dificulta la comparación adimensional por los cambios ocurridos. Así, para la investigación en métodos cuantitativos se hace necesario disponer de procedimientos que permitan recuperar la estructura factorial y que posibiliten la realización de comparaciones entre variables, la recodificación en las unidades de medida y en la escala, todo ello en distintos periodos.

Para comprobar la consistencia factorial y así evaluar en qué medida se mantiene la estructura factorial, en este trabajo se recurre a medidas de bondad del ajuste y contrastes propios del análisis factorial confirmatorio y de las ecuaciones estructurales (SEM), a lo que se añade una interpretación novedosa al principio de parsimonia a los estadísticos RMSEA (*Root Mean Error of Aproximation*), FIT (*Factor Fit of the Complete Model*) y BIC (*Bayesian Information Criterion*), parecida a la selección del número de factores a través de la conocida regla del codo en el gráfico de sedimentación.

La presente investigación formula una propuesta metodológica centrada en la estructura factorial entre dos publicaciones del Centro de Investigaciones Sociológicas (CIS): i) el estudio *CIS N° 2910, Opinión Pública y Política Fiscal (XXVIII)*, de 2011, y ii) el estudio *CIS N° 3259, Opinión Pública y Política Fiscal (XXXVI)*, de 2019, ambas centradas en el fenómeno del fraude fiscal. La selección de los años se justifica por el interés de comparar dos momentos muy distintos; por un lado, de recesión económica severa (año 2011), y por el otro de crecimiento acelerado (año 2019). En ambas las preguntas de los cuestionarios son variables Likert, o variables observadas de manera ordinal que proceden de variables continuas a nivel poblacional (Likert, 1932).

2. Metodología

La primera condición para realizar inferencia estadística sobre una población determinada es disponer de un tamaño muestral suficiente y representativo, si bien frecuentemente faltan datos y atípicos. Los datos faltantes en alguna o algunas de las variables privan de información y reducen el número de casos completos; sin embargo, para la extracción factorial basta con que el número de casos completos sea suficientemente grande para la extracción y la generalización en términos de inferencia (Fabrigar et al., 1999; Ferrando & Anguiano-Carrasco, 2010; Lloret-Segura et al., 2014).

De acuerdo con Peña (2002, pp. 120 y 121), se denominan datos atípicos a aquellas observaciones que parecen haberse generado de forma distinta al resto de los datos y, asimismo, las consecuencias de una sola observación atípica pueden ser graves: distorsionar las medias y desviaciones típicas de las variables y destruir las relaciones existentes entre ellas. Por tanto, en la presente investigación se ha optado por la identificación de los atípicos de acuerdo al criterio de la distancia de Mahalanobis, y su exclusión en análisis posteriores a la reducción de dimensiones cuando $sig \leq 10^{-3}$ (Smith, 1997). Los resultados del análisis factorial, tratamiento de perdidos, test de normalidad para el estudio CIS3269 se encuentran en el anexo: “Anexo 1. Análisis Factorial y pruebas de normalidad CIS3259.spv”.

Las técnicas de reducción de dimensiones tienen la capacidad de identificar un conjunto de variables y ponderarlas por su capacidad para explicar la varianza del modelo, varianza total de las variables consideradas; en segundo lugar, se obtienen variables latentes (no observadas) que el método hace que sean ortogonales respecto de la primera extraída que es la que tiene mayor capacidad explicativa, denominada componente o factor de fuerza, dependiendo de la técnica, lo que evita sobreponderar dimensionalmente una característica poblacional al impedir que más de una variable mida la misma característica. En definitiva, obtener una estructura dimensional lo suficientemente

simple para explicar la realidad encuestada a través de la extracción y la rotación más adecuada (Thurstone, 1947).

Hay dos tipos de técnicas de reducción de dimensiones ampliamente utilizadas en psicología (Bjelland et al., 2002), economía (Rahmer et al., 2020) y biología genética (Jombart et al., 2010), entre otras disciplinas:

- El Análisis de Componentes Principales (ACP) explica la varianza muestral multivariante con el menor número posible de componentes (principio de parsimonia), mediante la proyección ortogonal de las variables observadas en el espacio de componentes.
- El Análisis Factorial Exploratorio (AFE), además de estar sometido al principio de parsimonia, también lo está al principio de interpretabilidad, consistente en que las variables latentes deben ser nombradas e identificadas con un concepto teórico propio del fenómeno de estudio. A diferencia del ACP, el AF busca explicar la estructura de covarianzas entre variables y es posible comparar con un Análisis Factorial Confirmatorio (AFC) a través de un modelo de SEM: si ambas estructuras son similares esto respalda la consistencia de la estructura de variables latentes. Estudios recientes cuestionan que el carácter exploratorio o confirmatorio de estas técnicas sea la característica que las diferencia y proponen otra distinción respecto de las relaciones entre las variables y los propios factores: lo que denominamos Análisis Factorial Exploratorio deberíamos considerarlo un modelo no restringido, y el Confirmatorio un modelo restringido (Lloret-Segura et al., 2014).

En la presente investigación se ha optado por realizar un AFE debido a la necesidad en ciencias sociales de aplicar el principio de interpretabilidad.

A partir de lo anterior, en el desarrollo de la investigación se ha optado por una extracción por máxima verosimilitud (ML) por su carácter inferencial (Lawley & Maxwell, 1971), donde la estimación de los parámetros se realiza maximizando la probabilidad respecto de la matriz de correlaciones observada, bajo el supuesto de normalidad multivariante con m factores latentes. Además, la extracción ML tiene otras ventajas. Peña (2002) lo describe en los siguientes términos:

“... se comprueba que la estimación ML es invariante ante transformaciones lineales de las variables. En consecuencia, el resultado de la estimación no depende, como ocurre en componentes principales, del uso de la matriz de covarianzas o de correlaciones. Una ventaja adicional del método de máxima verosimilitud es que podemos obtener las varianzas asintóticas de los estimadores mediante la matriz de información en el óptimo” (p. 372).

Paso previo a la extracción (ML) hay que valorar la normalidad de las variables observadas a nivel poblacional a través de la distribución de las puntuaciones factoriales (F) (estimación del valor del factor para cada caso). Solo entonces se podrá realizar la extracción ML, de conformidad al cumplimiento del supuesto de normalidad multivariante, bien mediante la prueba de bondad del ajuste de Shapiro-Wilk (Shapiro & Wilk, 1965) o por el criterio de estadístico de asimetría y curtosis (Hogarty et al., 2005).

En el caso de las encuestas sociológicas la mayoría de las variables son de carácter politómico o dicotómico, lo que presenta varias dificultades a la hora de considerar que proceden de una población normal multivariante, tales como: i) que los coeficientes de asimetría y curtosis de las variables se encuentren entre $[-1,1]$ (Muthen & Kaplan, 1992; Ferrando & Anguiano-Carrasco, 2010), o ii) que el rango se encuentre entre $[-2,2]$ (Muthen & Kaplan, 1992; Bandalos & Finney, 2014). El impacto negativo de la asimetría puede paliarse con un gran tamaño muestral o el número de variables que participan en la extracción del factor (Muthen & Kaplan, 1992; West et al., 1995; Gao et al., 2008). El peor escenario se presenta cuando las comunales son bajas, inferiores a 0,30, por lo que entonces se precisa una muestra mínima de 400 casos (Conway & Huffcutt, 2003) o incluso de 500 (Hogarty et al., 2005).

Cuando los factores extraídos se distribuyen como una normal multivariante, la extracción ML consiste básicamente en minimizar la distancia F entre la matriz de covarianzas observada S y las puntuaciones asignadas a los casos por el modelo factorial. Esto equivale a maximizar la función de verosimilitud del modelo de m factores.

$$F = \ln|\Lambda\Lambda' + \Psi| + \text{traza}(S|\Lambda\Lambda' + \Psi|^{-1}) - \ln|S| - p$$

donde Λ es matriz de carga (covarianza entre los factores y las variables observadas), Ψ recoge las perturbaciones de todas las variables, la matriz de covarianzas queda expresada por $V = \Lambda\Lambda' + \Psi$, y p es el número de variables observadas.

Expresar la matriz de covarianzas como $V = \Lambda\Lambda' + \Psi$ permite que las varianzas de las variables observadas pueden descomponerse como:

$$\sigma_i^2 = \sum_{j=1}^m \lambda_{ij}^2 + \psi_i^2 = h_i^2 + \psi_i^2 \quad i = 2, \dots, p.$$

Se denomina comunalidad a $h_i^2 = \sum_{j=1}^m \lambda_{ij}^2$, que es la varianza de la variable observada explicada por el modelo factorial. ψ_i^2 es la perturbación o variabilidad específica de cada variable. Es deseable que la comunalidad sea elevada para que se verifique el supuesto base del análisis factorial, que la varianza explicada es debida a los factores y no a la causada por el ruido o componente aleatorio (Peña, 2002, p. 358).

Para facilitar la interpretación de los factores, tras realizar la extracción ML, se procede a realizar una rotación ortogonal Varimax (Kaiser, 1958) que trata de concentrar las saturaciones altas en pocos factores. La rotación Varimax busca que los factores tengan una alta correlación con un pequeño número de variables y bajas para con el resto y, en segundo lugar, maximiza las diferencias entre las cargas factoriales, manteniendo ejes ortogonales. El método Varimax es iterativo y la rotación puede expresarse (Stegmann et al., 2006) en relación a la matriz de correlaciones R y la matriz de cargas Λ como:

$$R_{VARIMAX} = \arg \max_R \left[\left(\frac{1}{P} \sum_{j=1}^K \sum_{J=1}^P (\Lambda R)_{ij}^4 - \sum_{j=1}^k \left(\frac{1}{p} \sum_{i=1}^p (\Lambda R)_{ij}^4 \right)^2 \right) \right]$$

En relación con la validación, se analiza:

- El valor del determinante de la matriz de correlaciones muestral (Lloret-Segura et al., 2014): para realizar una reducción de dimensiones es preciso que las variables observadas no sean ni ortogonales ni combinaciones lineales exactas de las demás, lo que conlleva consecuencias en el determinante de la matriz de correlaciones en dos sentidos: i) no puede ser nulo, debido a que implicaría que existe alguna combinación lineal perfecta entre variables, y ii) no puede ser igual a la unidad, debido a que ello únicamente ocurriría en el caso en que las variables fueran ortogonales.
- Las medidas de adecuación muestral (Kaiser, 1974; Cerny & Kaiser, 1977): el estadístico de Kaiser-Meyer-Olkin (KMO) mide si las correlaciones parciales entre el conjunto de variables son adecuadas para realizar el análisis factorial. En la mayoría de los casos se consideran adecuadas si el $KMO \geq 0'8$, por lo que las medidas de adecuación muestral para cada una de las variables (MSA) coinciden con los elementos de la diagonal de la inversa de la matriz anti-imagen, en cuyo caso: se consideran adecuadas si $MSA_i \geq 0'7$ (Lloret-Segura et al., 2014), siendo r_{ij}^2 los

coeficientes de determinación de Pearson, y $r_{ij \cdot m}^2$ los coeficientes de correlación parcial al cuadrado. Se expresa como sigue:

$$KMO = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} r_{ij \cdot m}^2} \quad MSA_i = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} r_{ij \cdot m}^2}$$

- Test de esfericidad de Bartlett (Jones, 1994; Rahmer et al., 2020): se contrasta la hipótesis nula en cuanto a que k muestras poseen igualdad de varianzas. Su aplicación al análisis factorial equivaldría a contrastar la hipótesis nula relativa a que la matriz de correlaciones poblacional es igual a la matriz identidad. Cuando el objetivo es acometer una reducción de dimensiones, es necesario que las variables poblacionales no sean ortogonales, por lo que se rechaza la hipótesis nula como sigue:

$$\begin{cases} H_0: V_i = V_{j \neq i} \\ H_1: V_i \neq V_{j \neq i} \end{cases} \quad \forall i = 1, \dots, k \quad \Leftrightarrow \quad \begin{cases} H_0: R = I \\ H_1: R \neq I \end{cases}$$

Así, el número de factores extraídos coincide con el número de variables observadas y cada factor en el autovalor muestra su capacidad explicativa respecto de la matriz de observaciones X . Por esta razón debe seleccionarse el menor número posible de factores que proporcionalmente tienen mayor capacidad para explicar la varianza total (principio de parsimonia). Se valoran simultáneamente tres criterios (Ibar, 2014):

- Criterio de Kaiser: selección de los factores cuyo autovalor sea igual o superior a la unidad:

$$\lambda_i \geq 1$$

- Criterio de la regla del codo en el gráfico de sedimentación: selección de aquellos factores donde la pendiente en el gráfico de sedimentación aumenta bruscamente de forma local.
- Criterio de la Varianza Total Explicada (VTE): selección de tantos factores como sean necesarios para explicar aproximadamente un 60% de la varianza total explicada por el modelo factorial.

En el ámbito de las ciencias sociales el principio de interpretabilidad tiene una especial relevancia; cuestión crucial cuando es necesario flexibilizar los criterios. Las pautas originalmente marcadas por Spearman (1927) en la determinación del factor inteligencia, son tan vigentes como entonces para medir una variable latente con las observaciones disponibles (Lorenzo-Seva et al., 2011), de manera que: i) la interpretación de los factores se realiza asociando conceptualmente las variables que tienen mayor carga factorial con el factor, al que se considera variable causante de las variables observadas; ii) la escala inversa o directa en el cuestionario debe ser tenida en cuenta en relación al incremento o decremento del factor con las variables que permiten su extracción; iii) si la extracción (de conformidad a los criterios meramente cuantitativos) no permite una interpretación correcta del factor, es aceptable la eliminación de aquellas variables cuya relación conceptual con el factor no está justificada; iv) una vez seleccionado el número de factores, las variables observadas que han participado en la extracción serán sustituidas por la estimación de los valores de los factores para cada caso; a estas estimaciones se las denomina puntuaciones factoriales y pueden obtenerse por distintos métodos, y v) cuando el tamaño muestral es enorme y las variables son ordinales, pueden estimarse las puntuaciones por regresión.

2.1. Propuesta metodológica

Sobre las bases teóricas anteriormente señaladas (medidas de bondad del ajuste, contrastes propios del análisis factorial confirmatorio y de las SEM), para evaluar en qué medida se mantiene la estructura factorial, se propone una interpretación novedosa: aplicar el principio de parsimonia a los estadísticos RMSEA (*Root Mean Error of Aproximation*), FIT (*Factor Fit of the Complete Model*) y BIC (*Bayesian*

Information Criterion), similar a la selección del número de factores a través de la conocida regla del codo en el gráfico de sedimentación, de amplia aceptación. El valor añadido de la presente propuesta es que además de los factores extraídos de la primera encuesta, variables observadas originalmente no incluidas pueden considerarse linealmente independientes en análisis ulteriores.

En lo relativo a la comparación entre periodos o años, se propone: i) argumentar por qué justamente estos son los periodos o años que se desean estudiar a partir de las semejanzas y divergencias entre ellos; ii) conocer la estructura dimensional de cada encuesta y a partir de ello analizar en qué medida puede reproducirse/compararse en el segundo periodo o año; iii) disponer de métodos robustos y aplicarlos al análisis individual de los periodos o años que permitan evaluar las diferencias significativas en términos de contraste de hipótesis de la estructura dimensional ortogonal respecto de características politómicas o dicotómicas no ordinales del periodo o año a comparar, y iv) en caso de no existir coincidencia entre las categorías, habrá que recodificar con la mínima interpretación posible la pregunta del periodo o año de referencia para que sea compatible con el periodo a comparar. Aunque algunos factores no sean extraídos simultáneamente en ambas encuestas, serán parcialmente comparables, si bien es preciso establecer un orden de preferencia, para lo que se propone:

- La mejor comparación posible se deriva de la estructura factorial del periodo o año de referencia al periodo o año que se desea comparar, para lo que es necesario que existan las mismas preguntas en escalas idénticas en ambos estudios, en cuyo caso se procederá a realizar un nuevo AFE con extracción ML sobre ambas encuestas simultáneamente y con las variables que hayan participado en la extracción del periodo o años de referencia.
- Comparación de las variables próximas (las de mayor carga factorial en el periodo de referencia respecto del factor) por ser las que mejor describen el comportamiento del factor a nivel muestral.
- Comparación de las variables que habiendo participado en el factor tengan la mayor carga factorial posible en el periodo o año de referencia.
- Cuando la variable no aparece en el factor del periodo de referencia, hay un paralelismo de contenido teórico entre la variable latente y la observada en ambos periodos o años. La comparación conceptual debe hacerse únicamente respecto de la variable seleccionada.
- Excepcionalmente, recurrir a variables que expresen una idea de complementariedad respecto de las variables del periodo o año de referencia.

En lo concerniente a la unidad de medida, aunque sólo se desee comparar el impacto de distintas variables en un periodo, los diferentes rangos en las escalas Likert y la heterogeneidad de variables hace necesario recurrir a la tipificación (Martín-Pliego, 2011) para comparar los efectos de las variables en términos de desviaciones típicas adimensionales. Esta tipificación debe ser corregida por la escala, inversa o directa, en que está realizada la pregunta del cuestionario para que la relación entre las variables sea coherente. Para conseguirlo se propone una tipificación corregida z^* que transforme todas las variables a una escala directa, como sigue:

$$z_i^* = (-1)^a * \left(\frac{x_i - E(x_i)}{\sqrt{V(x_i)}} \right); \quad \begin{cases} \text{relación inversa} \Rightarrow a = 1 \\ \text{relación directa} \Rightarrow a = 0 \end{cases}$$

Esta tipificación siempre es reversible y en cualquier caso puede facilitar la interpretación. En segundo lugar, cuando no se evalúan diferencias de medias la puntuación tipificada puede ser interpretada en términos de medidas de posición, lo que facilita comprender la posición relativa del dato de forma más intuitiva que cuando se recurre a la desviación estándar como unidad de medida. Y en tercer lugar, en el caso de los factores es innecesario hacer esta transformación porque, por definición, son variables de media cero y desviación típica unitaria, a excepción del signo cuando el factor está en

escala inversa (Los gráficos y transformaciones aparecen en el anexo “Diferencia de medias 2910 3259.xlsx”).

En lo relativo estrictamente al método de comparación, seleccionar un conjunto de variables objetivo de naturaleza politómica o dicotómica que midan de forma directa el objeto de estudio, si bien: i) cuando la variable politómica sea de naturaleza ordinal, pero no diferencia sus categorías de forma significativa en términos de contraste, procede dicotomizarla porque así se podrán comparar los subconjuntos de casos que se identifican con estas categorías; ii) cuando el objetivo es evaluar si las diferencias muestrales en términos medios son estadísticamente significativas, proceder a la prueba de igualdad de medias (siempre que ésta no tenga un sentido teórico relevante) *t* de Student (Welch, 1938, 1947) y asintótica por Monte Carlo (Fernández & Cao, 2020) y iii) como resultado, sólo resaltarán aquellas diferencias entre ambos periodos o años significativas en términos estadísticos y se eliminará la información irrelevante o redundante.

3. Resultados

En el desarrollo de la investigación empírica se utilizaron los siguientes software especializados: IBM SPSS 25 para Windows (IBM Corp, 2017), Microsoft Excel (Microsoft Corporation, 2020), R (R Core Team, 2013) y R-Studio (Rstudio Team, 2020). Las fuentes de información empleadas fueron:

- Para el periodo o año de referencia, el estudio CIS 3259 *Opinión Pública y Política Fiscal (XXXVI)* (CIS, 2019a), dirigido a la población española de ambos sexos de 18 años en adelante. El diseño del muestreo es estratificado por conglomerados con un diseño para 2.500 entrevistas, de las que se llevaron a cabo 2.464 (CIS, 2019b).
- Para el periodo o año a comparar, el estudio *CIS 2910 Opinión Pública y Política Fiscal (XXVIII)* (CIS, 2011a), de similares características respecto del diseño del muestreo: de las 2500 entrevistas previstas se realizan 2.468 (CIS, 2011b). Las variables identificadas que reflejan la percepción de los encuestados sobre el nivel de fraude fiscal en España son P15, P17, P18, P19 y P20 en el cuestionario del periodo de referencia (CIS, 2020). La base de datos recodificada para el periodo de referencia se encuentra como anexo: “3259 recod.sav”.

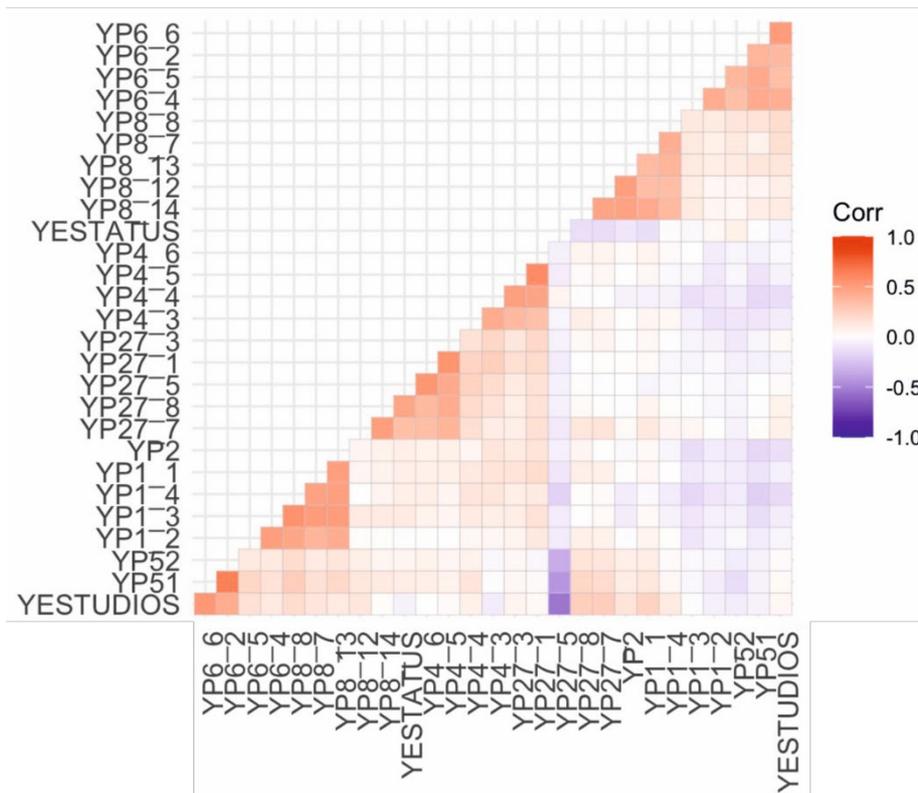
Tras seleccionar casos completos para realizar la extracción arrojó un tamaño muestral de 761 que verifica el criterio de Hogarty et al. (2005). Asimismo, la práctica totalidad de las variables mostró un coeficiente de asimetría y curtosis perteneciente al intervalo [-2,2].

3.1 Análisis factorial máximo verosímil para el Estudio CIS 3259.

3.1.1. Análisis exploratorio

En la Figura 1 se muestra que el determinante de la matriz de correlaciones, $|R| = 0.0001573812$, es próximo a cero (Lloret-Segura et al., 2014). La matriz policórica de las variables observadas muestra de forma visual como se relacionan linealmente. Al ser una matriz simétrica, sólo se muestra los elementos de la diagonal inferior (Kassambara, 2019).

Figura 1. Matriz de correlaciones policórica.



Fuente: Elaboración propia con ggcorrplot.

En este caso el estadístico B sigue una distribución χ_k^2 (lo que confirma a Bartlett, 1951 y Arsham & Lovric, 2011, entre otros), de forma asintótica con el tamaño muestral N ; siendo k el número variables (o grupos) independientes si la matriz de correlaciones población R es igual a la matriz identidad I . Para que exista una correlación entre variables suficiente a nivel poblacional para acometer el análisis factorial, interesa rechazar esta hipótesis como sigue:

$$H_0: R = I \Rightarrow B = \frac{(N - k) \ln \left(\frac{\sum_{i=1}^k (n_i - 1) s_i^2}{N - k} \right) - \sum_{i=1}^k (n_i - 1) \ln(s_i^2)}{1 + \frac{1}{3(k-1)} \left[\left(\sum_{i=1}^k \frac{1}{n_i - 1} \right) - \frac{1}{N - k} \right]} \xrightarrow{N \rightarrow \infty} \chi_k^2$$

En este caso, se obtiene un p-valor $sig = 0$, por lo que queda rechazada esta hipótesis nula a un nivel de significación $\alpha = 0'05$. Para un tamaño del efecto, $w = 0'3$, la potencia del contraste $(1 - \beta)$ calculada (en anexo “potencia bartlett CIS 2019.rtf”) con *G*Power* (Faul et al., 2007; Erdfelder et al., 2009), alcanza su valor máximo $(1 - \beta) = 1$ para un valor crítico del estadístico de Bartlett $B^* = 52,1923$ en una χ_{37}^2 .

El estadístico de adecuación muestral de Kaiser-Meyer- Olkin es adecuado ($KMO = 0'818$), así como las medidas de adecuación muestral para cada una de las variables que verifican:

$$MSA_i \geq 0'7 \quad \forall i = 1, \dots, p$$

Cálculos disponibles en el “Anexo 1. Análisis Factorial y pruebas de normalidad CIS3259.spv”. Los MSA en esta sección son en su mayoría superiores a 0,8. Cálculos realizados con IBM SPSS 25 para Windows (IBM Corp, 2017).

Todas las comunalidades h_i mejoran tras la extracción máximo-verosímil.

Para determinar el número de factores se ha valorado simultáneamente el criterio de Kaiser, el de la regla del codo en el gráfico de sedimentación y el de la varianza total explicada. En la Tabla 1 se observa cómo con seis factores se explica un 57,3% de la VTE y, tras la rotación Varimax, un 46%. Sin embargo, esto no afecta a la identidad de los factores que pueden ser interpretados con mayor facilidad en la matriz de factores rotados.

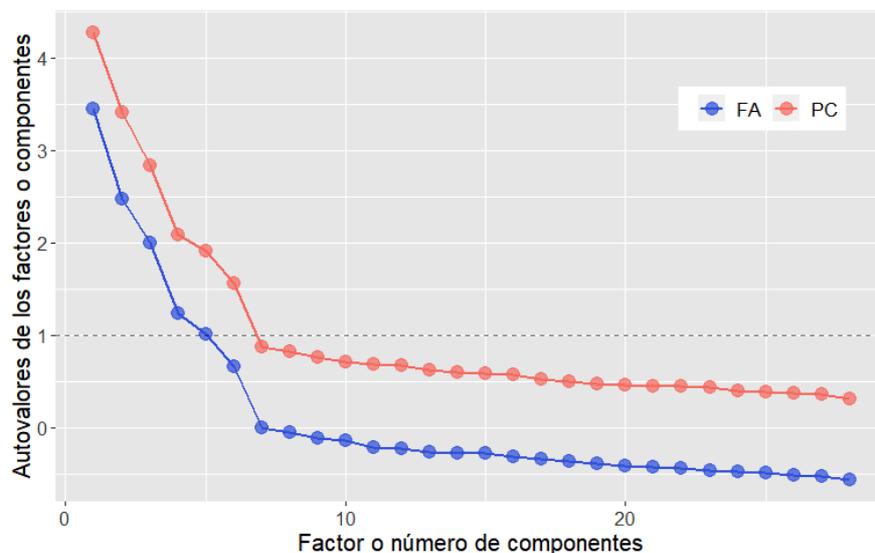
Tabla 1. Varianza total explicada (VTE) por el modelo factorial.

Factor (<i>i</i>)	Autovalores iniciales			Sumas de cargas al cuadrado de la extracción			Sumas de cargas al cuadrado de la rotación		
	λ_i	VTE	$\sum_i VTE$	λ_i	VTE	$\sum_i VTE$	λ_i	VTE	$\sum_i VTE$
ML1	4,27	15,3%	15,3%	3,74	13,4%	13,4%	2,48	8,9%	8,9%
ML2	3,42	12,2%	27,4%	0,00	10,0%	23,4%	2,28	8,1%	17,0%
ML3	2,84	10,1%	37,6%	2,33	8,3%	31,7%	2,15	0,0%	24,7%
ML4	2,08	7,4%	45,0%	1,55	5,5%	37,2%	2,11	7,5%	32,2%
ML5	1,90	6,8%	51,8%	1,40	5,0%	0,0%	2,03	7,3%	39,5%
ML6	1,55	5,5%	57,3%	1,07	3,8%	46,0%	1,84	6,6%	46,0%

Fuente: Elaboración propia.

En la Figura 2, para comprobar la consistencia del número de factores seleccionados, se ha comparado el número de componentes principales y de factores que se seleccionarían según la regla del codo. Aplicando el criterio de parsimonia, en sentido estricto, deberían seleccionarse 4 factores (primer codo en el gráfico de sedimentación); sin embargo, en el ámbito de las ciencias sociales el principio de interpretabilidad cobra una especial importancia. Con ánimo de encontrar el mayor número posible de dimensiones explicativas, se hubieran elegido 7 factores según regla del codo, aunque el autovalor del factor 7 no supera la unidad. Tanto el gráfico de sedimentación de los componentes principales como el de los factores recomiendan la misma selección de cuatro o siete (Figura 2), lo que da consistencia a decisión de la selección del número de latentes.

Figura 2. Gráfico de sedimentación de componentes principales y factores.

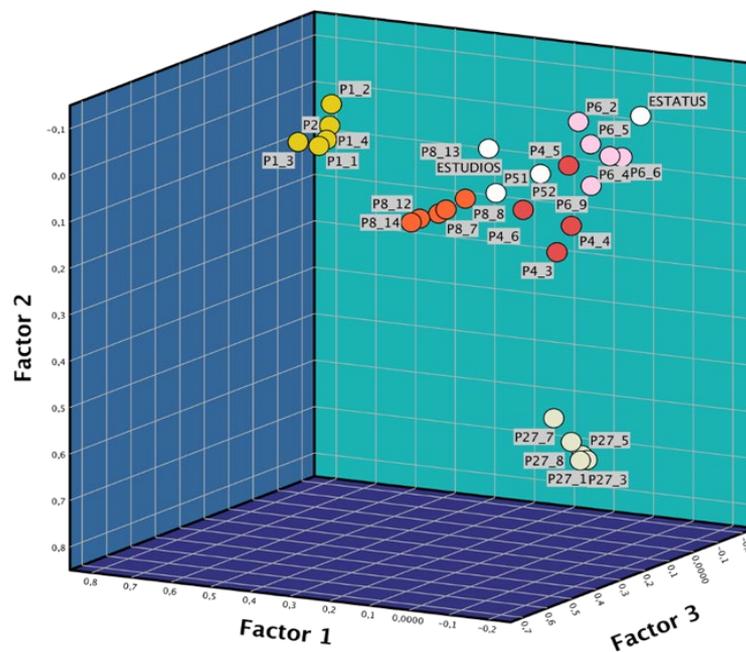


Fuente: Elaboración propia con *Psych* (Revelle, 2020).

Ahora bien, esta selección no cumple el principio de interpretabilidad porque en el factor 7 no satura ninguna variable en este factor. De acuerdo a Lorenzo-Seva et al. (2011), se han seleccionado 6 factores, que proporcionan una varianza total explicada próxima al 60% (criterio de la VTE) y la matriz de factores rotados tiene una interpretación clara: i) felicidad: medidas de forma, carga factorial y escala en cuestionario (ML1); ii) intolerancia con el fraude fiscal: medidas de forma, carga factorial y escala en cuestionario (ML2); iii) necesidad de financiación en políticas públicas de largo plazo: medidas de forma, carga factorial y escala en cuestionario (ML3); iv) insatisfacción con el funcionamiento de los SSPP del estado de bienestar: medidas de forma, carga factorial y escala en cuestionario (ML4); v) ascensor social: medidas de forma, carga factorial y escala en cuestionario (ML5), y vi) moralidad ciudadana: medidas de forma, carga factorial y escala en cuestionario (ML6).

En la Figura 3 se observa cómo las variables que son saturadas por los 6 factores quedan claramente diferenciadas en el espacio de los tres primeros factores rotados.

Figura 3. Espacio de factores rotados.

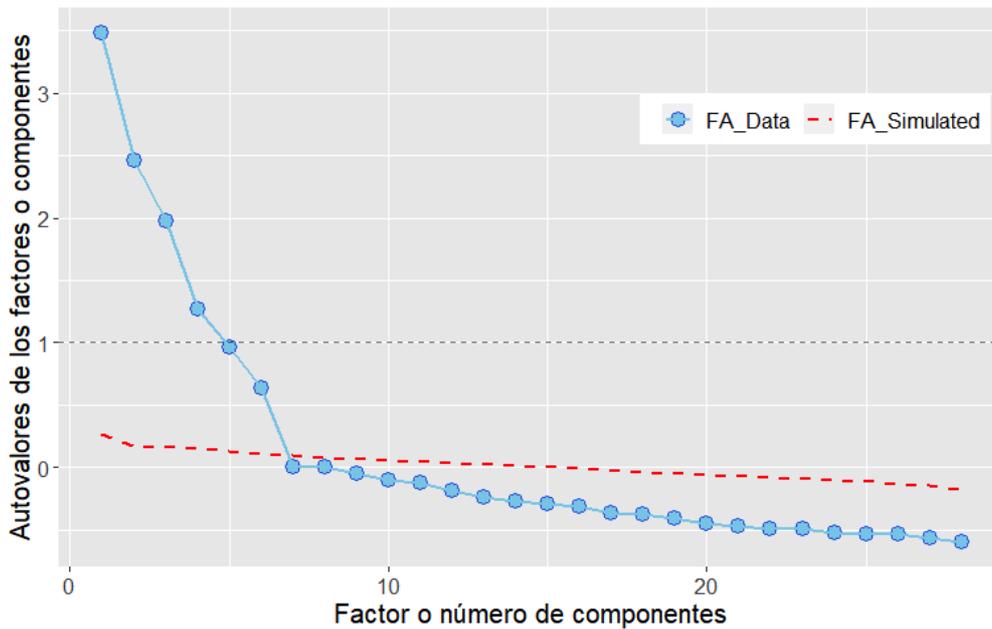


Fuente: Elaboración propia con *IBM SPSS25 para Windows*.

3.1.2. Análisis confirmatorio.

El análisis en paralelo (Velicer) selecciona los componentes o factores comunes que presentan autovalores superiores a los que se obtendrían aleatoriamente (Horn, 1965). En la Figura 4 se comparan los factores extraídos con la información muestral con los procedentes de la simulación con datos aleatorios: el número de factores seleccionados debe ser superior a la línea que indica la simulación (Revelle & Rocklin, 1979), por lo que la selección de 6 factores es adecuada.

Figura 4. Análisis paralelo de los gráficos de sedimentación.



Fuente: Elaboración propia con *Psych* y *ggplot2*.

Tabla 2. Cargas factoriales y proporción de varianza.

	ML2	ML1	ML3	ML6	ML5	ML4
Cargas factoriales	2,55	2,45	2,09	2,03	1,99	1,84
Proporción de varianza	0,09	0,09	0,07	0,07	0,07	0,07
Varianza acumulada	0,09	0,18	0,25	0,33	0,40	0,46
Proporción de varianza explicada	0,20	0,19	0,16	0,16	0,15	0,14
Proporción de varianza acumulada	0,20	0,39	0,55	0,70	0,86	1,00

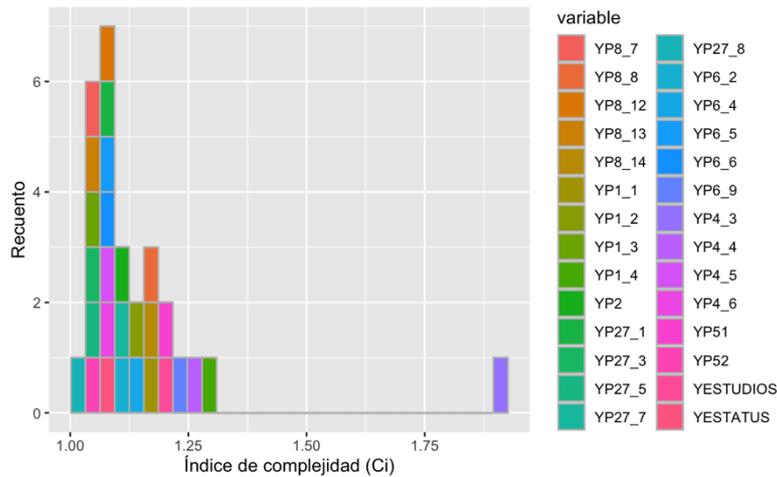
Fuente: Elaboración propia.

El índice de complejidad de Hoffman (Hofmann, 1978; Pettersson & Turkheimer, 2010) para cada ítem, siendo a_i la carga factorial para el factor i se define como:

$$c_i = \frac{(\sum_i a_i^2)^2}{\sum_i a_i^4}$$

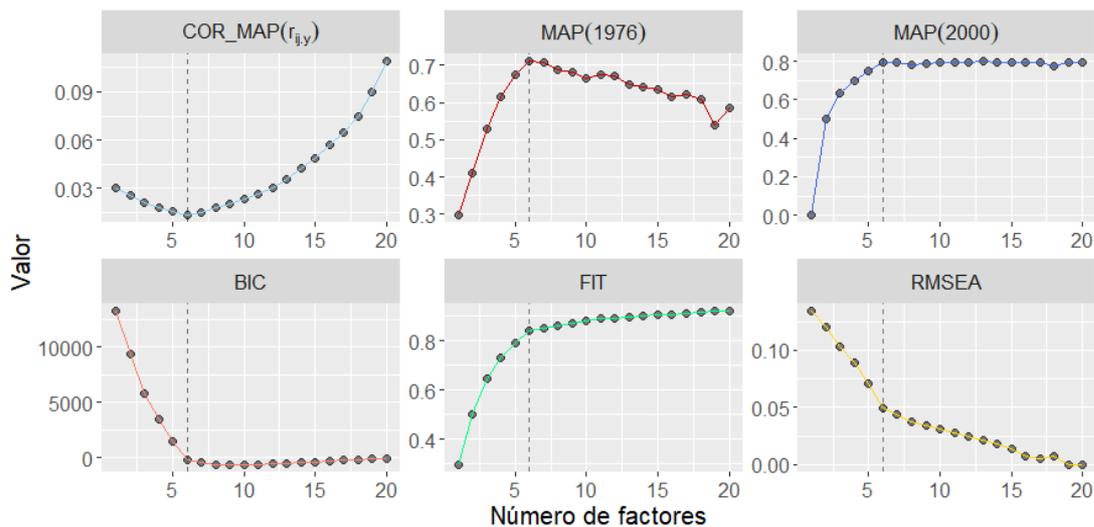
El índice de complejidad de Hofmann (1978) representa el número promedio de variables latentes necesarias para explicar las variables observadas. Mientras que una solución de estructura simple perfecta tiene una complejidad de uno, en el sentido de que cada elemento solo se cargaría en un factor, una solución con elementos distribuidos uniformemente tiene una complejidad mayor que 1. Nos interesa que la estructura no sea simple y perfecta porque entonces no tendría sentido la reducción de dimensiones. Cuanto mayor sea c_i , mejor será la calidad de la variable para participar en la extracción factorial. El índice de complejidad medio deberá ser superior a la unidad, en nuestro caso $\bar{c}_i = 1,1$ para un conjunto de seis factores.

Figura 5. Recuento de índices de complejidad de Hoffman.



Fuente: Elaboración propia con ggplot2 (Wickham, 2016).

Figura 6. Criterios para confirmación del número de factores para CIS 3259.



Fuente: Elaboración propia con ggplot2 y EFA.dimensions (O'Connor, 2020).

El gráfico $COR_MAP(r_{ij,y})$ muestra las correlaciones parciales del modelo factorial a medida que avanza la extracción:

$$r_{ij,y} = \frac{r_{ij} - r_{iy}r_{jy}}{[(1 - r_{iy}^2)(1 - r_{jy}^2)]^{1/2}} \in [-1, 1] \quad (\text{equation 11; Velicer, 1976})$$

En el numerador figura la covarianza entre cada par de variables (i, j) que irá decreciendo a medida que se extraigan factores porque cada vez será menor la varianza que queda sin explicar por el modelo. El denominador contiene los dos términos de correlación de cada variable y el componente y que se está eliminando. Tanto numerador como denominador se reducen con el número de factores extraídos, que van explicando progresivamente la varianza del modelo. Llega un momento en que el denominador comienza a reducirse más que el numerador, debido a que se está eliminando más varianza propia de cada variable que de la varianza conjunta explicada por el modelo factorial y, por tanto, aumenta la correlación parcial.

El gráfico MAP 1976 (Velicer, 1976), indica el número de factores óptimo para distintos niveles de complejidad. El gráfico MAP 2000, coincide con las aportaciones de Sharma (1996 - también véase Kaufman & Dunlap, 2000; O'Connor, 2000).

El criterio MAP de Velicer computa la media de las correlaciones parciales a medida que se avanza en la extracción, invitando a parar cuando se elimina más varianza propia de las variables que de la varianza común explicada por el modelo. En este caso, en 6 factores se alcanza el valor en que se explica proporcionalmente más varianza común que varianza propia de las variables.

Las medidas de bondad del ajuste del modelo factorial vienen expresadas por los gráficos FIT (*Factor Fit of the Complete Model*) o medida de ajuste del modelo respecto de la capacidad de explicar la varianza completa del modelo factorial (Clark & Bowles, 2018), el BIC o criterio bayesiano de comparación de modelos (Price, 2014), y el *Root Mean Error of Aproximation* o RMSEA (Smith et al., 1998; Steiger & Lind, 1980), expresado como sigue (Marais & Andrich, 2007):

$$RMSEA = \sqrt{\max \left[\frac{\chi^2}{df} - 1, 0 \right]} ; \begin{cases} \chi^2 = \text{chi - square RUMM2030} \\ df = \text{degrees of freedom} \\ N = \text{sample size} \end{cases}$$

Para un buen ajuste del modelo se precisa un valor del $RMSEA \leq 0,05$, y de al menos $RMSEA \leq 0,09$ (Tennant & Pallant, 2012). En el gráfico RMSEA de la Figura 6, se observa cómo al igual que en la regla del codo del gráfico de sedimentación puede evaluarse el principio de parsimonia por el cambio brusco en la pendiente de los tres indicadores propuestos de bondad del ajuste. Todos ellos recomiendan la extracción de seis factores, que en el caso de la presente investigación son los descritos en el apartado 3.1.1.

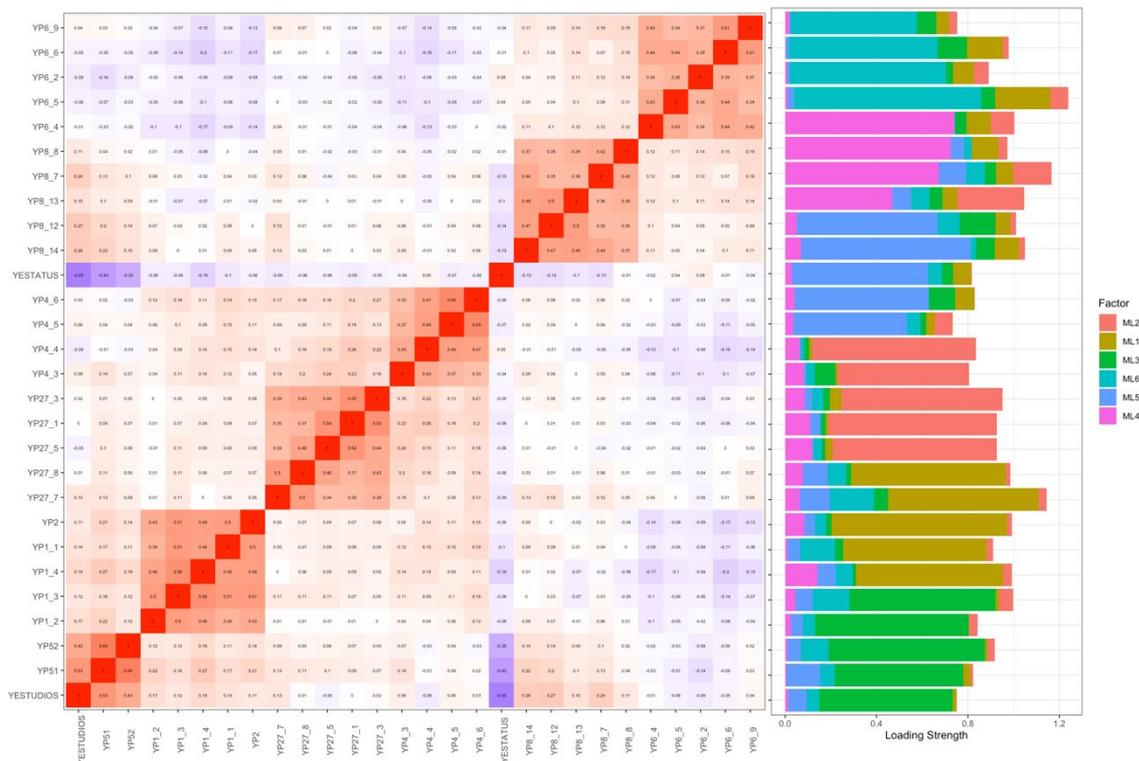
3.1.3. Interpretación de los factores

Simultáneamente se denominan los factores según el comportamiento de las variables que saturan en la extracción, se valora la normalidad multivariante de la distribución para posteriores contrastes y verificación de las hipótesis del modelo factorial.

En todos los contrastes para las variables, la prueba de Kolmogorov-Smirnov con simulación por Monte Carlo (se basa en 10.000 tablas de muestras con una semilla de inicio 2.000.000) rechaza la hipótesis nula de que las variables observadas sean normales univariantes con un nivel de significación $\alpha = 0'05$, y de hecho, sólo dos factores superan esta prueba: el factor 1 ($sig = 0'65$) y el factor 4 ($sig = 0'495$). No obstante, este contraste unidimensional es muy restrictivo desde el punto de vista multivariante (Fernández & Cao, 2020), razón por la que se recurre al criterio de asimetría y curtosis aprovechando la ventaja que otorga el tamaño muestral y cuyo resultado arroja que todos los estadísticos de forma de las variables que participan en el modelo pertenecen al intervalo [-1,2], a excepción de la variable p4_3 en el factor 6. Por tanto, de forma general no se rechaza que la distribución conjunta de los factores sea normal multivariante.

En las siguientes tablas se recogen los estadísticos de forma, junto a las variables ordenadas de mayor a menor carga factorial (en valor absoluto) en la matriz de factores rotados; por tanto, la primera variable será siempre la que mejor explica el factor (variable próxima). La escala se refiere al orden en que se formula la pregunta en la encuesta.

Figura 7. Matriz policórica y cargas factoriales CIS 3259.



Fuente: Elaboración propia.

Así, la variable más próxima al factor es la *PI_3* satisfacción con la vida social de entrevistado, seguido de *P1_4* su nivel de vida (relativo al bienestar económico), a su vez seguido de la escala de felicidad personal en una puntuación de cero a diez y, por último, seguido de la satisfacción con la vida familiar y la salud. A continuación, el factor relaciona satisfacción personal debida al establecimiento de vínculos con otras personas y la calidad del bienestar personal y, finalmente, el factor declara el nivel de felicidad del entrevistado, por tanto, se infiere que recoge la felicidad debida a estas dos cualidades y que podría denominarse Felicidad debida a la integración social y al bienestar personal. De forma reducida, Factor 1. Felicidad.

Tabla 3. Factor 1. Felicidad: medidas de forma, carga factorial y escala en cuestionario (ML1).

Etiqueta	Asimetría	Curtosis	Carga Factorial	Escala
P1_3. Su vida social	- 0,886	1,088	0,760	(+)
P1_4. Su nivel de vida	- 0,439	0,008	0,702	(+)
P2. Escala de felicidad personal (0-10)	-0,665	0,598	0,670	(+)
P1_1. Su vida familiar	-1,233	2,534	0,654	(+)
P1_2. Su salud	-0,987	1,271	0,630	(+)

Fuente: Elaboración propia.

La P27 refleja el nivel de intolerancia del entrevistado con distintas acciones de fraude tributario, que supuestamente son realizadas por terceros y se formula de forma impersonal. Las variables que saturan en el factor tienen en común los grandes impuestos: IRPF, IVA e impuesto de sociedades, pero no se recoge la evasión fiscal (P27_9), las infracciones puntuales como cobrar “en negro” (P27_2) o recibir una prestación a la que no se tiene derecho (P27_4). En resumen, la P27 se centra en la

intolerancia con aquellos que tienen obligación de declarar impuestos, en los términos descritos en la Ley 58/2003, de 17 de diciembre, General Tributaria (Jefatura del Estado, 2003). Éste se denomina: Factor 2: Intolerancia con el fraude fiscal de los contribuyentes.

Tabla 4. Factor 2. Intolerancia con el fraude fiscal de los contribuyentes: medidas de forma, carga factorial y escala en cuestionario (ML2).

Etiqueta	Asimetría	Curtosis	Carga Factorial	Escala
P27_1. No declarar todos los ingresos en el impuesto sobre la renta (IRPF)	-1,367	1,826	0,685	(-)
P27_5. Ser autónomo/a y no cobrar el IVA	-1,248	0,881	0,684	(-)
P27_3. Aplicarse una deducción que no le corresponde al realizar el pago de impuestos (declaración de IVA o IRPF)	-1,422	2,253	0,677	(-)
P27_8. Que una pequeña empresa eluda o evite pagar el impuesto de sociedades	-1,723	2,95	0,642	(-)
P27_7. Que una gran empresa eluda o evite pagar el impuesto de sociedades	-1,984	3,177	0,565	(-)

Fuente: Elaboración propia.

La P8 recoge el juicio del entrevistado sobre si la recaudación de los impuestos que paga se administra adecuadamente en relación con las necesidades de las distintas políticas. Este factor no incluye las políticas del Estado de bienestar, el mantenimiento de las fuerzas y cuerpos de seguridad del Estado ni la creación o el mantenimiento de infraestructuras. Sólo recoge políticas públicas de largo plazo, denominadas como: Factor 3. Necesidad de financiación en políticas públicas de largo plazo.

Tabla 5. Factor 3. Necesidad de financiación en políticas públicas de largo plazo: medidas de forma, carga factorial y escala en cuestionario (ML3).

Etiqueta	Asimetría	Curtosis	Carga Factorial	Escala
P8_14. Investigación en ciencia y tecnología	-1,626	1,756	0,678	(-)
P8_13. Cooperación al desarrollo	-0,862	-0,246	0,669	(-)
P8_12. Protección del medio ambiente	-1,230	0,502	0,657	(-)
P8_7. Cultura	-0,880	-0,256	0,579	(-)
P8_8. Vivienda	-0,996	-0,028	0,569	(-)

Fuente: Elaboración propia.

La P6 recoge el nivel de satisfacción en escala inversa con el funcionamiento de diversos servicios públicos. Los que recoge el factor están directamente relacionados con el Estado de bienestar, y denominados como: Factor 4. Insatisfacción con el funcionamiento de los servicios públicos del Estado de bienestar.

Tabla 6. Factor 4. Insatisfacción con el funcionamiento de los SSPP del Estado de bienestar: medidas de forma, carga factorial y escala en cuestionario (ML4).

Etiqueta	Asimetría	Curtosis	Carga Factorial	Escala
P6_6. Los servicios sociales	0,194	-0,413	0,708	(-)
P6_9. La ayuda a personas dependientes	-0,197	-0,413	0,649	(-)
P6_4. La Administración de Justicia	-0,309	-0,365	0,624	(-)
P6_5. La seguridad ciudadana	0,324	-0,35	0,603	(-)
P6_2. La asistencia sanitaria	0,123	-0,309	0,553	(-)

Fuente: Elaboración propia.

La variable que tiene mayor carga factorial es *Ingresos del hogar* (P51), seguida de *Ingresos personales* (P52) y después el nivel de titulación de estudios del entrevistado y su estatus estimado por el CIS; ambas denominadas: Factor 5. Ascensor social.

Tabla 7. Factor 5. Ascensor social: medidas de forma, carga factorial y escala en cuestionario (ML5).

Etiqueta	Asimetría	Curtosis	Carga Factorial	Escala
P51. Ingresos del hogar	-0,043	-1,182	0,794	(+)
P52. Ingresos personales	0,605	-0,494	0,698	(+)
ESTUDIOS. Estudios de la persona entrevistada [recodificada]	-0,317	-1,093	0,651	(+)
ESTATUS. Estatus socioeconómico de la persona entrevistada [recodificada]	0,122	-1,378	-0,568	(-)

Fuente: Elaboración propia.

La P4 recoge la exigencia del entrevistado para otorgar el calificativo de buen ciudadano debida a cierto comportamiento moral y cívico. En este factor se observa cómo la variable con mayor carga factorial refleja la tolerancia con *la opinión de los demás, aunque sean diferentes a las propias* (P4_5), seguido de variables referidas a las responsabilidades propias: responsabilidad y honestidad (P4_6), cumplir normativa (P4_4) y no evadir impuestos (P4_3). En P4_3 se establece una excepción respecto al criterio de curtosis para no rechazar la normalidad multivariante debido a la connotación que aporta a la interpretación del factor. Todo ello queda englobado bajo la denominación: Factor 6. Moralidad ciudadana.

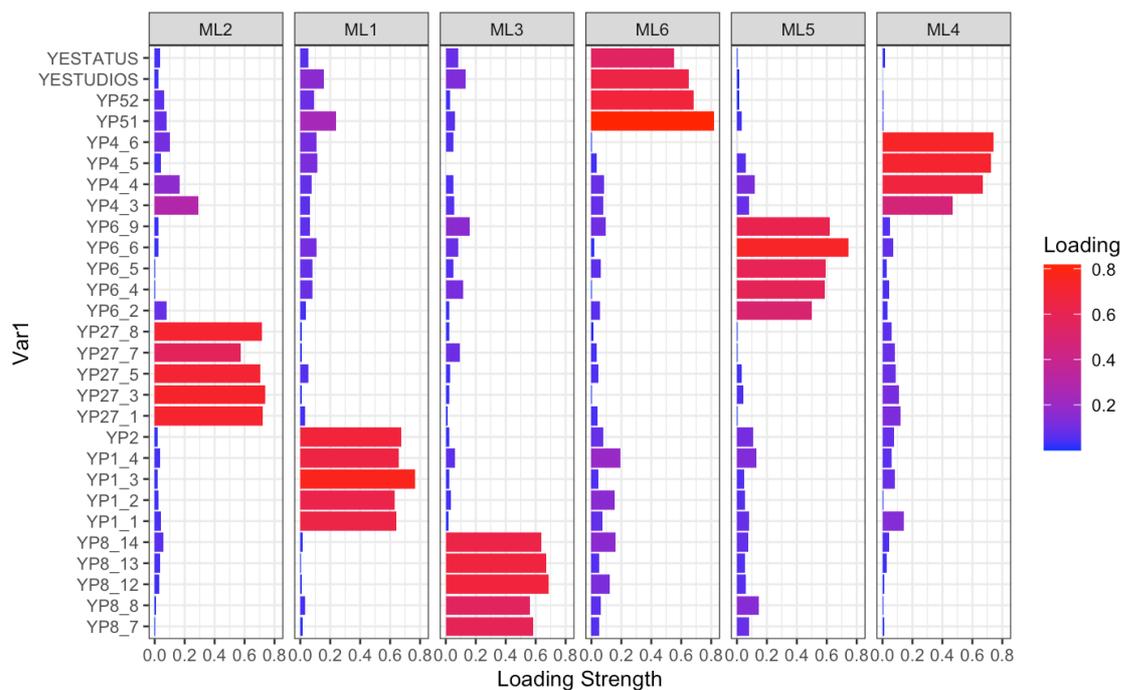
Tabla 8. Factor 6. Moralidad ciudadana: medidas de forma, carga factorial y escala en cuestionario (ML6).

Etiqueta	Asimetría	Curtosis	Carga Factorial	Escala
P4_5. Respetar las opiniones de los/as demás, aunque sean diferentes de las propias	-1,118	0,845	0,784	(+)
P4_6. Ser una persona responsable y honesta	-1,352	1,287	0,709	(+)
P4_4. Cumplir siempre las leyes y las normas	-1,208	1,147	0,626	(+)
P4_3. No evadir impuestos	-1,94	4,982	0,462	(+)

Fuente: Elaboración propia.

Una vez señalado lo anterior, en la Figura 8 se muestran los factores extraídos ordenados por las mayores cargas factoriales acumuladas respecto de las variables saturadas. Este gráfico permite comprender la estructura factorial respecto de las variables observadas de forma clara, valorando también las bajas cargas de las variables que no saturan y no se tienen en cuenta para dar nombre al factor. Esta estructura gráfica, que procede de una rotación Varimax, aporta indicios de relaciones estructurales relacionadas con los modelos de SEM (Bollen, 1989).

Figura 8. Representación policórica de cargas factoriales.



Fuente: Elaboración propia con ggplot 2 y Psych.

3.1.4. Contrastes de significación unidimensionales para el periodo de referencia

Las preguntas objetivo del Estudio CIS nº3259 que constan de 4 categorías se recodifican en dos categorías para proceder al simple contraste de hipótesis. Asimismo, las variables en el cuestionario aparecen en escala inversa, por lo que es menester recodificarlas de forma directa.

A continuación se procede al contraste de hipótesis de igualdad de medias siempre que se haya superado el contraste de normalidad univariante, o bien que los factores o variables procedan de una distribución normal multivariante siguiendo el criterio de asimetría y curtosis: rango $[-2, 2]$ (Muthen & Kaplan, 1992; Bandalos & Finney, 2014). Según este criterio no se rechaza cuando las puntuaciones factoriales F procedan de una población normal multivariante.

En las siguientes tablas únicamente se muestran los contrastes donde la diferencia de medias ha sido significativa, si bien cabe advertir: i) cuando se ha superado la prueba de Levene (Mandelbrot, 1961) de igualdad de varianzas, se ha procedido al contraste de la t de Student, y ii) cuando se ha rechazado la igualdad de varianzas, se ha procedido al contraste asintótico de Welch (1938; 1947).

De acuerdo a los resultados mostrados en la Tabla 9, los encuestados consideran que: i) los españoles son poco conscientes y responsables a la hora de pagar impuestos, lo que se traduce en una menor felicidad, en 0,21 desviaciones típicas (SD), y ii) hay menor interés en la financiación de políticas públicas de largo plazo, en 0,17 SD. Así, la menor felicidad probablemente no proviene del cumplimiento de las responsabilidades tributarias, sino del hecho de asumir comparativamente una mayor carga fiscal que las clases menos favorecidas, quienes en cambio más se benefician del Estado de bienestar (contrastes disponibles en anexo “Anexo 2 Factorial 3259 con contrastes de medias.spv”), y en el segundo caso, la falta de interés en la financiación de políticas de largo plazo probablemente obedece a la pertenencia a grupos sociales más favorecidos. Ambos resultados, sin embargo, hay que observarlos con cautela, a la espera de validación empírica ulterior que los confirmen.

Tabla 9. Contraste de la t de Student para la igualdad de medias de los factores sobre RP_16. ¿Cree Ud. que los/as españoles/as, a la hora de pagar impuestos, son muy conscientes y responsables, bastante conscientes y responsables, poco o nada conscientes y responsables? (0: poco o nada conscientes, 1: bastante o muy conscientes).

Variable	Sig.	Diferencia de Medias	IC (95%)	
			Inferior	Superior
F1. Felicidad	0,02	-0,21	-0,34	-0,08
F3. Necesidad de financiación de políticas públicas de largo plazo	0,004	0,19	0,06	0,32
F5. Ascensor Social	0,014	0,17	0,03	0,30

Fuente: Elaboración propia.

Los encuestados que se consideran a sí mismos muy conscientes y responsables en el pago de impuestos, son más intolerantes con el fraude fiscal (0,42 SD) y están más satisfechos con el funcionamiento de los servicios públicos del Estado de bienestar (0,35 SD). Este contraste: i) identifica el problema del fraude como algo externo al propio entrevistado, es decir, al resto de la ciudadanía; ii) las diferencias de medias son especialmente importantes, cercanas a 0,5 desviaciones típicas, y iii) la posición del propio entrevistado respecto de la obligación fiscal es especialmente diferenciadora de los prototipos de ciudadanía tributaria.

Tabla 10. Contraste de la t de Student para la igualdad de Medias de los factores sobre RP_17. ¿Y Ud., personalmente, se considera a sí mismo/a, muy consciente y responsable, bastante consciente y responsable, poco o nada consciente y responsable? (0: poco o nada conscientes, 1: bastante o muy conscientes).

Variable	Sig.	Diferencia de Medias	IC (95%)	
			Inferior	Superior
F2. Intolerancia con el fraude fiscal.	0	-0,42	-0,63	-0,21
F4. Insatisfacción con el funcionamiento de los servicios públicos del Estado de bienestar	0,001	0,35	0,14	0,56

Fuente: Elaboración propia.

Los encuestados que creen que existe mucho fraude fiscal en España consideran más necesaria la financiación de política públicas de largo plazo (0,48 SD), y están más satisfechos con los servicios públicos del Estado de bienestar (0,49 SD), con lo que se trata de ciudadanos moralmente muy comprometidos con el cumplimiento de las obligaciones tributarias.

Tabla 11. Contraste de la t de Student para la igualdad de medias de los factores sobre RP_18. En su opinión, ¿cree Ud. que en España existe mucho fraude fiscal, bastante, poco o muy poco fraude fiscal? (0 = poco o muy poco, 1 = mucho o bastante)

Variable	Sig.	Diferencia de Medias	IC (95%)	
			Inferior	Superior
F3. Necesidad de financiación de políticas públicas de largo plazo	0,005	-0,48	-0,82	-0,15
F4. Insatisfacción con el funcionamiento de los servicios públicos del Estado de bienestar	0,004	-0,49	-0,83	-0,15

Fuente: Elaboración propia.

Los encuestados que consideran que poca gente declara todos sus ingresos tributarios aprecian que hay una menor necesidad de financiación de las políticas públicas a largo plazo (0,24 SD), quienes a su vez están situados en una posición levemente inferior en la escala social (0,16 SD). Ambos contrastes, aunque significativos estadísticamente, muestran pequeñas diferencias entre ambos grupos. Esto viene a corroborar lo que se dedujo en el contraste en RP18, respecto de la necesidad de financiación en las políticas públicas de largo plazo. En segundo lugar, RP_19 sitúa al entrevistado en la valoración del volumen de fraude fiscal en su entorno social inmediato; por lo que, a medida que la posición en la escala social es inferior se percibe una mayor ocultación de rentas sobre los que tributar. Por último, en relación a la categoría recodificada "RP_20. Y entre la gente que Ud. conoce que está obligada a declarar el IVA (profesionales, autónomos/as, etc.), ¿cuánta cree Ud. que declara realmente todos sus ingresos al hacer la declaración del IVA: toda o casi toda, bastante, poca o ninguna?", no se detectan diferencias significativas en términos de contraste de igualdad de medias respecto de ninguno de los factores.

Tabla 12. Contraste de Welch para la igualdad de medias de los factores sobre RP_19. "Entre la gente que conoce, ¿cuánta cree Ud. que declara realmente todos sus ingresos al hacer la declaración de la renta (IRPF): toda o casi toda, bastante, poca o ninguna?" (0 = poca o ninguna, 1 = casi toda o bastante)

Variable	Sig.	Diferencia de Medias	IC (95%)	
			Inferior	Superior
F3. Necesidad de financiación de políticas públicas de largo plazo	0,004	-0,24	-0,41	-0,08
F5. Ascensor social	0,024	-0,16	-0,31	-0,02

Fuente: Elaboración propia.

Por último, los participantes encuestados que piensan que los impuestos se cobran con justicia están más satisfechos con el funcionamiento de los servicios públicos del Estado de bienestar (0,44 SD), de lo que se infiere que el cobro de impuestos se considera más justo en la medida en que se reciben contraprestaciones del Estado de carácter social.

Tabla 13. Contraste de la t de Student para la igualdad de medias de los factores sobre RP_15. Y, ¿cree Ud. que, en general, los impuestos se cobran con justicia esto es, que pagan más quienes más tienen, o no lo cree así?

Variable	Sig.	Diferencia de Medias	IC (95%)	
			Inferior	Superior
F4. Insatisfacción con el funcionamiento de los servicios públicos del Estado de bienestar	0	-0,44	-0,65	-0,23

Fuente: Elaboración propia.

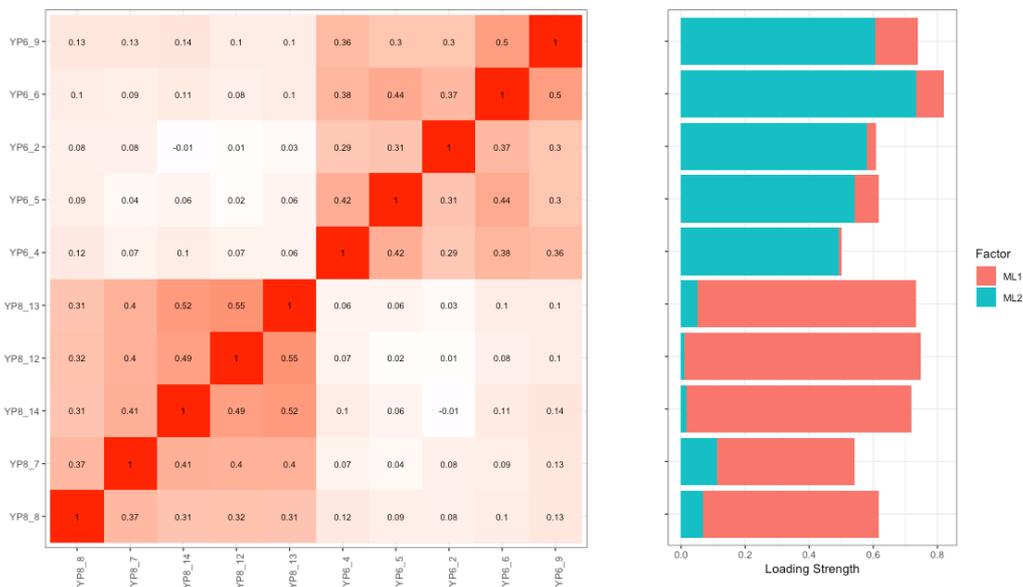
3.2. Análisis factorial máximo verosímil con extracción Varimax simultáneo para los estudios CIS N°3259 y CIS N°2910.

A continuación, se ha procedido a recodificar en términos de la encuesta de 2019 las variables de 2011 que son comparables, con lo que arroja una base de datos de tamaño efectivo de 4.932 líneas (Sintaxis de recodificación en anexo “Sintaxis 2 RECOD CIS3259 CIS2910.sps”. Nótese que debemos transformar la variable ESTUDIOS de CIS3259 para que coincida con las categorías de CIS2910. En el resto de las transformaciones se adecúan las variables de CIS2910 a CIS3259). El punto de partida es la extracción por máxima verosimilitud de forma análoga al apartado anterior, pero únicamente con las variables que han participado en la extracción de los factores, lo que se resume en F3 y F4 (Extracción ML en anexo “Anexo 3 AF ML CIS3259 CIS2910.spv”). A las variables se les añade el prefijo “Y” (tras su recodificación para R sustituyendo los valores perdidos por NA).

Tras seleccionar casos completos para realizar la extracción de ambos factores queda un tamaño muestral por casos completos aún mayor, de 2.447 que verifica el criterio de Hogarty. La práctica totalidad de las variables tiene un coeficiente de asimetría y curtosis perteneciente al intervalo [-2,2].

En la Figura 9, se observan las correlaciones entre pares de variables y las cargas en el modelo factorial.

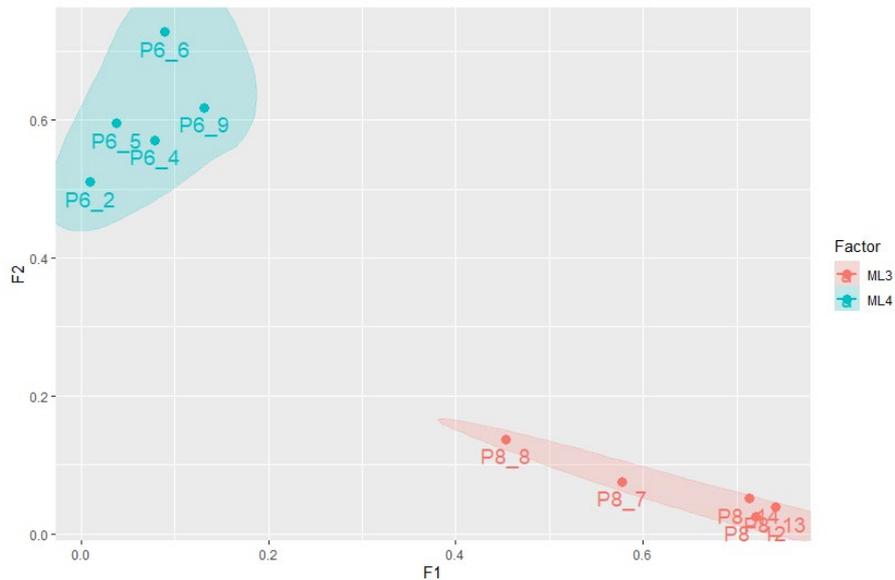
Figura 9. Matriz policórica de correlaciones y cargas factoriales acumuladas para el modelo conjunto AF ML CIS3259 y CIS2910.



Fuente: Elaboración propia con ggplot2 y Psych.

En la Figura 10 se observa cómo la proyección conjunta de las variables de ambos periodos en el espacio de factores rotados por Varimax, reconstruida respecto de los factores del periodo de referencia (ML3 y ML4), se mantiene bien diferenciada.

Figura 10. Espacio de factores rotados.



Fuente: Elaboración propia con ggplot2.

En la Tabla 14 se muestran los principales estadísticos de validación del modelo (cálculos en “Anexo 6 af stats CIS 2910 y 3259.spv” y “FA.pdf” y “Pconjunto19.pdf”). Teniendo en cuenta el pequeño número de variables que participan en la extracción del modelo factorial reproducido (AF ML 2), el determinante de la matriz de correlaciones es próximo a cero (Lloret-Segura et al., 2014). Así, todos los modelos quedan validados, sin bien es apreciable un cierto descenso en la bondad de los resultados del modelo referido al estudio CIS2910 (año 2011), respecto del modelo para CIS3259 (año 2019) que está directamente relacionado con el primer modelo (AF ML 6) que únicamente tiene en cuenta la estructura para el 2019 (Cálculos disponibles en el anexo “Anexo 1. Análisis Factorial y pruebas de normalidad CIS3259.spv”. Los MSA en esta sección son en su mayoría superiores a 0’8).

Tabla 14. Estadísticos y contrastes de validación del modelo factorial conjunto.

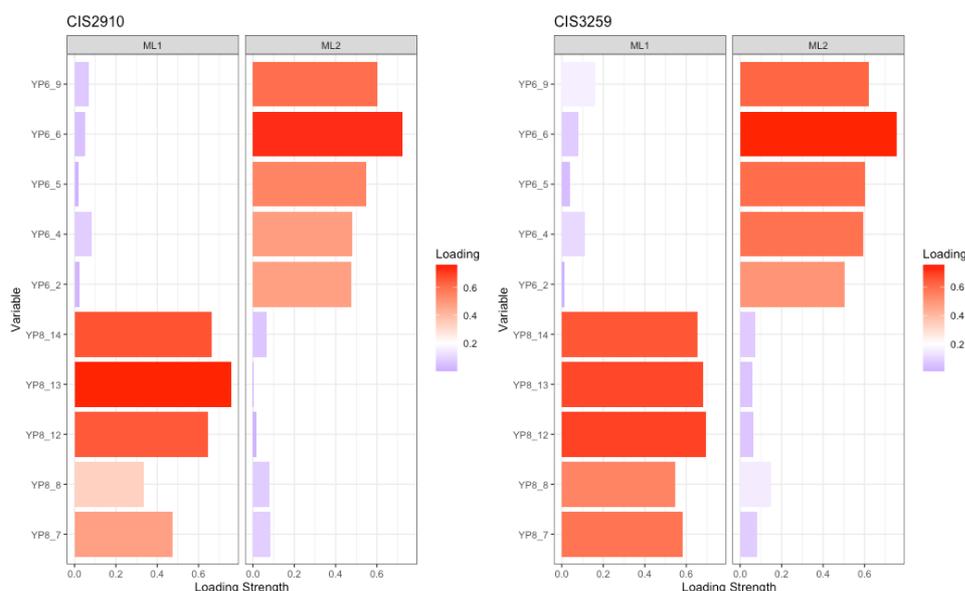
	AF ML 6 CIS3259	AF ML 2 CIS3259 CIS2010	CIS3259	CIS2910
N = casos completos	719	2474	1297	1177
R	0,000	0,099	0,07	0,153
KMO	0,818	0,802	0,829	0,757
Sig. Bartlett	0	0	0	0
Mean Item complexity	1,1	1	1	1
Número de factores suficientes (Velicer test)	6	2	2	2
RMSEA	0,046	0,053	0,052	0,058
IC_{α=0'1}(RMSEA)	(0'043, 0'048)	(0'048, 0'058)	(0'046, 0'059)	(0'052, 0'065)
RSMR	0,02	0,03	0,03	0,03
RSMR Corregido	0,03	0,04	0,03	0,04
BIC	-383,16	166,36	-1,23	39,51

FIT (based off upon diagonal values)	0,99	0,99	0,99	0,98
NNFI (Tucker & Lewis, 1973)	0,909	0,940	0,947	0,915

Fuente: Elaboración propia con SPSS 25 y *Psych*.

En la Figura 11 se compara la estructura de cargas factoriales para los modelos de cada uno de los periodos y, según se observa, se mantiene, aunque hay un cambio de variables próximas: en el periodo de referencia (CIS3259) la variable con mayor saturación es P8_12, mientras que en CIS2910 es P8_13.

Figura 11. Comparativa de cargas factoriales AF ML Varimax CIS3259 y CIS2910.



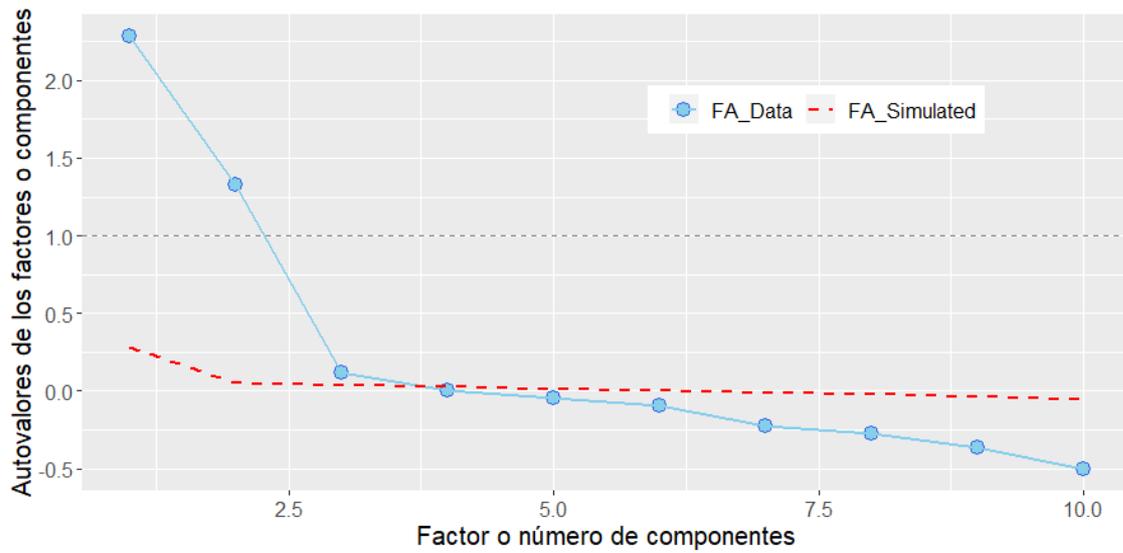
Fuente: Elaboración propia con ggplot2 y *Psych*.

La Figura 12 muestra el gráfico de sedimentación con el criterio de selección de Velicer, con una única extracción para ambos periodos. Se observa que la regla del codo indica la selección de 3 factores, así como el criterio del análisis paralelo. No obstante, no se verifica con tres factores, ni el criterio de Kaiser y tampoco es interpretable el tercer factor.

La Figura 13 muestra los gráficos de sedimentación y el análisis en paralelo para cada uno de los periodos con la misma estructura factorial. Nótese que el primer y el segundo autovalor son mayores respecto del periodo de referencia. Sin embargo, la estructura respecto del periodo 2011 es consistente, y de nuevo se seleccionarían en ambos casos dos factores por el mismo razonamiento antes explicado respecto del modelo factorial conjunto para ambos periodos.

En relación a la selección del número de factores, la Figura 14 se interpreta de idéntica manera a como se hizo para el modelo factorial para el periodo de referencia en la Figura 4. También se observa cómo el principio de parsimonia queda reflejado en el codo de los gráficos de bondad del ajuste.

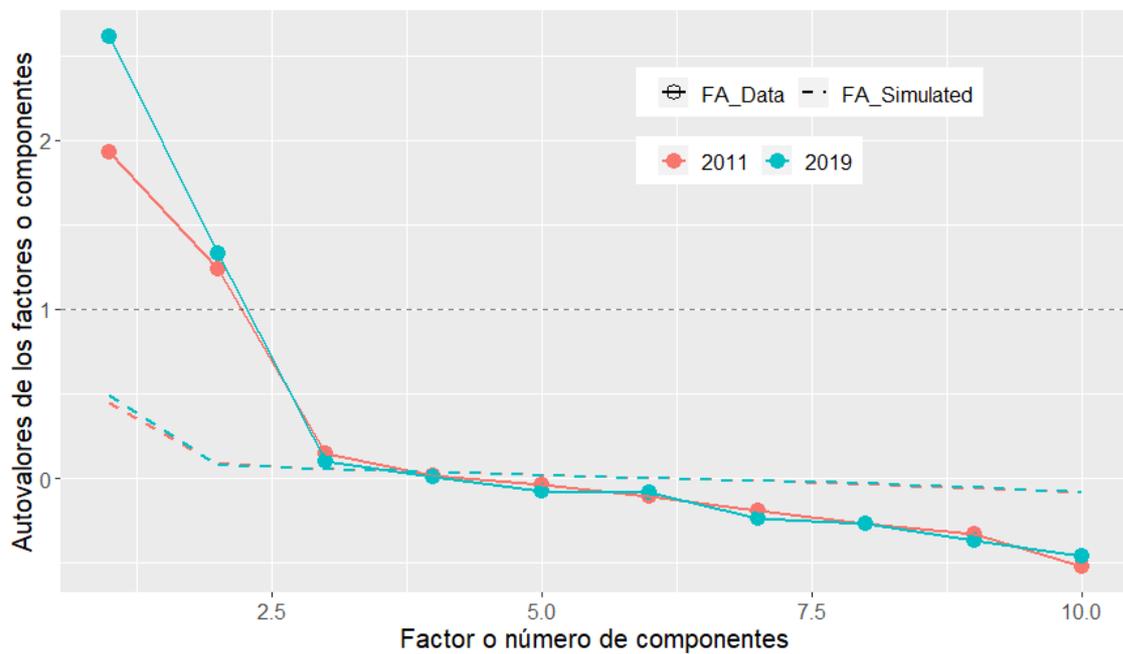
Figura 12. Gráfico MAP FA ML Varimax conjunto CIS3259 y CIS2910.



Fuente: Elaboración propia con *Psych* y *ggplot2*.

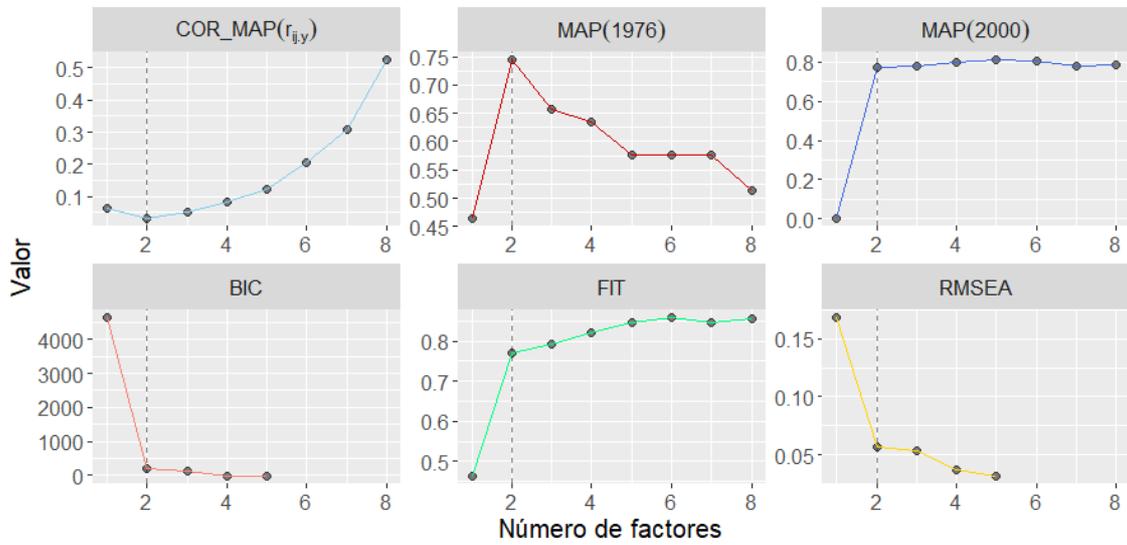
Finalmente, la estructura de selección se mantiene al comparar los criterios en ambos periodos, como muestra la Figura 13.

Figura 13. Gráfico MAP FA ML Varimax conjunto CIS3259 y CIS2910.



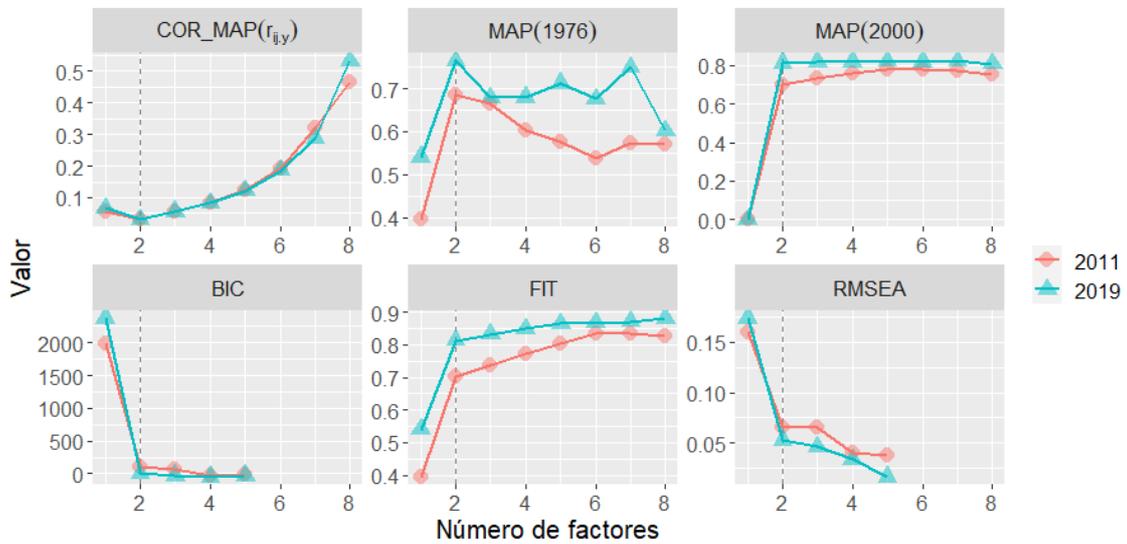
Fuente: Elaboración propia con *Psych* y *ggplot2*.

Figura 14. Criterios para confirmación del número de factores para el modelo factorial AF ML 2 conjunto CIS3259 CIS2910.



Fuente: Elaboración propia con *ggplot2* y *EFA.dimensions*.

Figura 15. Criterios para confirmación del número de factores para los modelos factoriales AF ML 2 CIS3259 CIS2910.



Fuente: Elaboración propia con *ggplot2* y *EFA.dimensions*.

3.3. Comparativa de las diferencias significativas entre ambos periodos.

A continuación, se procede a hacer un contraste de igualdad de medias. En esta sección sólo se muestran aquellas variables cuya diferencia es significativa en términos de contraste de hipótesis (Contrastes de hipótesis y estadísticos en anexo “Anexo 4 Contrastes y estadísticos CIS3259 CIS2910.spv”), con especial relevancia las que representan a los factores de 2019.

Tabla 15. Tabla de correspondencia para comparación de variables en 2011 y 2019.

Factor 2019	Relación con el modelo AF ML 6 CIS 3259 (Año 2019)	
ML1	Sin información	
ML2	Proxy	P27_1. No declarar todos los ingresos en el impuesto sobre la renta (IRPF)
ML3	Reconstruido	F3. Necesidad de Financiación en políticas públicas de largo plazo
ML4	Reconstruido	F4. Insatisfacción con el funcionamiento de los SSPP del Estado de bienestar
	Conceptual	P44. Valoración de la situación económica personal actual
ML5	3ª en carga (recodificada)	ESTUDIOSMIX.- ESTUDIOS 3359 en escala 2910
ML6	Sin información	

Fuente: Elaboración propia.

La mejor comparación posible se obtiene recuperando la estructura factorial de 2019: esto ocurre con ML3 y ML4. La segunda mejor alternativa es comparar las variables próximas: esto ocurre con ML2. La tercera mejor opción es comparar aquellas variables que habiendo participado en el factor tengan la mayor carga factorial posible respecto al modelo de 2019: esto ocurre con F5, representada por ESTUDIOSMIX. La cuarta opción es la comparación conceptual; si bien la variable no aparece en el factor original hay un paralelismo de contenido teórico entre la variable latente y la observada en ambos periodos: en este caso, la P44 (CIS2910) es conceptualmente aproximada a las variables P51(CIS3259) ingresos del hogar y P52(CIS3259) ingresos personales.

De acuerdo a metodología formulada en la sección 2.1, en la Tabla 16 se muestran las variables de auto identificación de los encuestados relevantes para el análisis del fraude fiscal y otros aspectos importantes.

Tabla 16. Tabla de variables objetivo en 2011 y 2019 y otras.

	Variable
Variables Objetivo	P10. Contraprestaciones que recibe la sociedad por el pago a las administraciones públicas de los impuestos y cotizaciones
	P18. Grado de fraude fiscal que existe en España
	P24. Valoración del esfuerzo de la Administración para la lucha contra el fraude fiscal
Curiosidad	P35. Edad de la persona entrevistada
Complementarios de F4	P8_1. Enseñanza, P8_2. Obras públicas, P8_3. Protección al desempleo
	P8_4. Defensa, P8_5. Seguridad ciudadana, P8_6. Sanidad, P8_9. Justicia
	P8_10. Seguridad Social/Pensiones, P8_11. Transporte y comunicaciones
	P8_15. Ayuda a personas dependientes

Fuente: Elaboración propia.

Las diferencias más relevantes para los años 2011 y 2019 quedan reflejadas en la Tabla 17, donde se pone de manifiesto la heterogeneidad de la unidad de medida desde variables politómicas de preferencia a variables como la edad con unidad de medida en años, e incluso en escalas directas e inversas. Asimismo, se observa que: i) ha aumentado la edad media de la población entrevistada en lo

relativo a la percepción de necesidad de llevar a cabo políticas públicas de largo plazo y ii) hay correlación entre el mayor nivel de estudios y la mejor situación económica personal, con el descenso en la percepción de necesidad de financiación de obras públicas.

Tabla 17. Prueba T de Student o asintótica de Welch de diferencias de medias (DM) significativas respecto a los periodos 2011 y 2019.

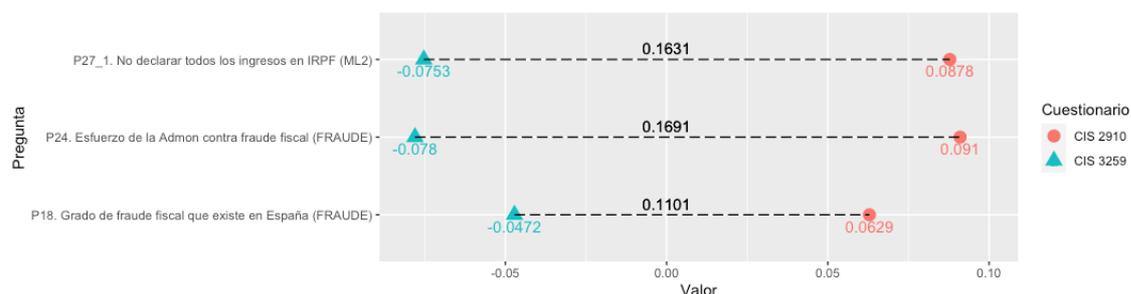
	P8_2. Obras públicas	ESTUDIOSMIX.- ESTUDIOS CIS3359 en escala CIS2910	F3. Necesidad de Financiación en Políticas Públicas de Largo Plazo.	P35. Edad de la persona entrevistada	P44. Situación económica personal (INV)
Unidad de medida	Likert: (3 categorías)	Likert (5 categorías)	SD	Años	Likert (5 categorías)
Media CIS 2910	1,98	3,0772	-0,2717347	47,16	2,94
Media CIS 3259	2,25	3,5713	0,2465935	50,62	2,16
DM	0,27	0,4941	0,5183282	3,46	-0,78

Fuente: Elaboración propia.

En la Figura 16 se muestran las posiciones de la media de cada variable en desviaciones típicas respecto a la media de las variables tipificadas con corrección de signo de escala (Z^*) en ambos años. Esta medida permite evaluar cuáles han sido los cambios más relevantes en el comportamiento medio de las variables.

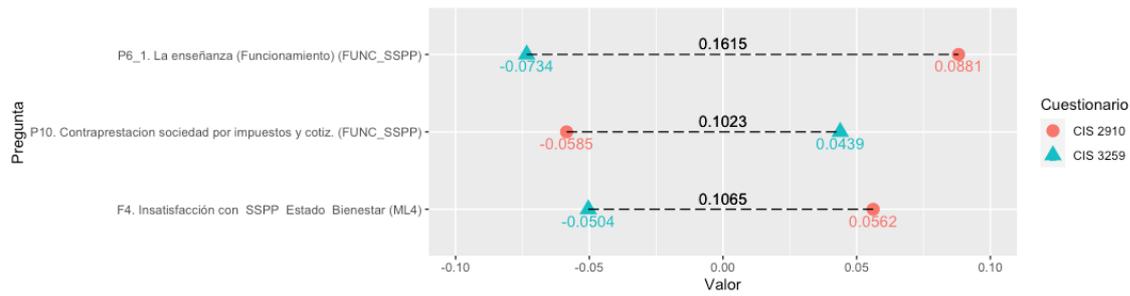
Finalmente, aunque la diferencia de medias no es elevada en las variables ESTATUS y ESTUDIOSMIX, se observa un cambio en la forma de la distribución, como se muestra en las figuras 16 a 19.

Figura 16. Diferencias estadísticamente significativas Fraude Fiscal CIS 2910 y CIS3259 en puntuaciones Z^* .



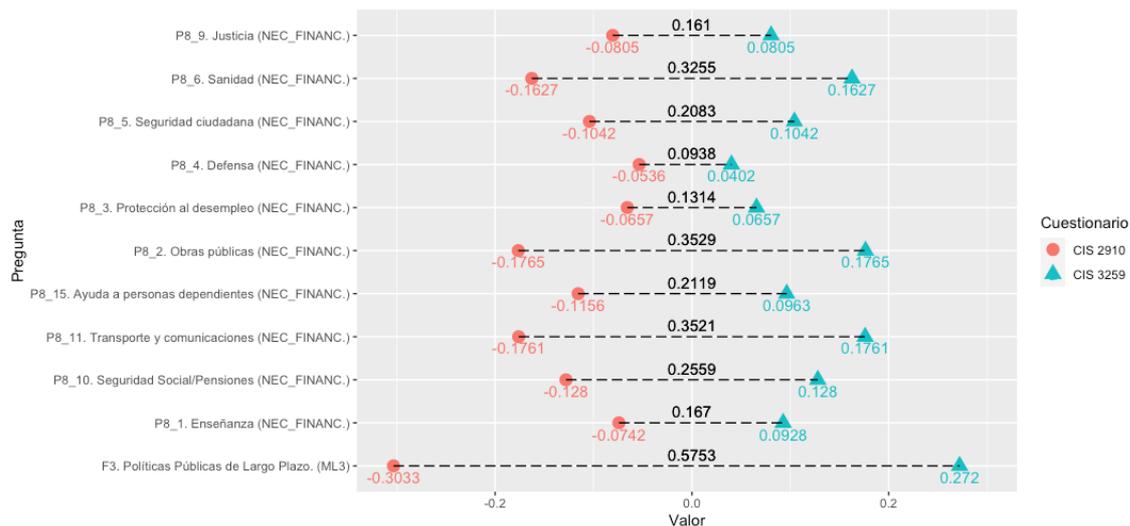
Fuente: Elaboración propia.

Figura 17. Diferencias estadísticamente significativas funcionamiento de los Servicios Públicos CIS 2910 y CIS3259 en puntuaciones Z*.



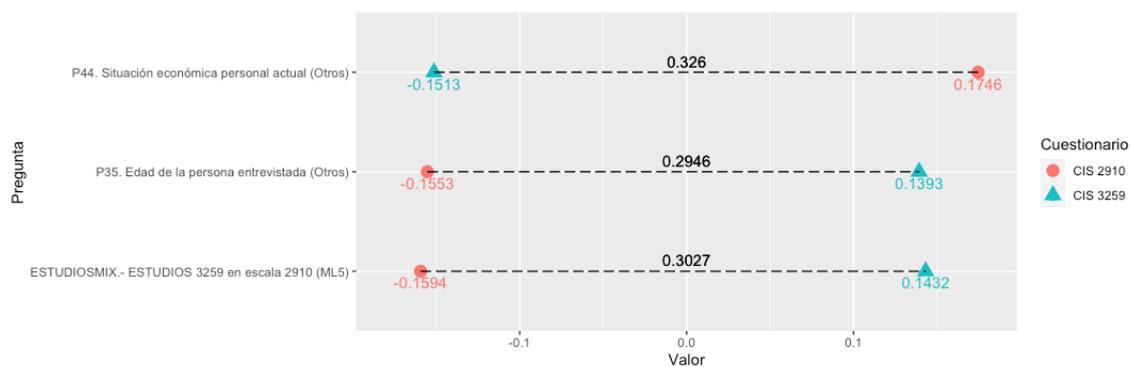
Fuente: Elaboración propia.

Figura 18. Diferencias estadísticamente significativas Necesidad de Financiación de los Servicios Públicos CIS 2910 y CIS3259 en puntuaciones Z*.



Fuente: Elaboración propia.

Figura 19. Diferencias estadísticamente significativas perfil personal CIS 2910 y CIS3259 en puntuaciones Z*.



Fuente: Elaboración propia.

4. Discusión

A lo largo de la presente investigación se han constatado las dificultades metodológicas en la comparación de encuestas, tales como: i) el añadido de nuevas preguntas o la eliminación de otras; ii) la selección de las variables a comparar; iii) la recodificación entre cuestionarios no coincidentes y iv) la discrecionalidad en las unidades de medida de las variables. Con estas dificultades presentes, el trabajo empírico realizado es un proceso de transición desde el análisis exploratorio al análisis confirmatorio (entendido como un análisis factorial restringido).

La propuesta metodológica validada en los resultados es un procedimiento que, en efecto, permite recuperar la estructura factorial, realizar comparaciones entre variables y recodificar las unidades de medida y en la escala, en dos periodos o años. Para verificar tal relación, se ha aumentado el nivel de exigencia habitual para la selección de variables que participan en el modelo exploratorio, tanto en el valor de la carga factorial (para considerar que las variables son saturadas por el factor la carga debía ser de al menos 0,7 en términos generales), como en la mejora en la extracción maximoverosimil en todas las comunalidades, lo que en conjunto ha permitido comparar el modelo exploratorio con contrastes y criterios del número de factores adecuado, propio del análisis factorial confirmatorio y las SEM.

En ciencias sociales el principio de interpretabilidad tiene una importancia medular, si bien está sujeto a multitud de dificultades. No obstante, en virtud de los resultados antes mostrados, se confirma la validez del criterio experto del análisis fundamentado en el principio de interpretabilidad al margen del criterio exploratorio, a través de los criterios de selección cuantitativos (Kaiser, codo en el gráfico de sedimentación y VTE), las pruebas (COR_MAP, MAP76 y MAP2000) y los criterios (FIT, VSS, RMSEA).

En lo relativo al principio de parsimonia en la regla del codo del gráfico de sedimentación (Figuras 2, 4, 12, 13), se ha demostrado ser aplicable a los estadísticos confirmatorios (BIC, FIT y RMSEA), si bien es necesario tomar en consideración dos criterios ampliamente aceptados en el modelado de las SEM: i) rigurosidad con la separación en las cargas factoriales, donde se propone un criterio de una diferencia de 0,3 entre la carga máxima y la anterior (particularmente véase la separación mostrada en las figuras de matrices policóricas con cargas factoriales asociadas (Figuras 7 y 9), y en los gráficos de cargas (figuras 8 y 11), y ii) optimización en las comunalidades en la extracción por máxima verosimilitud.

Los resultados antes mostrados constatan que la regla del codo aplicada a los estadísticos de confirmación del número de factores (BIC, FIT y RMESA) (véase las Figuras 6, 14 y 15), el número de factores que recomienda la aplicación del principio de parsimonia coincide con los criterios de selección de correlaciones de Velicer [$\min(r_{ij \cdot y})$], MAP (1976), MAP (2000). El gráfico del índice de complejidad de Hoffman (Figura 5) es especialmente interesante para detectar las variables menos adecuadas para incluirlas en el modelo factorial desde punto de vista cuantitativo, al margen del principio de interpretabilidad.

La metodología propuesta permite comprobar la consistencia de la estructura factorial para la reconstrucción de un periodo anterior (CIS2910) respecto a uno de referencia (CIS3259) y se manifiesta en los siguientes aspectos: i) mantenimiento o cambio de la variable próxima al factor; ii) posibilidad de medir la diferencia en las saturaciones de las variables que participan en la extracción para ponderar su nivel de importancia en cada periodo y iii) constatar si hay un debilitamiento o una mejora en los estadísticos de calidad del modelo factorial (BIC, FIT y RMSE<A) en el periodo en el que reconstruimos la estructura factorial.

En virtud de los resultados, se confirma la validez de la metodología propuesta en la sección 2.1 para la selección de las diferencias más significativas en términos de contraste y, en segundo lugar, que las transformaciones de las variables en la forma z_i^* facilitan considerablemente la interpretación y comparación de las diferencias.

Referencias

- Arsham, H., & Lovric, M. (2011). Bartlett's Test. In *International Encyclopedia of Statistical Science* (Issue March, pp. 87-88). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-04898-2_132
- Bandalos, D. L., & Finney, S.J. (2014). The Reviewer's Guide to Quantitative Methods in the Social Sciences. In G. R. Hancock & R. O. Mueller (Eds.), *The Reviewer's Guide to Quantitative Methods in the Social Sciences* (pp. 93-114). Routledge. <https://doi.org/10.4324/9780203861554>
- Bartlett, M.S. (1951). A further note on tests of significance in factor analysis. *British Journal of Statistical Psychology*, 4(1), 1-2. <https://doi.org/10.1111/j.2044-8317.1951.tb00299.x>
- Bjelland, I., Dahl, A.A., Haug, T.T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale. *Journal of Psychosomatic Research*, 52(2), 69-77. [https://doi.org/10.1016/S0022-3999\(01\)00296-3](https://doi.org/10.1016/S0022-3999(01)00296-3)
- Bollen, K.A. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons.
- Cerny, B.A., & Kaiser, H.F. (1977). A Study Of A Measure Of Sampling Adequacy For Factor-Analytic Correlation Matrices. *Multivariate Behavioral Research*, 12(1), 43-47. https://doi.org/10.1207/s15327906mbr1201_3
- CIS (Centro de Investigaciones Sociológicas) (2011a). 2910| *Opinión Pública y Política Fiscal (XXVIII)*. Catálogo de Encuestas.
- CIS (2011b). *Ficha Técnica 2910| Opinión Pública y Política Fiscal (XXVIII)*. http://www.cis.es/cis/export/sites/default/-Archivos/Marginales/2900_2919/2910/Ft2910.pdf
- CIS (2020). *Cuestionario del Estudio: Opinión Pública y Política Fiscal XXXVI SEP 19*. http://www.cis.es/cis/export/sites/default/-Archivos/Marginales/3240_3259/3259/cues3259.pdf
- CIS (2019a). *Estudio Nº 3259 Opinión Pública y Política Fiscal (XXXVI)*. http://www.cis.es/cis/opencm/ES/1_encuestas/estudios/ver.jsp?estudio=14464
- CIS (2019b). *Ficha Técnica Estudio CIS 3259* (pp. 1-2). CIS. http://www.cis.es/cis/export/sites/default/-Archivos/Marginales/3240_3259/3259/FT3259.pdf
- Clark, D.A., & Bowles, R.P. (2018). Model Fit and Item Factor Analysis: Overfactoring, Underfactoring, and a Program to Guide Interpretation. *Multivariate Behavioral Research*, 53(4), 544-558. <https://doi.org/10.1080/00273171.2018.1461058>
- Conway, J.M., & Huffcutt, A.I. (2003). A Review and Evaluation of Exploratory Factor Analysis Practices in Organizational Research. *Organizational Research Methods*, 6(2), 147-168. <https://doi.org/10.1177/1094428103251541>
- Erdfelder, E., Faul, F., Buchner, A., & Lang, A.G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149-1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: A flexible statistical power

- analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191. <https://doi.org/10.3758/BF03193146>
- Fernández, R., & Cao, R. (2020). *Simulación Estadística*. Universidad de la Coruña. <https://rubenfcasal.github.io/simbook/Simulacion.pdf>
- Ferrando, P.J., & Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología. *Papeles del Psicólogo*, 31(1), 18-33.
- Gao, S., Mokhtarian, P.L., & Johnston, R.A. (2008). Nonnormality of data in structural equation models. *Transportation Research Record*, 2082, 116-124. <https://doi.org/10.3141/2082-14>
- Hofmann, R.J. (1978). Complexity and Simplicity as Objective Indices Descriptive of Factor Solutions. *Multivariate Behavioral Research*, 13(2), 247-250. https://doi.org/10.1207/s15327906mbr1302_9
- Hogarty, K.Y., Hines, C.V., Kromrey, J.D., Ferron, J.M., & Mumford, K.R. (2005). The Quality of Factor Solutions in Exploratory Factor Analysis: The Influence of Sample Size, Communalities, and Overdetermination. *Educational and Psychological Measurement*, 65(2), 202-226. <https://doi.org/10.1177/0013164404267287>
- Horn, J.L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185. <https://doi.org/10.1007/BF02289447>
- Ibar, R. (2014). *Nueva Metodología de recogida de información para su tratamiento a través del análisis multivariante y los modelos de ecuaciones estructurales. Aplicación en el ámbito universitario*. CEU San Pablo.
- IBM Corp. (2017). *IBM SPSS Statistics for Windows, Version 25.0*. IBM Corp.
- Jefatura del Estado. (2003). Legislación Consolidada Ley 58/2003, de 17 de diciembre, General Tributaria. *BOE*, 302, 1-168.
- Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, 11(1), 94. <https://doi.org/10.1186/1471-2156-11-94>
- Jones, D.H. (1994). Book Review: Statistical Methods, 8th Edition George W. Snedecor and William G. Cochran Ames: Iowa State University Press, 1989. xix + 491 pp. *Journal of Educational Statistics*, 19(3), 304-307. <https://doi.org/10.3102/10769986019003304>
- Kaiser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3), 187-200. <https://doi.org/10.1007/BF02289233>
- Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36. <https://doi.org/10.1007/BF02291575>
- Kassambara, A. (2019). *ggcorrplot: Visualization of a Correlation Matrix using "ggplot2."*
- Kaufman, J.D., & Dunlap, W.P. (2000). Determining the number of factors to retain: A Windows-based FORTRAN-IMSL program for parallel analysis. *Behavior Research Methods, Instruments, and Computers*, 32(3), 389-394. <https://doi.org/10.3758/bf03200806>
- Lawley, R.D., & Maxwell, A.E. (1971). *Factor analysis as a statistical method* (2ª Edición). Butterworths.

- Likert, R. (1932). A technique for the measurement of attitudes [Thesis]. In R. S. Woodworth (Ed.), *Archives of Psychology* (Vol. 22, Issue 140). The Science Press.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169. <https://doi.org/10.6018/analesps.30.3.199361>
- Lorenzo-Seva, U., Timmerman, M.E., & Kiers, H.A.L. (2011). The Hull Method for Selecting the Number of Common Factors. *Multivariate Behavioral Research*, 46(2), 340-364. <https://doi.org/10.1080/00273171.2011.564527>
- Mandelbrot, B. (1961). Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling (Ingram Olkin, Sudhist G. Ghurye, Wassily Hoeffding, William G. Madow, and Henry B. Mann, eds.). In *SIAM Review* (Vol. 3, Issue 1). <https://doi.org/10.1137/1003016>
- Marais, I., & Andrich, D. (2007). *RUMMs. Rasch Unidimensional Measurement Models Simulation Studies Software, Version 2.0* (2nd ed.). The University of Western Australia.
- Martín-Pliego, F.J. (2011). *Introducción a la Estadística Económica y Empresarial*. Paraninfo.
- Microsoft Corporation. (2020). *Microsoft Excel*.
- Muthen, B., & Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model. *British Journal of Mathematical and Statistical Psychology*, 45(1), 19-30. <https://doi.org/10.1111/j.2044-8317.1992.tb00975.x>
- O'Connor, B.P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, and Computers*, 32(3), 396-402. <https://doi.org/10.3758/bf03200807>
- O'Connor, B.P. (2020). *EFA.dimensions: Exploratory Factor Analysis Functions for Assessing Dimensionality*.
- Peña, D. (2002). *Análisis de Datos Multivariantes*. McGraw-Hill.
- Pettersson, E., & Turkheimer, E. (2010). Item selection, evaluation, and simple structure in personality data. *Journal of Research in Personality*, 44(4), 407-420. <https://doi.org/10.1016/j.jrp.2010.03.002>
- Price, L.R. (2014). *Investigating Perfect Model Fit in Higher-Order Confirmatory Factor Analysis through Exploratory / Confirmatory and Bayesian Approaches HCFA 1 Investigating Perfect Model Fit in Higher-Order Confirmatory Factor Analysis through Exploratory / Confirmatory*. May.
- R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- Rahmer, B.D.J., Garzón Saénz, H., & Solana, J. (2020). Análisis comparativo de indicadores de capacidad multivariante. El caso del clúster manufacturero cartagenero Comparative analysis of multivariate capacity indicators. The case of the Cartagena manufacturing cluster. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 29, 172-189.
- Revelle, W. (2020). *psych: Procedures for Psychological, Psychometric, and Personality Research*.
- Revelle, W., & Rocklin, T. (1979). Very Simple Structure: An Alternative Procedure For Estimating The Optimal Number Of Interpretable Factors. *Multivariate Behavioral Research*, 14(4), 403-414. https://doi.org/10.1207/s15327906mbr1404_2

- Rstudio Team (2020). *RStudio: Integrated Development for R*. Rstudio,PBC. <http://www.rstudio.com>
- Shapiro, S.S., & Wilk, M.B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), 591-611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Sharma, S. (1996). *Applied Multivariate Techniques*. John Wiley and Sons Inc.
- Smith, E.P. (1997). Methods of multivariate analysis. In *Journal of Statistical Planning and Inference* (Vol. 59, Issue 1). [https://doi.org/10.1016/S0378-3758\(96\)00098-5](https://doi.org/10.1016/S0378-3758(96)00098-5)
- Smith, R.M., Schumacker, R.E., & Bush, M.J. (1998). Using item mean squares to evaluate fit to the Rasch model. *Journal of Outcome Measurement*, 2(1), 66-78.
- Spearman, C. (1927). The Abilities of Man their Nature and Measurement. *Nature*, 120(3014), 181-183. <https://doi.org/10.1038/120181a0>
- Stegmann, M.B., Sjöstrand, K., & Larsen, R. (2006). Sparse modeling of landmark and texture variability using the orthomax criterion. *Medical Imaging 2006: Image Processing*, 6144, 61441G. <https://doi.org/10.1117/12.651293>
- Steiger, J.H. and Lind, J. (1980). Statistically-based tests for the number of common factors. *Paper Presented at the Annual Spring Meeting of the Psychometric Society*.
- Tennant, A., & Pallant, J. F. (2012). The Root Mean Square Error of Approximation (RMSEA) as a supplementary statistic to determine fit to the Rasch model with large sample sizes. *Rasch Measurement Transactions*, 25(4), 1348-1349. <https://www.rasch.org/rmt/rmt254d.htm>
- Thurstone, L.L. (1947). *Multiple-factor Analysis*. The University of Chicago Press.
- Velicer, W.F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41(3), 321-327. <https://doi.org/10.1007/BF02293557>
- Welch, B.L. (1947). The Generalization of Student's, Problem when Several Different Population Variances are Involved. *Biometrika*, 34(1/2), 28. <https://doi.org/10.2307/2332510>
- Welch, B.L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, 29(3-4), 350-362. <https://doi.org/10.1093/biomet/29.3-4.350>
- West, S.G., Finch, J.F., & Curran, P.J. (1995). Structural Equation Models with Non Normal Variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues, and Applications* (pp. 56-75). Sage Publications.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (New York). Springer-Verlag New York. <https://ggplot2.tidyverse.org>