



Dpto. de Estadística
Universidad de Salamanca

DRA. M^a PURIFICACIÓN GALINDO VILLARDÓN

*Profesora Titular del Departamento de Estadística de la Universidad de
Salamanca*

CERTIFICA que **D. Juan Antonio Carbonell** ha realizado en la Universidad de Salamanca, bajo su dirección, el trabajo que para optar título de Máster en Análisis Avanzado de Datos Multivariantes, presenta con el título ***Análisis multivariante de la precariedad en España, Italia, Francia, Portugal y Bélgica***, autorizando expresamente su lectura y defensa.

Y para que conste, firma el presente certificado en Salamanca a 19 de julio de 2013.

M^a Purificacion Galindo Villardón

ANÁLISIS MULTIVARIANTE DE LA PRECARIEDAD EN ESPAÑA, ITALIA, FRANCIA, PORTUGAL Y BÉLGICA



Dpto. de Estadística
Universidad de Salamanca

Trabajo para optar al título de
Máster en
Análisis Avanzado de Datos
Multivariantes
por la Universidad de Salamanca.

Presenta:

Juan Antonio Carbonell Asíns

Salamanca
2013
UNIVERSIDAD DE SALAMANCA

Departamento de Estadística
Máster en Análisis Avanzado de Datos Multivariantes
Trabajo Fin de Máster

Análisis multivariante de la precariedad en España, Italia, Francia, Portugal y Bélgica

Autor: D. Juan Antonio Carbonell Asíns
Tutora: Dra. M^a Purificación Galindo Villardón

2013

Índice

Resumen.....	5
1. Introducción	6
2. Marco Teórico	9
3. Encuestas.....	15
3.1. Encuesta de Condiciones de Vida (2009)	16
3.2. Encuesta de Condiciones de Vida de Eurostat (2009)	16
3.3. Organización de la encuesta	17
3.4. Variables usadas para el estudio de los aspectos de la precariedad	18
4. Métodos Estadísticos	26
4.1. Análisis Factorial Exploratorio	27
4.2. Modelo Lineal General	29
4.3. HJ-Biplot	33
4.4. Statis/Statis Dual.....	35
5. Resultados.....	41
5.1. Descriptivos	42
5.2. Análisis de la dimensión de la precariedad	44
5.3. Evaluación de la eficiencia de los indicadores	46
5.4. Caracterización multivariante de la precariedad en Comunidades Autónomas ...	49
5.5. Estudio de las estructuras multifactoriales en España, Francia, Italia, Bélgica y Portugal.....	56
6. Discusión	64
7. Conclusiones	68
8. Bibliografía	72

Resumen

Se lleva a cabo un estudio multivariante de la precariedad en sus múltiples dimensiones tanto para España como para Italia, Portugal, Bélgica y Francia para el año 2009. Este trabajo se basó en el proyecto de investigación del Plan Nacional de I+D+I *Respuestas sociales a la crisis y procesos de precarización de la vida en la sociedad contemporánea: Bélgica, España, Francia, Italia y Portugal* financiado por el Ministerio de Ciencia e Innovación, (Ref. CSO 2011-23252). Construcción de indicadores, que aquí hemos sometido a la revisión desde la teoría estadística, se han discutido indicadores para medir la multidimensionalidad de la precariedad y se han realizado propuestas alternativas. Una vez obtenidos los indicadores y validados mediante modelos lineales con variables de control se propone un HJ-Biplot para el estudio de la relación de las Comunidades Autónomas con siete indicadores de precariedad. Posteriormente, se estudian las estructuras de correlación entre los indicadores, en los cinco países, mediante un STATIS DUAL y se encuentra la estructura consenso que es convenientemente analizada, siendo posible así, explicar los patrones de precariedad comunes en todos los países. Ninguna de las dos últimas técnicas ha sido utilizada anteriormente en el estudio de la precariedad.

Palabras Clave: STATIS, HJ-Biplot, Precariedad, Sociología

1.Introducción

El presente trabajo es fruto de la colaboración de tres departamentos: el departamento de Sociología II de la Universidad del País Vasco, el departamento de Sociología y Antropología Social de la Universidad de Valencia, y el departamento de Estadística de la Universidad de Salamanca. Durante la realización del master estuve colaborando en la elaboración del proyecto de investigación “Respuestas Sociales a la Crisis y Procesos de Precarización de la Vida en la Sociedad Contemporánea: Bélgica, España, Francia, Italia y Portugal”.

Una parte de ese Proyecto tenía como objetivo la creación de indicadores que permitiesen evaluar los diferentes aspectos de la precariedad, cuyo desarrollo ha sido plasmado en un documento de trabajo (nos referiremos a él como documento base). Este documento “*Code Book of the Precariousness Indicators Built with the EU-SILC 2009 Survey (only the Belgian, French, Italian, Portuguese and Spanish Samples)*”, redactado por Carles Simó (UV) y Martin Wagener (UCL) contiene, además, la sintaxis del proceso de tratamientos de datos que ha llevado a su construcción.

Mi motivación para elegir este tema es múltiple. Por un lado, es un trabajo en el que he participado desde la reflexión sociológica y, en consecuencia, ello me permite acercarme a la elaboración de alternativas de creación de los indicadores que tengan en cuenta la teoría sociológica junto con la teoría estadística y, finalmente, la posibilidad de experimentar con técnicas innovadoras para el estudio cuantitativo de la precariedad.

En este trabajo se proponen cuatro objetivos generales:

1. Abordar el estudio de los indicadores de precariedad desde una perspectiva multifactorial y proponer indicadores alternativos, sociológicamente eficientes, si procede.
2. Analizar las componentes de precariedad preponderantes en las diferentes Comunidades Autónomas del Estado Español.
3. Estudiar las estructuras multifactoriales en diferentes países, y encontrar las componentes estables en todos ellos definiendo una configuración consenso.
4. Probar que las técnicas BIPLLOT y STATIS, nunca utilizadas hasta ahora en estudios de precariedad, pueden ser una herramienta útil en estos

estudios y en cualquier otro del campo de la Sociología que requiera inspeccionar matrices multivariantes de dos y tres vías.

El desarrollo de este trabajo está estructurado de la siguiente forma: En primer lugar se contextualiza el objeto de estudio mediante un marco teórico en el que centraremos la precariedad dentro del contexto de la exclusión social y explicaremos las diversas definiciones de precariedad y aquella en la que nos basaremos para este estudio. Una vez hecho esto pasaremos al apartado de encuestas donde explicaremos las dos encuestas utilizadas así como la preparación que han requerido para el estudio y las variables que finalmente se usan. El siguiente apartado abordará los métodos estadísticos en donde se explicarán las técnicas usadas de forma exhaustiva así como el software usado. Posteriormente entraremos en el análisis y discusión de los resultados. A continuación se debatirán las discusiones que surgen a raíz de este trabajo. Finalmente se concluirá el trabajo con las conclusiones y la bibliografía utilizada.

2. Marco Teórico

La precariedad es un concepto complejo cuya definición ha ido cambiando con el tiempo.

“Considerando su procedencia etimológica precarius proviene de precor —igual que (to) pray: orar—, «aquello que se obtiene mediante la súplica». Con la modernidad —cuando ya no hay divinidad a la que orar que garantice seguridad—, surge un nuevo protector de la comunidad, del individuo-ciudadano, cristalizado en la figura del Estado-Nación. La transformación del Estado moderno en una sociedad salarial, en la época fordista¹, supuso «la disposición de una base de recursos y garantías sobre la cual el trabajador puede apoyarse para gobernar el presente y dominar el futuro»“(Castel, 1996).

La precariedad va unida al concepto de **clase**. Cuando se habla de **clase** en sentido sociológico se distinguen dos tipos de clase social. Por un lado tenemos la clase social objetiva y, por el otro, la clase social subjetiva. En el primer caso se refiere a la medición de la clase por el investigador y, en el otro, la percepción del individuo dentro de su clase social. En este trabajo se aplicara la perspectiva objetiva de la clase social dado que el objetivo es la medición del nuevo concepto de precariedad vital.

“Las dos grandes teorías son la marxista y la weberiana, a partir de ellas se han desarrollado diversos avances y operacionalizaciones. Actualmente podríamos decir que las divisiones de clase se basan en tres criterios principales: la posición de la persona en la **estructura ocupacional**, la posición en las **estructuras de autoridad** (de cuántas personas recibe órdenes y a cuántas les da) y la posesión de propiedades que den beneficios (la **estructura de la propiedad**). La intersección de estos tres criterios es lo que determina las divisiones de clase más o menos diferenciadas y las posiciones de los individuos en ellas. Atendiendo a estas tres dimensiones se han establecido diferentes clases sociales como compartimentos estanco.” (Simo, Sanchez, Llopis y Castelló, 2011).

Además es necesario describir otros conceptos que están íntimamente relacionados con la precariedad, estos son: *exclusión social*, *pobreza* y *vulnerabilidad*. Todos ellos tienen en común un concepto: **desigualdad social**. Definimos este concepto como “la condición por la cual las personas tienen un acceso desigual a los recursos, servicios y posiciones que la sociedad valora.” (Simo, Sanchez, Llopis y Castelló, 2011).

¹ El término fordismo se refiere al modo de producción en cadena que llevó a la práctica Henry Ford; fabricante de automóviles de Estados Unidos. Para más información sobre este tema consultar el libro “Trabajo y paro en la sociedad postindustrial” (Sanchis, 2011).

Actualmente es necesario tener en cuenta mucho más otras desigualdades, a parte del factor económico, para la definición de las desigualdades sociales. Existen nuevos mecanismos de segregación que impiden o que dificultan a las personas el acceso a las diferentes esferas sociales como el trabajo, la formación o la sanidad. Desde la crisis financiera que estalló en España en el 2008 pero que fue gestándose mucho tiempo atrás, el Estado de Bienestar ha ido debilitándose facilitando nuevas formas de **exclusión social**.

Existen pues nuevos desequilibrios que se están consolidando como fundamentales en la marginación social, política, económica y laboral para ciertos individuos. Por lo tanto, aunque el factor económico es crucial a la hora de determinar la exclusión social, no es posible centrarse solo en esta dimensión sino que los nuevos tiempos nos fuerzan a buscar respuestas en otras dimensiones de la exclusión.

Existen tres nuevos procesos de exclusión social:

- En primer lugar, la fragmentación de la sociedad en un nuevo cuerpo mucho más complejo, debido a la creciente diversidad étnica y cultural que es producto de las migraciones que las políticas públicas no han podido manejar. Por otro lado, el envejecimiento de la población que tiene una clara dependencia económica y sanitaria son nuevos colectivos con alta precariedad.
- En segundo lugar, la nueva economía postindustrial ha desencadenado una transición hacia la información en vez de en a la producción mercantil con un capitalismo menos regulado que aprovecha las ventajas de la economía global ha hecho que existan nuevos grupos que sufren estas consecuencias. La forma más fácil de entender esto es que los jóvenes de hoy en día ya no se insertan en el mercado laboral industrial estable sino que requieren de trayectorias laborales mucho más complejas y precarias.
- Finalmente, la flexibilidad del mercado mediante la desregulación laboral ha generado que muchos individuos tengan condiciones precarias contribuyendo a la aparición de nuevos espacios de exclusión y vulnerabilidad social.

La **precariedad** proviene del concepto de “exclusión social”.

Son muchas las definiciones existentes sobre la precariedad pero vamos a basarnos en los escritos del catedrático de Ciencias Políticas de la UAB, Joan Subirats debido a que es el que ha sido usado para definir el concepto dentro del marco de la

investigación en la que se basa este trabajo. Según un artículo publicado en la versión catalana de El País en Mayo del 2008, Subirats define el marco histórico de la precariedad de la siguiente forma:

“Hace ya años, la Fundación Friedrich Ebert, vinculada al partido socialdemócrata alemán y para nada susceptible de ser tachada de radicalismo alternativo, publicó un estudio, *Sociedad en proceso de reforma*, en el que utilizaba el término *Prekariat* para referirse a los grandes cambios que se estaban produciendo en un mercado de trabajo que ha ido transformando a los asalariados estables en trabajadores permanentemente en precario. Si buscamos en fuentes autorizadas, encontramos “inseguro, apurado, escaso,... se aplica a la manera de estar en un cargo o situación cuando no se está con plena seguridad o derecho” (*Diccionario María Moliner*). El término había sido ya utilizado por un grupo de activistas italianos vinculados al movimiento libertario y fue después popularizado a través del Fórum Social Europeo. Hoy la expresión nos recuerda la proliferación casi estructural de formas de vinculación laboral que poco tienen que ver con lo que conocíamos en plena época fordista.”

Vemos pues que, originariamente, la precariedad tenía una sola dimensión, la dimensión laboral.

Progresivamente la precariedad aparece vinculada al Estado de Bienestar y sus políticas de integración y exclusión. “Siguiendo como ejemplo el caso francés, la expansión del término se produce primero en relación a la pobreza (Pitrou, 1978, es la primera en asociarlo a las familias vulnerables), después con el estatus del empleo (Schnapper y Villac, 1989) y en la década de 1990 en relación con el trabajo (Paugam, 2007). Mientras, en el ámbito anglosajón, a partir de la década de 1990, la precariedad se pone en relación con la flexibilidad y la corrosión” (Sennett, 2000).

Siguiendo con la definición de precariedad podemos extraer del artículo publicado por Benjamin Tejerina y otros: “Precariedad vital y juventud vasca Condiciones sociales y estrategias biográficas para llevar una vida normal” otra aproximación al concepto de precariedad:

“La precariedad es un concepto que aúna la condición y la situación personal en la relación del individuo con el medio. Es una forma de definir la relación del individuo con el medio social. La precariedad es el estado al que se llega mediante procesos de precarización entendidos como ocupaciones de espacios de vida desinstitucionalizados. Los procesos de precarización afectan a la identidad en la medida en que los individuos pierden o ven alteradas las modalidades de apego del yo al nosotros y a los tú. La precarización como proceso toca diferentes dimensiones que tienen que ver con pérdidas o entradas en zonas de riesgo y que hacen referencia a

limitaciones en recursos y capacidades de los individuos: trabajo, remuneración, consumo, residencia, cualificación educativa, entorno, vida familiar y afectiva, relaciones sociales, salud y participación cívica. En la precarización como proceso también participan las instituciones por medio de las propias prácticas de la acción pública o por la ausencia de ésta. Además, las instituciones normativizan la precariedad en la medida en que enseñan a los individuos a moverse en ella.”

Ésta será la definición de precariedad que usaremos en el estudio y que contempla cinco aspectos de la precariedad:

1. **Problemas financieros:** Dentro del ámbito económico podemos distinguir tres factores esenciales de exclusión: la pobreza, las dificultades financieras del hogar y la dependencia económica de la protección social. Son factores, los tres, complementarios entre sí y que nos muestran distintos grados, momentos o aspectos de la exclusión económica.
2. **Condiciones:** Se refiere a las condiciones de la casa. Ejemplos de esta precariedad sería que la casa fuese demasiado oscura o no fuese lo suficiente caliente para vivir dignamente.
3. **Vecindad:** Entendemos esta precariedad como la relacionada con los alrededores de la vivienda. Ejemplo de este tipo de precariedad sería el hecho que existiese mucho ruido o crimen.
4. **Salud:** La precariedad en salud es entendida como la falta de asistencia sanitaria, ya sea por no poder permitirse la asistencia como por ser un paciente crónico.
5. **Socio-económico:** Usemos la definición de Subirats para este apartado: “En este factor se concentran todos los elementos de precariedad laboral: el trabajo a tiempo parcial, el trabajo sin contrato, los empleos de bajo salario, e incluso parte de los empleos de baja cualificación que, en muchos casos, coinciden con las anteriores fórmulas. Se trata pues de un factor muy claro de vulnerabilidad social, que viene determinado por una inclusión relativa o muy frágil en el mercado de trabajo. Relativa por el hecho de que algunas de estas formas de precariedad dejan a las personas en una situación de desprotección prácticamente total; y frágil porque las somete a una presencia parcial, intermitente y en definitiva no plenamente garantizada en el mercado de trabajo. Así, por ejemplo, el trabajo sin contrato, al no cotizar, no está sujeto a garantías ni derechos laborales. De modo similar ocurre con el trabajo a tiempo parcial que, aunque pudiera ser estable, a menudo no llega a los mínimos necesarios

de cotización para poder ser beneficiario de una prestación por desocupación o una pensión de jubilación.”

3. Encuestas

En este estudio abordaremos el análisis de la precariedad en España, y además compararemos estos resultados con los de otros países. Por esta razón, además de analizar la Encuesta de Condiciones de Vida, facilitada por el Instituto Nacional de Estadística (2009), código de contrato (Nº EU-SILC/2011/17), consideraremos también la Encuesta de Condiciones de Vida Eurostat (2009), código de contrato (Nº EU-SILC/2011/17), ya que en la primera solo están los datos españoles.

3.1. Encuesta de Condiciones de Vida (2009)

La Encuesta de Condiciones de Vida ha sido utilizada para medir la privación material y privación múltiple (Ayala, 2008), sin embargo la vamos a utilizar aquí para operacionalizar nuestro concepto de precariedad, construir indicadores sintéticos por área con el fin de poder medir niveles de precariedad. Esta encuesta no sólo tiene una perspectiva monetaria, ocupacional y laboral de los individuos y hogares como las que incluye la Encuesta de Población Activa (EPA) sino que incluye además información acerca de las condiciones de vida (salud, vivienda, relaciones, etc.), lo que permite acercarnos a la totalidad de las dimensiones propuestas por Subirats (et al 2004). Finalmente, puesto que el apartado de información sociodemográfica de la encuesta es muy completo, ello va a permitir la localización de los grupos más vulnerables (Ayala et al 2008).

El número total de individuos encuestados fue de 36.865 y la encuesta cuenta con un total de 740 variables.

3.2. Encuesta de Condiciones de Vida de Eurostat (2009)

La agencia Europea Eurostat ha elaborado una encuesta de condiciones de vida cogiendo las diversas encuestas de condiciones de vida de cada país y uniéndolas en una sola por lo que contiene menos variables la versión española, sin embargo, ha sido posible elaborar los mismos indicadores sin problema alguno. Esta encuesta con código de contrato (Nº EU-SILC/2011/17) ha sido usada para el análisis del STATIS DUAL al tener individuos x variables x países y también viene protegida por derechos de explotación. Los países que han sido analizados son: Bélgica, España, Francia, Italia y Portugal. Es muy importante comparar estos cinco países pues todos ellos están dentro del modelo conservador y de tradición católica (Esping Andersen, 2000), pero con una clara diferencia entre los más familiaristas (España,

Portugal e Italia) y los que tienen un régimen de bienestar más desarrollado (Bélgica y Francia).

Se realizaron un total de 141.406 entrevistas desglosadas en: España (36.865), Bélgica (14.721), Francia (25.611), Italia (51.196) y Portugal (13.013). El número total de variables en la encuesta es de 710.

3.3. Organización de la encuesta

Con el objetivo de poder trabajar al mismo tiempo con las informaciones relativas a personas y las relativas al hogar, hemos realizado la fusión de las cuatro bases de datos que contiene la encuesta. Nuestro trabajo tiene como unidad de análisis a las personas entrevistadas, individuos por tanto. Sin embargo, existen informaciones relativas al hogar que no pueden ser obviadas en el proceso de análisis. Por ello, hemos procedido a integrar en la misma base de datos las informaciones relativas a los ficheros:

1. Fichero D: Fichero de datos básicos del hogar.
2. Fichero R: Fichero de datos básicos de la persona.
3. Fichero H: Fichero de datos detallados del hogar.
4. Fichero P: Fichero de datos detallados de los adultos.

Para realizar dicha fusión de variables en un mismo fichero completo de personas, hemos utilizado como variable de referencia la RB030 (identificador transversal de la persona), que es coincidente con la PB030, y además contiene la información de la DB030 (identificador del hogar), que es la misma que la de la HB030.

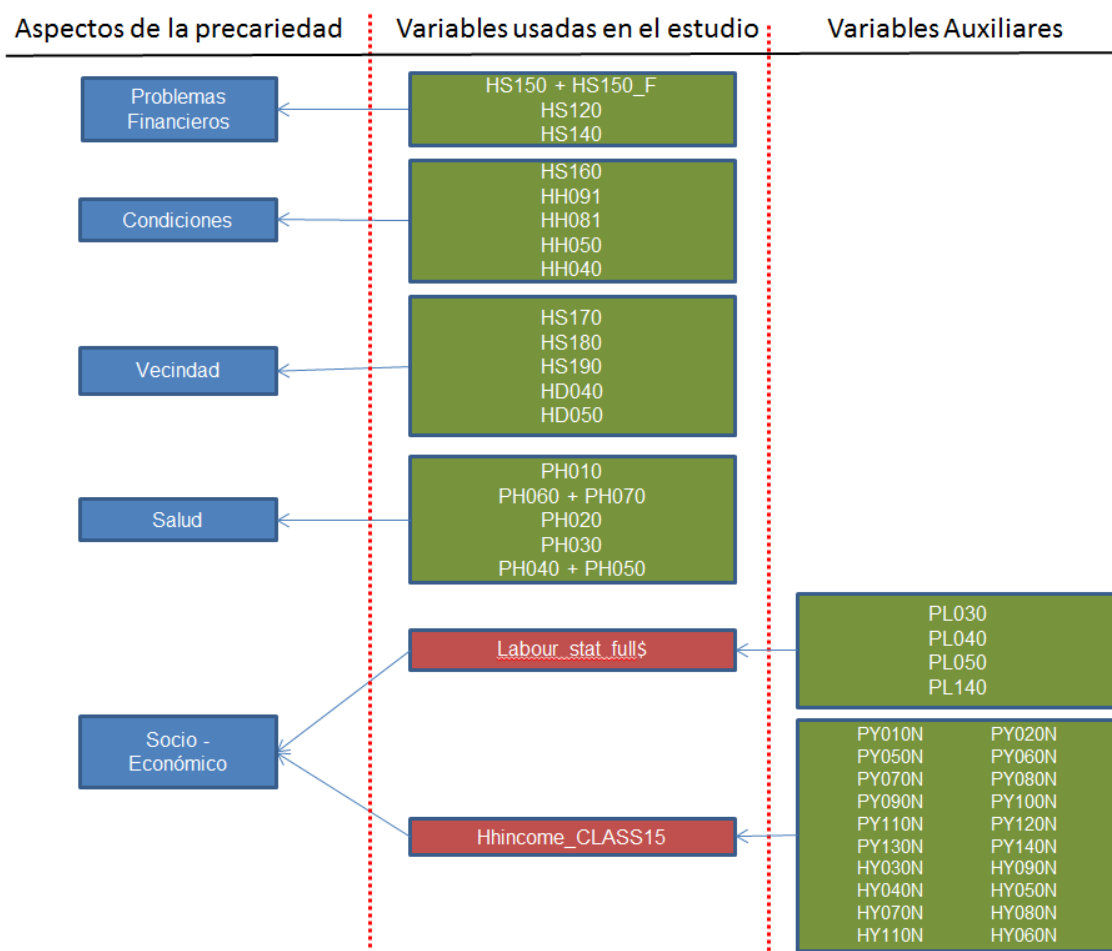
Esto nos permite imputar datos del hogar a cada individuo, pero hemos de tener en cuenta que las medias de las variables de hogar que puedan calcularse con esta nueva base de datos fusionada estarán ponderadas por el número de miembros del hogar.

Las muestras han sido ponderadas, según se indica en la documentación, con la variable RB050 (factor transversal de ponderación. Aunque las bases de datos incluyen información parcial sobre la población de menos de 16 años, la hemos excluido con la ponderación, por tanto, todo el trabajo hace referencia a la población española de 16 años o más.

3.4. Variables usadas para el estudio de los aspectos de la precariedad

La figura 1 muestra la relación entre aspectos de la precariedad, variables usadas y variables auxiliares.

Figura 1. Esquema de los aspectos de precariedad y las variables usadas



Teóricamente estudiamos cinco aspectos de la precariedad tal y como se comentó en el marco teórico. Un grupo de expertos sociólogos fueron consultados para identificar que variables formarían cada aspecto y se llegó a las siguientes conclusiones:

1. Problemas Financieros:

- i. Carga financiera debido al pago de compras a plazos o préstamos. Está formada por dos variables, por un lado, HS150 que se refiere a cuánta carga existe y la variable *flag* HS150_F que indica si existe o no carga.
- ii. Carga financiera debido al pago de la hipoteca (HS140).
- iii. Capacidad para llegar a fin de mes (HS120).

2. **Condiciones:**

- i. Demasiado oscuro, poca luz (HS160).
- ii. Váter de uso privado (HH091).
- iii. Baño o ducha (HH081).
- iv. Posibilidad de tener la casa caliente (HH050).
- v. Existencia de goteras (HH040).

3. **Vecindad:**

- i. Demasiado ruido (HS170).
- ii. Polución (HS180).
- iii. Crimen (HS190).
- iv. Vecindario sucio (HD040).
- v. Servicios públicos en mal estado (HD050).

4. **Salud:**

- i. Autopercepción del estado general de salud (PH010).
- ii. Motivos por los que no se ha visitado un médico. Está formada por dos variables, por un lado, PH060 que se refiere a los motivos en sí y la variable *flag* PH070 que indica si existe o no enfermedad que requiera ir al médico.
- iii. Sufre de una enfermedad crónica (PH020).
- iv. Limitación en las actividades debido a tener problemas de salud (PH030).
- v. Motivos por los que no se ha visitado un dentista. Está formada por dos variables, por un lado, PH040 que se refiere a los motivos en sí y la variable *flag* PH050 que indica si existe o no enfermedad que requiera ir al dentista.

Tenemos pues 18 variables directamente consideradas en la encuesta y dos creadas a partir de variables de la encuesta (*labour_stat_full\$* y *hhincome_CLASS15*). A continuación describimos el proceso de creación de dichas variables.

3.4.1. *Dimensión Socio-económica:*

Este apartado está subdividido en dos partes, por un lado la variable *labour_stat_full\$* relacionada con la precariedad laboral y, por el otro, la variable *Hhincome_CLASS15* relacionada con la precariedad de ingresos:

La variable *labour_stat_full\$* ha sido elaborada mediante varias variables de la propia encuesta:

1. **“La situación laboral:** Se trata de la situación laboral y del trabajo principal actuales que define la propia persona entrevistada (variable PL030).
2. **La situación profesional.** Esta pregunta únicamente puede ser respondida por aquellas personas que han trabajado alguna vez (variable PL040). Aquéllas que habían entrado en el mercado laboral pero que en el momento de la encuesta no trabajaban se les preguntó sobre su último trabajo principal.
3. **La ocupación.** La ocupación se define como la clase o tipo de trabajo desarrollado con especificación del puesto de trabajo desempeñado y se codifica siguiendo la Clasificación Nacional de Ocupaciones de 1994 (CNO-94) a dos dígitos. Sin embargo, en los ficheros de usuario dicha codificación se ha adaptado a la clasificación Internacional de Ocupaciones (ISCO88) a excepción de los trabajadores de las fuerzas armadas no han sido incluidos en la encuesta. Igual que en la situación profesional, a las personas que no trabajan en el momento de la encuesta también se les pregunta. En este caso se trata de la ocupación en el último trabajo principal. La variable que recoge esta información es la PL050.
4. **El tipo de contrato (PL140)** En aras de una simplificación, esta pregunta utiliza rúbricas muy generales en las que se encuadran los distintos tipos de contratos existentes. Se trata de distinguir entre contratos fijos de duración indefinida (cuando no existen criterios objetivos para la finalización de la relación laboral) o contrato temporal de duración determinada (el fin de la relación laboral queda reflejado en el contrato o está determinado por la expiración de cierto plazo, la realización de una tarea determinada, la reincorporación de un empleado/a reemplazado/a temporalmente, la realización de un periodo de prácticas).”

Por otro lado, la variable hh_income_class15 ha sido construida de la siguiente forma:

Hemos organizado los ingresos en tres grandes capítulos, en función de su origen:

1. **Rentas totales individuales**, en las que consideramos los ingresos netos obtenidos por medio de la actividad económica del individuo y que, por tanto, son ingresos relacionados con la concurrencia del entrevistado en el mercado y su aportación a los factores de producción

de la economía (trabajo y capital), o su autoconsumo. En ellas se incluyen:

- i. Rentas salariales (monetarias y no monetarias): PY010N y PY020N. Se entiende por rentas salariales las remuneraciones totales, monetarias o no monetarias, que debe pagar un empleador a su asalariado como contrapartida del trabajo realizado por éste durante el período de referencia de los ingresos.
 - ii. Beneficios o pérdidas: hace referencia a los rentas obtenidas por la actividad empresarial, PY050N. Ingresos percibidos durante el período de referencia de los ingresos por las personas, para sí o en relación con los miembros de su familia, como resultado del ejercicio presente o pasado de una actividad por cuenta propia. Se consideran actividades por cuenta propia aquellas en las que
 - iii. la remuneración depende directamente de los beneficios (o posibles beneficios) derivados de los bienes y servicios producidos. (PY060N)
 - iv. Autoconsumo: se refiere a aquello consumido de la propia producción, PY070N. Se refiere al valor de los alimentos y las bebidas producidos y consumidos dentro del mismo hogar. El valor de los bienes producidos para autoconsumo corresponde al precio de mercado de los bienes producidos, una vez deducidos todos los gastos soportados en la producción
 - v. Rentas privadas de pensiones: PY080N. Prestaciones procedentes de los sistemas en los que el beneficiario ha cotizado únicamente de manera voluntaria, independientemente de su empleador o de la Administración.
2. **Ayudas sociales recibidas**, aquellas de titularidad individual que recibe el entrevistado de resultados de la aplicación de las políticas sociales del estado. Concretamente se refiere a las transferencias del estado, ya que, de hecho, entre ellas no se cuantifican las políticas de carácter general, como por ejemplo el acceso a la educación o a la sanidad pública. Se incluyen:

- i. Prestaciones por desempleo: PY090N. Prestaciones que sustituyen, íntegra o parcialmente, la pérdida de renta de un trabajador debido a la pérdida de un empleo remunerado, compensan por la pérdida de ingresos derivada del desempleo parcial; sustituyen, íntegra o parcialmente, la pérdida de renta de un trabajador que se jubila de un empleo remunerado antes de la edad legal de jubilación como consecuencia de una reducción de plantilla por motivos económicos; contribuyen a costear los gastos de formación o reciclaje de personas que buscan trabajo; o ayudan a las personas desempleadas a hacer frente a los gastos de viaje o traslado para conseguir trabajo.
- ii. Prestaciones por jubilación: PY100N. Esta prestación se refiere a la protección social que se concede para afrontar el riesgo vinculado a la edad, la pérdida de ingresos, la renta inadecuada, la falta de independencia para llevar a cabo las tareas cotidianas, la participación reducida en la vida social, etc. Las prestaciones por vejez incluyen aquellas que proporcionan una renta de sustitución cuando la persona mayor se retira del mercado laboral o garantizan una renta determinada cuando una persona ha alcanzado una determinada edad
- iii. Prestaciones por supervivencia: PY110N. Se refieren a aquellas que aportan una renta temporal o permanente a personas que no han alcanzado la edad de jubilación y han sufrido la pérdida del cónyuge, la pareja o el pariente más cercano, generalmente cuando éste último representaba el sustento principal del beneficiario. Los supervivientes con derecho a la prestación pueden ser el cónyuge o ex cónyuge de la persona fallecida, sus hijos, sus nietos, sus padres y demás parientes. En algunos casos, la prestación también puede concederse a una persona ajena a la familia.
- iv. Prestaciones por enfermedad: PY120N. Se refieren a las prestaciones monetarias que sustituyen, íntegra o

parcialmente, la pérdida de ingresos durante un período de incapacidad temporal para el trabajo, por causa de enfermedad o lesión.

- v. Prestaciones por invalidez: PY130N. Se refieren a aquéllas que aportan una renta a personas que no han alcanzado la edad normal de jubilación y cuya capacidad para trabajar y percibir ingresos se ha visto deteriorada por encima de un nivel mínimo establecido por la Ley a causa de una incapacidad física o mental
- vi. Ayudas de estudio: PY140N. Se refieren a subvenciones, becas y demás ayudas para estudios que reciben los estudiantes.

3. **Ingresos del hogar imputados a individuos.** A parte de los ingresos individualizados declarados por los entrevistados, también existen ingresos que se recogen en el cuestionario del hogar, como ingresos de carácter colectivo, y que hemos procedido a imputar de forma individualizada equitativa a los miembros de cada hogar. Son ingresos que provienen tanto del mercado como del estado, pero que se considera que benefician al conjunto del hogar. Como hacemos un tratamiento individualizado de la información, hemos considerado conveniente hacer la imputación per capita de estos ingresos. Esta forma de proceder facilita que todos los individuos tengan algún ingreso. En estos ingresos se incluyen:

- i. Alquiler imputado del hogar: HY030N. Se refiere al valor que se imputará a todos los hogares que no declaren pagar un alquiler completo, ya sea porque son ocupantes-propietarios, porque ocupan una vivienda alquilada a un precio inferior al de mercado porque ocupan una vivienda a título gratuito. El alquiler imputado sólo se calcula en el caso de las viviendas (y demás locales asociados, como garajes) utilizadas por los hogares como residencia principal.
- ii. Intereses, dividendos y ganancias netos de inversiones de capital en empresas no constituidas en sociedad: HY090N. Se refiere al importe de los intereses procedentes de los activos como cuentas bancarias, certificados de depósito, bonos, etc., los dividendos, y

- ganancias de inversiones de capital en una empresa no constituida en sociedad, en la cual no trabaja la persona, percibidos durante el período de referencia de los ingresos, una vez deducidos los gastos soportados.
- iii. Renta neta procedente del alquiler de una propiedad o terreno: HY040N. Se refiere a los ingresos percibidos durante el período de referencia en concepto de alquiler de una propiedad (por ejemplo, el alquiler de una vivienda) una vez deducidos los gastos, como el reembolso de los intereses de la hipoteca, las reparaciones menores, el mantenimiento, el seguro, etc.
 - iv. Ayuda por familia/hijos: HY050N. Proporcionan ayuda económica a los hogares para criar a los hijos u otras personas menores a su cargo.
 - v. Ayuda para vivienda: HY070N. Ayudas a los hogares para hacer frente a los gastos asociados a la vivienda.
 - vi. Transferencias periódicas monetarias percibidas de otros hogares: HY080N. Se refieren a cantidades periódicas monetarias percibidas durante el período de referencia procedentes de otros hogares o personas y que incluyen: la pensión de manutención y las ayudas obligatorias a los hijos, la pensión de manutención y las ayudas voluntarias a los hijos, recibidas de manera periódica, la ayuda monetaria periódica procedente de personas distintas de los miembros del hogar, la ayuda monetaria periódica procedente de hogares residentes en otros países...
 - vii. Renta neta percibida por los menores de 16 años: HY110N. Incluye la renta bruta percibida por todos los miembros del hogar menores de 16 años durante el período de referencia de los ingresos
 - viii. Ingresos por asistencia social: HY060N. Se refieren a los « excluidos sociales» o a «aquellos que corren el riesgo de convertirse en excluidos sociales». El carácter general de este apartado permite identificar diversos grupos objetivo como indigentes, emigrantes, refugiados, toxicómanos, alcohólicos, víctimas de la violencia, entre otros.

Para el cálculo per cápita de cada uno de estos componentes se ha dividido la cantidad consignada en las variables de hogar citadas por el número de miembros del hogar, recogido en la variable HX040.

4. Métodos Estadísticos

4.1. Análisis Factorial Exploratorio

Recordamos el primer objetivo de este trabajo: Abordar el estudio de los indicadores de precariedad desde una perspectiva multifactorial. Para ello usaremos la técnica llamada Análisis Factorial (AF).

El objetivo principal del AF es la reducción de dimensionalidad de los datos. Esto es, a partir de un conjunto de variables observadas, explicarlas mediante un número menor de factores latentes o variables no observadas.

Esta técnica surge a manos del psicólogo inglés Spearman, C. (1904) pero no será usada hasta los años 30 en el contexto del conocimiento de las dimensiones de la inteligencia humana. Otros autores que contribuyeron al AF fueron Pearson (1901) y Hotelling (1933) y Thurstone (1947).

Las principales características de esta técnica son tres:

1. Estudiar la estructura de correlaciones entre variables.
2. Buscar factores latentes o hipotéticos que expliquen las variables originales. En nuestro caso tenemos cinco factores latentes: Socio-económico, Salud, Vecindad, Problemas Financieros y Condiciones. Por otro lado, tenemos 20 variables originales.
3. Representar correlaciones entre variables y entre variables y factores.

A continuación presentamos el método de una forma más profunda. Sean X_1, X_2, \dots, X_{20} las 20 variables objeto de estudio tipificadas.

Tabla 1. Esquema de la matriz de datos

Individuos	Variables			
	X_1	X_2	...	X_{20}
1	X_{11}	X_{12}	...	X_{120}
2	X_{21}	X_{22}	...	X_{220}
...
n (36.865)	X_{n1}	X_{n2}	...	X_{n20}

El modelo viene dado por las ecuaciones:

$$X_1 = a_{11}F_1 + a_{12}F_2 + \dots + a_{1k}F_k + u_1$$

$$X_2 = a_{21}F_1 + a_{22}F_2 + \dots + a_{2k}F_k + u_2$$

... ..

$$X_{20} = a_{201}F_1 + a_{202}F_2 + \dots + a_{20k}F_k + u_{20}$$

Donde

F_1, \dots, F_k son los factores comunes con $k \ll p$.

u_1, \dots, u_{20} son los factores únicos o específicos.

$\{a_{ij}; i=1, \dots, p; j=1, \dots, k\}$ son las cargas factoriales.

Para la extracción de los factores se usará el método de las componentes principales con rotación varimax para poder comparar los resultados con el documento base.

“El método [de las componentes principales] consiste en estimar las puntuaciones factoriales mediante las puntuaciones tipificadas de las k primeras componentes principales y la matriz de cargas factoriales mediante las correlaciones de las variables originales con dichas componentes. Este método tiene la ventaja de que siempre proporciona una solución. Tiene el inconveniente, sin embargo, de que al no estar basado en el modelo de Análisis Factorial puede llevar a estimadores muy sesgados de la matriz de cargas factoriales, particularmente, si existen variables con comunalidades bajas.” (Salvador Figueras, M., Gargallo Valero, P., 2006).

La rotación varimax minimiza el número de variables con cargas altas en un factor para conseguir así una mejor interpretación de los factores. Esta rotación consigue obtener cargas más extremas (± 1) y otras cercanas al 0.

En principio, es necesario realizar la prueba de esfericidad de Bartlett que contrasta si la matriz de correlaciones es una matriz identidad, lo cual indicaría que el modelo factorial es inadecuado. El estadístico de Bartlett se obtiene a partir de una transformación χ^2 del determinante de la matriz de correlaciones y cuanto mayor sea, y por tanto menor el nivel de significación, más improbable es que la matriz sea una matriz identidad y más adecuado resulta el análisis factorial.

Por otro lado, La medida de la adecuación muestral de Kaiser-Meyer-Olkin (Coeficiente KMO) contrasta si las correlaciones parciales entre las variables son pequeñas, toma valores entre 0 y 1, e indica que el análisis factorial es tanto más adecuado cuanto mayor sea su valor. Así, Kaiser propuso en 1974 el siguiente criterio para decidir sobre la adecuación del análisis factorial de un conjunto de datos:

- $0,9 < KMO \leq 1,0$ = Excelente adecuación muestral.
- $0,8 < KMO \leq 0,9$ = Buena adecuación muestral.
- $0,7 < KMO \leq 0,8$ = Aceptable adecuación muestral.
- $0,6 < KMO \leq 0,7$ = Regular adecuación muestral.
- $0,5 < KMO \leq 0,6$ = Mala adecuación muestral.
- $0,0 < KMO \leq 0,5$ = Adecuación muestral inaceptable.

El software usado fue el SPSS en su versión 21.

4.2. Modelo Lineal General

En el caso del análisis de varianza en donde se estudia el efecto de unas variables clasificadas sobre una variable independiente, si los niveles de la variable clasificadora son cuantitativos (por ejemplo, la edad), puede pensarse en un enfoque de regresión donde los valores de la variable regresora coincidirán con los valores correspondientes a dichos niveles. Por otro lado, si los niveles del factor son categorías (por ejemplo, distintas clases sociales), se puede sustituir la variable clasificadora por tantas variables regresoras como niveles tenga y codificar cada observación por la pertenencia (un 1) o no (un 0) a un nivel determinado. Así, si tenemos tres clases sociales ($V_1 =$ alta, $V_2 =$ media y $V_3 =$ baja) y dos individuos por clase en un diseño simple jerárquico, el modelo que representa este diseño es $Y_{ij} = \mu + V_i + \varepsilon_{ij}$, en donde Y_{ij} es la observación de la variable independiente correspondiente al j -ésimo individuo de la i -ésima clase social. Particularizando el modelo para cada observación se puede escribir:

$$\begin{aligned} Y_{11} &= 1\mu + 1V_1 + 0V_2 + 0V_3 + \varepsilon_{11} \\ Y_{12} &= 1\mu + 1V_1 + 0V_2 + 0V_3 + \varepsilon_{12} \\ Y_{21} &= 1\mu + 0V_1 + 1V_2 + 0V_3 + \varepsilon_{21} \\ &\dots\dots\dots \\ &\dots\dots\dots \\ Y_{32} &= 1\mu + 0V_1 + 0V_2 + 1V_3 + \varepsilon_{32} \end{aligned}$$

Este conjunto de ecuaciones se puede representar en forma matricial como

$$Y = X\beta + \varepsilon$$

siendo

$$Y = \begin{pmatrix} Y_{11} \\ Y_{12} \\ Y_{21} \\ \cdot \\ \cdot \\ Y_{32} \end{pmatrix} \quad X = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & 0 & 0 & 1 \end{pmatrix} \quad \beta = \begin{pmatrix} \mu \\ V_1 \\ V_2 \\ V_3 \end{pmatrix} \quad \varepsilon = \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{21} \\ \cdot \\ \cdot \\ \varepsilon_{32} \end{pmatrix}$$

Por tanto, este problema de Análisis de Varianza se puede tratar como un modelo clásico de regresión múltiple pudiendo calcular las sumas de cuadrados del

modelo completo (μ y V_i) y del reducido (solo con μ). A partir de ellas, por diferencia, las sumas de cuadrados atribuible al factor clase social. El único (e importante) problema es que la matriz \mathbf{X} produce una matriz $\mathbf{X}'\mathbf{X}$ que es singular (la primera columna de \mathbf{X} es la suma de las demás) por lo que es necesario recurrir a métodos alternativos. Uno, tradicionalmente empleado en Análisis de Varianza, es definir los efectos V_i de forma que sumen cero. En este caso se codifica V_3 como $-1V_1 -1V_2$ y la matriz \mathbf{X} tendrá solo tres columnas en vez de cuatro. En este caso, al particularizar para cada observación, el modelo es:

$$\begin{aligned}
 Y_{11} &= 1\mu + 1V_1 + 0V_2 + \varepsilon_{11} \\
 Y_{12} &= 1\mu + 1V_1 + 0V_2 + \varepsilon_{12} \\
 Y_{21} &= 1\mu + 0V_1 + 1V_2 + \varepsilon_{21} \\
 &\dots\dots\dots \\
 &\dots\dots\dots \\
 Y_{32} &= 1\mu - 1V_1 - 1V_2 + \varepsilon_{32}
 \end{aligned}$$

Con esta recodificación, los vectores y matrices son:

$$\mathbf{Y} = \begin{pmatrix} Y_{11} \\ Y_{12} \\ Y_{21} \\ \cdot \\ \cdot \\ Y_{32} \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & -1 & -1 \end{pmatrix} \quad \beta = \begin{pmatrix} \mu \\ V_1 \\ V_2 \end{pmatrix} \quad \varepsilon = \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{21} \\ \cdot \\ \cdot \\ \varepsilon_{32} \end{pmatrix}$$

lo que produce que la matriz $\mathbf{X}'\mathbf{X}$ sea ahora de rango completo y, por tanto, pueda calcularse su inversa.

Otros enfoques, en vez de recodificar para que la suma de los efectos sea cero, recodifican a base de que uno de los niveles (generalmente el último) valga cero.

En este caso, los vectores y matrices son:

$$\mathbf{Y} = \begin{pmatrix} Y_{11} \\ Y_{12} \\ Y_{21} \\ Y_{22} \\ Y_{31} \\ Y_{32} \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \beta = \begin{pmatrix} \mu \\ V_1 \\ V_2 \\ (V_3 = 0) \end{pmatrix} \quad \varepsilon = \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{31} \\ \varepsilon_{32} \end{pmatrix}$$

A partir de ahí, el proceso es similar al análisis de regresión. Por tanto, ambos tipos de análisis son realmente casos particulares de un enfoque más amplio que se llama Modelo Lineal General. En este modelo se asume que el vector de los errores ε se distribuye normalmente con vector de medias cero y matriz de varianzas covarianzas, $V = I\sigma^2$ en donde I es la matriz identidad (para incluir las suposiciones de independencia y homogeneidad de varianzas). Si se relajan las suposiciones de independencia y homogeneidad, el modelo todavía es válido si se supone que la matriz V tiene la forma:

$$V = \begin{pmatrix} \sigma_1^2 & S_{12} & \dots & S_{1t} \\ S_{12} & \sigma_2^2 & & \\ \dots & & \dots & \\ S_{1t} & S_{2t} & & \sigma_t^2 \end{pmatrix}$$

en donde σ_1^2 es la varianza del error experimental dentro del primer nivel del tratamiento y $\sigma_{1,2}$ es la covarianza entre los errores de los dos primeros niveles. Por tanto, este es un modelo lineal general en donde no se incluyen como suposiciones ni la homogeneidad de varianzas ni la independencia. A partir de él, aplicando estimación por mínimos cuadrados se puede obtener la descomposición de la variación y la estimación de los efectos.

Habiendo explicado un modelo sencillo con un solo factor, se procederá con un modelo más complejos en el que se estudiará la influencia de diversos factores demográficos sobre los indicadores obtenidos en el análisis factorial.

El criterio para decidir el factores fue mediante la consulta a numerosos sociólogos, expertos en el campo de la precariedad y se concluyó que los factores que deben incluirse son:

1. **Sexo:** recogida en la variable RB090, con valores de varón o mujer.
2. **Grupos Edad:** las edades anuales se han calculado a partir de la fecha de nacimiento de cada entrevistado (RB070 y RB080) y el año de referencia de la encuesta. Luego se han agrupado en grandes grupos: menos de 16 años, de 16 a 24 años, de 25 a 34 años, de 35 a 49 años, de 50 a 64 años y de 65 años y más. Hay que tener en cuenta que los valores superiores a 80 años han sido homogeneizados a 80 por el INE en la misma base de datos.

3. **Grupos de hogar:** en ella se recoge la composición del hogar en función de la edad. A partir de la variable HX060, hemos generado una agregación de categorías de manera que los hogares han quedado clasificados como sigue:

- i. Unipersonal con 64 años o menos: recoge las categorías «Una persona: hombre de menos de 30 años», «Una persona: hombre de entre 30 y 64 años», «Una persona: mujer de menos de 30 años» y «Una persona: mujer de entre 30 y 64 años».
- ii. Unipersonal 65 años o más: recoge las categorías «Una persona: hombre de 65 o más años» y «Una persona: mujer de 65 o más años».
- iii. Dos adultos sin niños: recoge las categorías «2 adultos sin niños dependientes económicamente, al menos una persona de 65 o más años» y «2 adultos sin niños dependientes económicamente, teniendo ambos menos de 65 años».
- iv. Tres o más adultos sin niños: recoge la categoría «Otros hogares sin niños dependientes económicamente».
- v. Un adulto con niños: recoge la categoría «Un adulto con al menos un niño dependiente».
- vi. Dos adultos con niños: recoge las categorías «Dos adultos con un niño dependiente», «Dos adultos con dos niños dependientes» y «Dos adultos con tres o más niños dependientes».
- vii. Tres adultos o más con niños: recoge la categoría «Otros hogares con niños dependientes».

En esta variable se consideran «niños dependientes» todos los menores de 16 años y los que tienen 16 y más años pero menos de 25 y son económicamente inactivos.

4. **Situación Profesional:** Esta variable se refiere a la situación profesional del individuo y contiene cuatro categorías diferentes

- i. Autónomo con empleados.
- ii. Autónomo sin empleados.
- iii. Empleado.
- iv. Empleado del hogar.

5. **Comunidad Autónoma:** la distribución territorial según la comunidad autónoma de residencia puede tener relevancia en la distribución de precariedades de la población española. Para recoger esta información hemos hecho uso de la variable DB040, en la que se recogen las regiones europeas (NUTS-II).”

Además se discutió cuáles eran las interacciones más relevantes y, finalmente, fueron decididas las siguientes: Sexo con Edad, Sexo con Tipo de Hogar, Sexo con Situación Profesional, Sexo con Comunidades Autónomas, Edad con Situación Profesional y Tipo de Hogar con Situación Profesional.

El modelo propuesto será de la forma:

$$Y_{ijklmn} = \mu + S_i + E_j + H_k + P_l + C_m + SE_{ij} + SH_{ik} + SP_{il} + SC_{im} + EP_{jl} + HP_{kl} + \varepsilon_{ijklmn}$$

Donde

$S_i =$ Efecto i – ésimo de la variable Sexo

$E_j =$ Efecto j – ésimo de la variable Edad

$H_k =$ Efecto k – ésimo de la variable Tipo de Hogar

$P_l =$ Efecto l – ésimo de la variable Situación Profesional

$C_m =$ Efecto m – ésimo de la variable Comunidad Autónoma

El programa usado para la aplicación de esta técnica fue el Infostat.

4.3. HJ-Biplot

El segundo objetivo propuesto es analizar las componentes de precariedad preponderantes en las diferentes Comunidades Autónomas del Estado Español y, por ello, optamos por la técnica HJ-Biplot. Tenemos pues las comunidades autónomas como filas y los índices calculados a partir del análisis factorial como columnas.

Gabriel (1971) fue el creador de los métodos Biplot con objetivo esencial de la representación gráfica en baja dimensión de una matriz de datos X (I filas y J columnas); es, por tanto, una técnica multivariante. La idea se sustenta mediante la descomposición en valores y vectores singulares de una matriz con la diferencia fundamental de reproducir el dato y la representación conjunta de filas y columnas. La idea de Biplot debe ser vista como un plot – Bi. Plot entendido como gráfico y Bi como representación simultánea de 2 objetos (filas – columnas) y es, por tanto, un método multivariante gráfico. Las dos factorizaciones más importantes propuestas fueron el

GH-Biplot y el JK-Biplot. Ambas tienen ciertas ventajas y desventajas. En el caso del GH-Biplot se consigue una alta calidad en la representación de las columnas mientras que en el caso de JK-Biplot se logra una alta calidad en la representación de filas con posibilidad de representar el elemento original X_{ij} .

Galindo (1985,1986) basándose en la teoría del AFC consigue una mejora del Biplot al permitir la representación con máxima calidad tanto en columnas como en filas aunque, desafortunadamente, sin posibilidad de representación del elemento original X_{ij} . Esta técnica multivariante fue bautizada como: HJ-Biplot. En este Biplot los marcadores para las columnas coinciden con los marcadores columnas del GH-Biplot y los marcadores para las filas coinciden con los del JK-Biplot.

Supongamos una matriz $X_{n \times p}$ con filas n centradas y columnas p centradas .

Tomamos $U\Sigma$ como marcadores para las filas y ΣV como marcadores para las columnas y teniendo en cuenta las relaciones entre U y V podemos escribir:

$$U\Sigma = XV\Sigma^{-1}\Sigma = XV$$

Entonces $U\Sigma$ coincide con la proyección de los puntos fila que tiene mejor ajuste en sentido de mínimos cuadrados.

Análogamente:

$$\Sigma V = \Sigma\Sigma^{-1}X^T U = X^T U$$

En este caso ΣV coincide con la proyección de los puntos columna sobre el espacio de máxima inercia.

Podemos relacionar ambas relaciones si tenemos en cuenta que:

$$X^T U = X^T XV\Sigma^{-1} \begin{cases} B = X^T A\Sigma^{-1} \\ A = XB\Sigma^{-1} \end{cases}$$

De aquí podemos expresar la h -ésima coordenada de la columna j -ésima en función de las h -ésimas coordenadas de las n filas.

$$b_{jh} = \{x_{1j}a_{1h} + \dots + x_{nj}a_{nh}\} \left\{ \frac{1}{\sqrt{\lambda_h}} \right\}$$

Análogamente:

$$a_{ih} = \{x_{1i}b_{1h} + \dots + x_{ip}a_{ph}\} \left\{ \frac{1}{\sqrt{\lambda_h}} \right\}$$

Esto nos permite representar las coordenadas de las filas y las coordenadas de las columnas en un mismo sistema de referencia.

El software usado para este apartado será el MultiBiplot de Vicente-Villardón (2010) (<http://biplot.usal.es/ClassicalBiplot/index.html>)

4.4. Stasis/Stasis Dual

STATIS es un acrónimo de una expresión francesa '*Structuration de Tableaux à Trois Indices de la Statistique*' (que podría traducirse como, estructurando tablas estadísticas de tres vías). Ésta técnica también se conoce bajo el acrónimo 'ACT' que significa en francés '*Analyse Conjointe de Tableaux*' (análisis conjunto de tablas). Bajo estos oscuros acrónimos esta la generalización del análisis de componentes principales cuyo objetivo es analizar varias matrices de datos de variables cogidas en el mismo conjunto de observaciones o- como en su versión dual llamada STATIS DUAL- varios conjuntos de observaciones medidas con el mismo set de variables, el cual es nuestro caso. Por lo tanto, STATIS es parte de la multitabla (también llamado multibloque o análisis de consenso) de la familia ACP en donde hay técnicas similares como análisis factorial múltiple (AFC), o análisis multibloque, discriminante de correspondencias (MUDICA, en inglés).

STATIS se originó en la escuela francesa por los trabajos de Escouffier (1980) y fue descrito por primera vez por L'Hermier des Plantes (1976) que luego fue desarrollado por Lavit (1988). Un acercamiento similar considerado en la comunidad de habla inglesa es el trabajo con título: *Procrustes matching by congruence coefficients* basado en los trabajos de Gower (1971).

Esta técnica estadística de carácter multivariante tiene tres objetivos. En primer lugar, comparar y analizar la relación entre diferentes conjuntos de datos, por otro lado, integrar estos conjuntos en uno llamado *compromiso* o *consenso* que es luego analizado vía ACP para mostrar la estructura común entre observaciones. Finalmente, proyectar cada conjunto de datos original en el espacio del compromiso para observar las discrepancias y comunalidades. A continuación se explican los dos tipos de metodología STATIS y descritos gráficamente en las figuras 2 y 3:

1. STATIS: Es un método para analizar múltiples conjuntos de variables diferentes medidos con mismas observaciones y ha sido aplicado en

varios campos de conocimiento como en biología molecular, ecología, ciencias de la información, neuroimagen y otros campos. Sin embargo jamás se ha usado para el estudio en el campo de la sociología y, menos aún, en el estudio de la precariedad. Esta técnica puede ser innovadora en este campo dado la cantidad de datos en tres vías que se disponen. Es por esto por lo que con este trabajo se abre un nuevo frente de trabajo y técnicas más allá de los clásicos modelos lineales con medidas repetidas.

2. STATIS DUAL: En nuestro caso se usa una variante del STATIS debido a la estructura de nuestro trabajo llamado STATIS DUAL. En el STATIS DUAL, los datos consisten en K conjuntos de observaciones diferentes medidas con el mismo conjunto de variables. Aquí, en vez de computar K matrices producto-cruzadas entre observaciones, se computa K matrices de covarianzas entre las variables (una por cada conjunto de observaciones). El STATIS-DUAL usa los mismos pasos que el STATIS y proporcionara un mapa compromiso de las variables (en vez de observaciones) y las cargas parciales de cada tabla.

Figura 2 Statis Dual

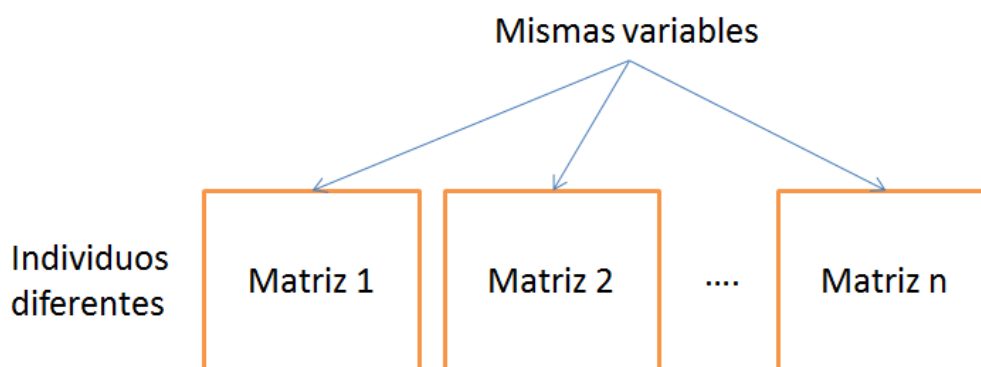
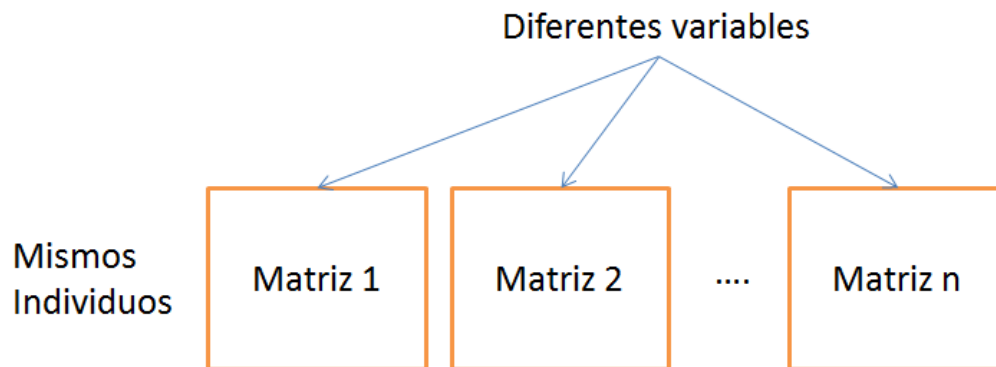
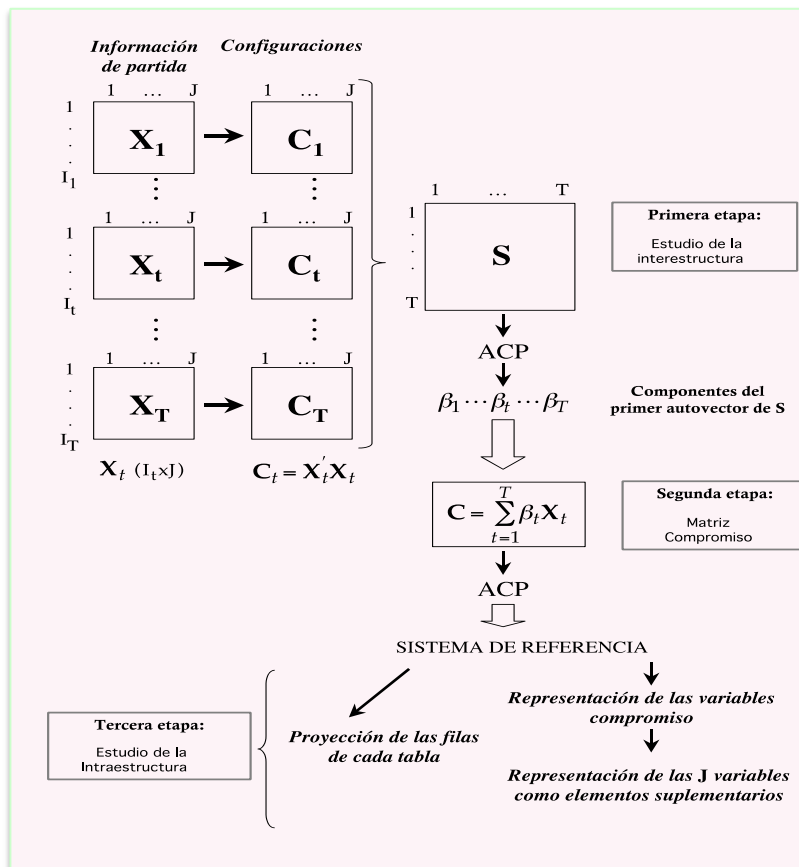


Figura 3. Statís



La idea general detrás de STATIS es el analizar la estructura de cada conjunto de datos y derivar de ésta una estructura óptima de los pesos para computar la mejor representación común llamada compromiso. Para computar este compromiso, los elementos de cada tabla son multiplicados por un peso óptimo de ésta y el compromiso es obtenido por la adición de estos “pesos” de las K tablas (la tabla compromiso es una combinación lineal del resto tablas). Estos pesos se eligen de forma que la estructura compromiso tiene la mejor representación (en sentido de mínimos cuadrados) del conjunto de tablas analizadas. El ACP del compromiso descompone la varianza en un conjunto de nuevas variables ortogonales llamadas, ejes, factores, o incluso variables latentes ordenadas por la cantidad de varianza que cada componente explica. Las coordenadas de las variables se llaman *puntuaciones* y pueden ser dibujadas en un mapa. La posición de estas de cada tabla de datos puede ser representada como puntos en el compromiso. Como las componentes obtenidas por combinación de las variables originales cada variable componente cierta cantidad a cada componente. Esta cantidad se llama carga de variable en componente y refleja la importancia de esa variable para el componente y también es posible hacer proyecciones de variables que reflejan la asociación. Finalmente, las tablas pueden ser representadas como puntos en un espacio multidimensional. La figura 4 muestra un esquema del algoritmo que usa el STATIS DUAL.

Figura 4. Esquema de las etapas del Statist Dual.



Tal y como muestra la figura 4, la estructura factorial en cada tabla X_t puede ser capturada por la covariación/matriz de correlaciones entre variables (es entre variables debido a tratarse de un STATIS DUAL y no de un STATIS clásico), tal y como se muestra a continuación:

$$C_t = X_t'X_t$$

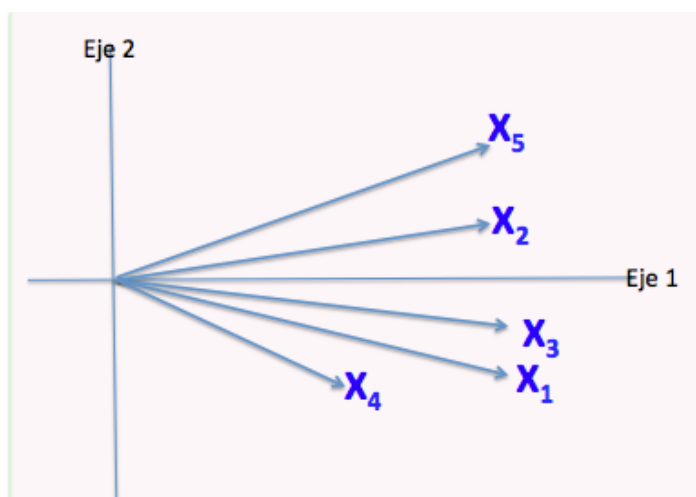
Para comprara las configuración de C_t se usa un coeficiente de correlación entre matrices y, para ellos, es necesario la definición de un producto interno. En este caso el producto interno de *Hilber-Schmidt*:

$$\langle C_t | C_{t'} \rangle_{\text{Hilber-Schmidt}} = \text{tr}(C_t C_{t'})$$

Una vez obtenido esta matriz de correlaciones se obtienen los valores y vectores propios de esta matriz mediante un ACP y se reduce así la dimensionalidad produciendo una imagen euclídea, donde cada matriz de partida vendrá representada como un punto en un espacio euclídeo de baja dimensión.

En el espacio estructurado por la distancia de *Hilber-Schmidt* cada matriz de datos viene representada por un punto y, conectando cada punto con el origen de coordenadas, podemos obtener una estimación grafica de coeficiente de correlación vectorial entre pares de matriz de la siguiente forma:

Figura 5. Representación hipotética de 5 matrices de datos, tras aplicar STATIS



En este caso, el ángulo formado con el eje 1 nos permite estimar el peso que tendrá esa matriz con la matriz consenso. Si tenemos ángulos agudos como el caso de la matriz ejemplo X_2 o X_3 , tendremos un peso importante en la matriz de compromiso. Sin embargo, podemos observar en la matriz X_4 tiene poco peso o, peso residual, en la matriz de compromiso además de tener poca información dada la escasa longitud del vector.

La matriz consenso se crea como una media ponderada de las matrices de partida donde los pesos son los factores de carga del primer vector propio de la matriz de correlaciones vectoriales S de la forma:

$$C = \sum_{t=1}^T \beta_t X_t$$

A partir de la diagonalización de la matriz compromiso se obtienen las coordenadas de las variables compromiso.

Una vez construida la matriz consenso, o matriz compromiso, es posible representar las variables de las matrices originales como elementos suplementarios y de esta forma analizar las trayectorias de las variables en los diferentes tiempos o situaciones.

Ver detalle de las etapas en el esquema de la figura 4.

En nuestro caso cada matriz X contendrá información sobre los individuos de los 5 países en estudio Tendremos pues 5 matrices (X_1, X_2, X_3, X_4 y X_5). Trataremos de ver si las estructuras de covariación de las diferentes dimensiones de la precariedad son similares en los distintos países, o sea, en las 5 matrices, y en caso de que así sea buscaremos la estructura consenso para los países en estudio.

El software usado para este apartado será la versión autoejecutable del ADE-4.

5. Resultados

5.1. Descriptivos

Tabla 2. Descriptivos de cada variable por países

Dimension	Variable	Categorías	Bélgica	España	Francia	Italia	Portugal
Problemas financieros	HS150 + HS150_F	No hay deudas	70,2%	70,4%	58,2%	79,8%	79,4%
		El pago no es una carga	10,2%	1,2%	22,6%	,6%	2,9%
		El pago tiene algo de carga	10,3%	11,4%	8,6%	9,7%	10,5%
		El pago es una gran carga	9,4%	17,0%	10,5%	9,9%	7,1%
	HS140	No hay carga	36,4%	3,7%	48,5%	1,4%	18,5%
		Poca carga	33,7%	45,1%	24,1%	45,1%	49,7%
		Gran Carga	29,9%	51,2%	27,5%	53,5%	31,9%
	HS120	Muy facilmente	5,4%	1,2%	1,9%	,8%	,3%
		Facilmente	26,3%	12,9%	10,7%	6,4%	4,0%
		Medianamente fácil	26,4%	26,3%	26,6%	20,8%	14,8%
Con cierta dificultad		20,9%	28,1%	42,5%	38,2%	33,8%	
Con dificultad		13,2%	17,1%	14,4%	19,0%	23,6%	
	Con gran dificultad	7,9%	14,5%	3,9%	14,8%	23,5%	
Condiciones	HS160	Suficiente luz	90,6%	93,5%	92,7%	92,6%	91,0%
		Demasiado oscuro	9,4%	6,5%	7,3%	7,4%	9,0%
	HH050	Suficiente calor	94,7%	93,5%	94,6%	90,8%	69,0%
		No hay suficiente calor	5,3%	6,5%	5,4%	9,2%	31,0%
	HH091	Si hay baño	99,3%	100,0%	98,6%	99,6%	97,5%
		Si hay baño pero es compartido	0,0%	0,0%	0,8%	0,1%	0,2%
		No hay baño	0,7%	0,0%	0,6%	0,2%	2,3%
	HH081	Si hay ducha	99,3%	100,0%	98,7%	99,5%	97,2%
		Si hay ducha pero es compartida	0,0%	0,0%	0,8%	0,1%	0,3%
		No hay ducha	0,7%	0,0%	0,5%	0,4%	2,6%
HH040	Goteras	85,3%	81,5%	87,5%	79,3%	81,2%	
	No tengo goteras	14,7%	18,5%	12,5%	20,7%	18,8%	
Vecindad	HS170	No tengo este problema	80,3%	78,9%	81,8%	75,8%	78,4%
		Tengo este problema	19,7%	21,1%	18,2%	24,2%	21,6%
		No tengo este problema	84,5%	86,9%	87,2%	81,4%	82,7%
		Tengo este problema	15,5%	13,1%	12,8%	18,6%	17,3%
	HS180	No tengo este problema	84,5%	86,9%	87,2%	81,4%	82,7%
		Tengo este problema	15,5%	13,1%	12,8%	18,6%	17,3%
	HS190	No tengo este problema	82,4%	84,8%	85,3%	86,6%	87,6%
		Tengo este problema	17,6%	15,2%	14,7%	13,4%	12,4%
	HD040	Raramente o nunca	66,2%	73,9%	91,5%	64,8%	58,3%
		A veces	18,1%	18,6%	2,3%	21,9%	27,2%
Frecuentemente		9,8%	5,2%	2,9%	8,3%	8,7%	
	Muy Frecuentemente	5,9%	2,3%	3,3%	5,1%	5,7%	
HD050	Raramente o nunca	76,8%	78,2%	82,8%	68,7%	71,2%	
	A veces	14,1%	13,8%	10,7%	21,4%	20,5%	
	Frecuentemente	6,0%	5,7%	4,5%	6,7%	5,6%	
	Muy Frecuentemente	3,1%	2,2%	2,0%	3,3%	2,7%	
Salud	PH010	Muy bien	28,4%	14,9%	25,5%	13,0%	7,8%
		Bien	45,8%	54,4%	44,0%	50,7%	35,3%
		Ni bien ni mal	17,5%	21,4%	21,9%	25,8%	34,8%
		Mal	6,8%	7,0%	7,6%	8,2%	16,2%
		Muy mal	1,6%	2,3%	1,1%	2,3%	5,9%
	PH020	No	75,7%	67,3%	63,6%	77,6%	62,9%
		Si	24,3%	32,7%	36,4%	22,4%	37,1%
	PH030	Si, muy limitado	77,8%	74,1%	76,5%	73,0%	65,3%
		Si, limitado	15,2%	20,0%	14,8%	19,0%	22,5%
		No, no limitado	7,0%	5,9%	8,8%	8,0%	12,3%
PH040 + PH050	Cumplido	99,1%	91,9%	95,0%	93,1%	95,4%	
	El motivo es estar en lista de espera	0,0%	,2%	,3%	1,4%	,5%	
	Miedo a examinacion médica	,3%	5,6%	2,2%	1,3%	,6%	
	No conocia ningun médico	,0%	,1%	,0%	,1%	,0%	
	Demasiado lejos para viajar o no podia pagar el viaje	0,0%	,1%	,0%	,1%	,1%	
	No podia faltar al trabajo	,1%	2,0%	1,0%	,6%	,3%	
	No tenia medios economicos	,5%	,2%	1,5%	3,5%	3,1%	
PH060+PH070	Cumplido	97,5%	90,9%	92,1%	89,8%	86,5%	
	El motivo es estar en lista de espera	0,0%	,0%	,2%	0,6%	,3%	
	Miedo a examinacion médica	,8%	4,0%	2,5%	2,4%	1,7%	
	No conocia ningun médico	,1%	,0%	,1%	,0%	,0%	
	Demasiado lejos para viajar o no podia pagar el viaje	0,0%	,0%	,1%	,1%	,1%	
	No podia faltar al trabajo	,4%	0,8%	,8%	,8%	,6%	
	No tenia medios economicos	1,2%	4,2%	4,2%	6,2%	10,8%	
Socio-Económico	Labour_stat_full\$	Nivel de precariedad 0	20,9%	16,4%	21,4%	15,8%	14,7%
		Nivel de precariedad 1	23,2%	21,4%	25,1%	20,8%	24,6%
		Nivel de precariedad 2	6,6%	12,9%	7,3%	11,6%	11,5%
		Nivel de precariedad 3	8,4%	1,9%	5,0%	2,4%	0,8%
		Nivel de precariedad 4	1,8%	2,1%	2,4%	1,8%	2,0%
		Nivel de precariedad 5	17,9%	14,2%	20,9%	19,7%	25,1%
		Nivel de precariedad 6	6,6%	6,0%	6,7%	5,9%	5,7%
		Nivel de precariedad 7	5,0%	8,9%	4,5%	5,0%	5,9%
		Nivel de precariedad 8	5,8%	12,5%	3,0%	10,7%	6,1%
		Nivel de precariedad 9	2,6%	2,1%	2,4%	1,1%	1,4%
	Nivel de precariedad 10	1,1%	1,5%	1,2%	5,1%	2,3%	
	HHIncome_CLASS15	Nivel de precariedad 0	,3%	1,6%	,1%	,5%	0,0%
		Nivel de precariedad 1	,2%	1,6%	,1%	1,0%	2,1%
		Nivel de precariedad 2	,5%	3,5%	,5%	2,1%	12,2%
		Nivel de precariedad 3	2,4%	6,7%	1,5%	4,0%	14,8%
		Nivel de precariedad 4	6,5%	9,0%	3,6%	6,8%	17,9%
		Nivel de precariedad 5	7,3%	10,7%	5,8%	8,9%	14,3%
		Nivel de precariedad 6	7,3%	10,4%	7,2%	9,4%	11,1%
		Nivel de precariedad 7	9,3%	11,6%	9,4%	11,1%	8,5%
		Nivel de precariedad 8	6,8%	8,6%	9,0%	8,4%	5,1%
Nivel de precariedad 9		7,1%	7,4%	9,2%	7,6%	3,3%	
Nivel de precariedad 10	7,8%	5,8%	8,4%	7,0%	2,8%		
Nivel de precariedad 11	16,7%	10,7%	17,1%	13,3%	3,6%		
Nivel de precariedad 12	11,3%	6,1%	10,7%	7,9%	1,7%		
Nivel de precariedad 13	9,1%	3,7%	7,8%	6,2%	1,4%		
Nivel de precariedad 14	4,3%	1,6%	3,9%	2,7%	,5%		
Nivel de precariedad 15	2,8%	1,0%	5,7%	3,2%	,7%		

La tabla muestra la distribución en porcentajes de las categorías de todas las variables usadas en el estudio, así como los cinco países estudiados. La tabla 2 resalta, en rojo, la información más destacable.

En el caso de la carga financiera debido al pago de compras a plazos o préstamos (variable HS150 + HS150_F) observamos que todos los países tienen una distribución de frecuencias muy similar excepto en el caso de Francia pues son los que más deudas tienen dado que un 58,2% de ellos no tiene deuda mientras que el resto de los países rondan el 70%. Sin embargo, el 22,6% de ellos no les supone carga alguna.

En la variable sobre la carga hipotecaria (HS140) se observan grandes diferencias entre un país y otro, teniendo a España en un extremo. En este país un 48,5% de los encuestados contestan que no tienen carga mientras que en el extremo opuesto, Italia, sólo un 1,4% no tienen carga. Además, la hipoteca supone una gran carga para el 53,5% de los italianos.

En la dimensión de Condiciones destacamos la poca variabilidad de los datos, pues una categoría acapara gran parte de la distribución de frecuencias.

En la dimensión de Vecindad, describimos la variable HD040 y la variable HD050. En el primer caso, tenemos dos extremos en la distribución de frecuencias. Por un lado Francia, en el que el 91,5% afirma que su vecindario no está sucio y, en el otro, el 58,3% de los portugueses afirma que su vecindario no está sucio.

Italia y Portugal son los que menos afirman que su vecindario no tenga los servicios públicos (variable HD050) en mal estado con un 68,7% y un 71,2% respectivamente. Ciertamente es que el 21,4% y el 20,5% de los encuestados afirman que los servicios públicos están en mal estado solo "a veces".

En el área de salud destacamos la poca variabilidad de las preguntas relacionadas con los motivos por la no asistencia al médico o dentista (variable PH040+PH050 y PH060+PH070). Bélgica es el país en el que menos problemas para la asistencia médica o dentista han tenido pues el 99,1% de ellos han visto sus necesidades cubiertas (variable PH040+PH050) y lo mismo sucede con su asistencia al dentista (variable PH060+PH070) con un 97,5%.

En la variable labour_stat_full\$ todos los países tienen patrones similares y alrededor del 30% de ellos no tienen precariedad laboral.

Finalmente, la variable hhincome_CLASS15 tiene una distribución muy similar para todos los países menos en el caso de Portugal que, como se señala en rojo, es el país con menos precariedad de ingresos.

5.2. Análisis de la dimensión de la precariedad

Los autores antes citados, que habían abordado el análisis previo de estos datos, asumen que la estructura latente hipotetizada se encuentra en nuestros datos y por tanto realizan un análisis de cada una de esas dimensiones sometiendo los ítems de la encuesta que teóricamente conforman cada dimensión a un Análisis Factorial Exploratorio (solución en componentes principales y rotación varimax) y probando que presentan una estructura unidimensional.

Nuestro planteamiento es diferente. Queremos estudiar toda la encuesta globalmente y explorar si esas dimensiones latentes existen realmente. El método estadístico utilizado es el mismo, un análisis factorial exploratorio (solución en componentes principales y rotación varimax).

La prueba de Bartlett es estadísticamente significativa ($<0,000$) y la medida de adecuación KMO nos da un valor de 0,713 que, como se explicó anteriormente, existe una “aceptable adecuación muestral.”

La siguiente tabla muestra la absorción de inercia.

Tabla 3. Absorción de inercia

Componente	Valor propio	% de la varianza	% acumulado
1	2,273	12,628	12,628
2	1,765	9,804	22,432
3	1,627	9,041	31,473
4	1,535	8,526	39,999
5	1,463	8,127	48,126
6	1,387	7,704	55,830
7	1,255	6,969	62,800

Vemos pues que se obtiene una varianza total explicada con 7 factores de 62,80%.

La tabla siguiente recoge las cargas factoriales y pone de manifiesto que existen siete dimensiones latentes, en su mayoría coincidentes con las supuestas por los autores del estudio base, que son las que aparecen en la figura 1, pero no todas las variables iniciales fueron usadas. El criterio para la eliminación de las variables HS160 y HS170 responde a que no cargaban en componentes sociológicamente interpretables e introducían demasiado ruido en la estructura factorial resultante.

Tabla 4. Cargas factoriales. Se han suprimido los valores menores a 0,30

<i>Variables</i>	<i>Salud1</i>	<i>Problemas Financieros</i>	<i>Vecindad1</i>	<i>Vecindad2</i>	<i>Salud2</i>	<i>Socio- Económico</i>	<i>Condiciones</i>
Labour_stat_full\$ Precariedad laboral Hhincome_CLASS15 Precariedad de ingresos						0,639 0,738	
HS160 Demasiado oscuro, poca luz HH050 Posibilidad de tener la casa caliente HH040 Existencia de goteras							0,697 0,461 0,707
HD040 Vecindario sucio HD050 Servicios públicos en mal estado				0,815 0,855			
HS150 + HS150_F Carga financiera debido a prestamos HS140 Pago hipoteca HS120 Llegar a fin de mes		0,605 0,78 0,798					
PH010 Autopercepción del estado general de salud PH020 Sufre de una enfermedad crónica PH030 Limitación en las actividades debido a tener problemas de salud	0,843 0,838 0,867						
PH040 + PH050 Motivos por los que no se ha visitado un dentista PH060+PH070 Motivos por los que no se ha visitado un médico					0,853 0,833		
HS170 Demasiado ruido HS180 Polución HS190 Crimen			0,786 0,722 0,54				

Las columnas muestran los nombres de las dimensiones latentes capturadas que se corresponden con los 7 ejes factoriales. Las dos variables construidas a partir de varias variables recogidas en la encuesta (ver apartado de encuestas para

consultar detalles) conforman una dimensión latente de la precariedad, bien diferenciada del resto. Cargan ambas en el eje 6, lo que indica que no es una dimensión con alto poder explicativo.

La mayor diferencia con respecto al documento base es que los ítems de salud conforman dos ejes a los que hemos llamado Salud 1 y Salud 2 y que se corresponden con información relativa a la salud general y, con las causas que han motivado la falta de salud, respectivamente.

Es importante resaltar que el primer eje es el que más información aporta, la precariedad de Salud1, formada por las variables: salud general, crónico y limitaciones en las actividades, es la dimensión de precariedad con mayor capacidad explicativa, mientras que la precariedad en condiciones aparece en el eje 7, lo que indica que es en la dimensión que existe menos variabilidad en el estudio. En el documento original la forma de aplicar el análisis factorial fue diferente pues fue aplicado a cada subdimensión (condiciones, vivienda, etc.) por separado. Esto conlleva el inconveniente de no saber la importancia (en términos de información/absorción de inercia) de los siete indicadores.

El indicador precariedad Salud1, está formado por variables que indican la precariedad general en salud. Por otro lado, la precariedad en Salud2, se refiere a los motivos por los que no se asistió al médico y está íntimamente relacionada con la precariedad socio-económica tal y como se puede observar en la tabla 4 ambas dimensiones están una al lado de la otra.

En definitiva, las tres dimensiones más importantes para la caracterización de la precariedad son: precariedad en salud general (Salud1), precariedad financiera (Problemas Financieros), y precariedad del vecindario1 (vecindario1). Tenemos pues que los tres pilares de la precariedad son: salud general, el tener deudas, y problemas en el vecindario (crimen, contaminación y ruido).

A continuación se propone probar que los indicadores son sociológicamente consistentes mediante la aplicación del modelo lineal general a las siete variables latentes resultado del análisis factorial.

5.3. Evaluación de la eficiencia de los indicadores

Esta técnica fue usada con objeto de probar que los indicadores son sociológicamente consistentes (y no únicamente estadísticamente consistentes) para ello podemos optar por diferentes métodos: descriptivo, anova/modelo lineal general. Nos hemos decantado por el modelo lineal general con múltiples factores por diversos

motivos. El análisis descriptivo, no permite hacer inferencia y por ello no es adecuado para probar la consistencia sociológica y, finalmente, se decide usar múltiples factores para evitar tener un error experimental tan alto como si solo se usase un factor. Además, es posible, que al usar un factor, se observen diferencias de medias cuando, en realidad, no las hay sino que existe interacción con otro factor. Recordemos el modelo:

$$Y_{ijklmn} = \mu + S_i + E_j + H_k + P_l + C_m + SE_{ij} + SH_{ik} + SP_{il} + SC_{im} + EP_{jl} + HP_{kl} + \varepsilon_{ijklmn}$$

Donde

$S_i =$ Efecto i – ésimo de la variable Sexo

$E_j =$ Efecto j – ésimo e la variable Edad

$H_k =$ Efecto k – ésimo de la variable Tipo de Hogar

$P_l =$ Efecto l – ésimo de la variable Situación Profesional

$C_m =$ Efecto m – ésimo de la variable Comunidad Autónoma

A continuación exponemos los resultados de los p valores y hacemos un breve comentario sobre los resultados obtenidos pues no es objetivo de este trabajo el análisis exhaustivo del modelo lineal. Debido a la cantidad de modelos realizados y que no son objeto específico de este estudio solo van a ser comentados aquellos resultados que sean relevante.

Tabla 5. P valores del modelo lineal general con las 7 dimensiones extraídas del análisis factorial

	Salud1	Problemas Financieros1	Vecindad1	Vecindad2	Salud2	Socio-Económico	Condiciones
Intercepto	0,000	0,770	0,410	0,150	0,010	0,000	0,770
Sexo	0,201	0,250	0,860	0,880	0,180	0,120	0,130
Edad (grupos)	0,000	0,000	0,010	0,270	0,010	0,000	0,010
Tipo de Hogar	0,413	0,000	0,020	0,030	0,010	0,000	0,220
Situación Profesional	0,004	0,030	0,240	0,040	0,230	0,000	0,330
CCAA	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Sexo * Edad (grupos)	0,000	0,050	0,250	0,830	0,870	0,000	0,320
Sexo * Tipo de Hogar	0,006	0,000	0,970	0,930	0,030	0,000	0,330
Sexo * Situación Profesional	0,183	0,710	0,490	0,580	0,440	0,000	0,000
Sexo * CCAA	0,091	0,870	0,930	1,000	0,000	0,720	1,000
Edad (grupos) * Situación Profesional	0,605	0,040	0,350	0,570	0,820	0,000	0,150
Tipo de Hogar * Situación Profesional	0,535	0,040	0,410	0,240	0,820	0,000	0,050

En verde se señalan aquellos resultados con un p valor menor de 0,05.

Lo primero a destacar es que, en todos los casos, los resultados de los contrastes de las comunidades autónomas no sale estadísticamente significativos pero hay que tener en cuenta que en el caso de salud2 existe una interacción sexo * comunidad autónoma por lo tanto no podemos analizar, en este caso, la comunidad autónoma sola. Según los resultados obtenidos las comunidades autónomas más precarias son Galicia, Extremadura, Ceuta, Melilla y Canarias, siendo esta ultima la que obtiene peores resultados en términos de precariedad.

Por otro lado, al analizar las interacciones con sexo vemos que la mujer suele estar peor situada tal y como muestra la teoría sociológica pero no podemos decir que la mujer este siempre peor que el hombre en todas las categorías debido a la existencia interacción.

En lo referente a grupos de edad se observa que los mayores de 65 son los que peores resultados (en términos de precariedad) obtienen en los indicadores de salud (salud1 y salud2) mientras que los jóvenes son los que peores resultados obtienen en socio-económico y problemas financieros, tal y como se esperaba y que contrasta con la teoría.

En definitiva concluimos que conseguimos unos indicadores sociológicamente consistentes con la teoría.

A continuación se propone la realización del HJ-Biplot para observar las relaciones entre las dimensiones latentes y las diferentes comunidades de España y ahondar más en los resultados obtenidos en este apartado.

5.4. Caracterización multivariante de la precariedad en Comunidades Autónomas

Otro de los objetivos de este trabajo era encontrar los perfiles de precariedad de las distintas Comunidades Autónomas españolas, una vez probado en el apartado anterior que existen diferencias estadísticamente significativas entre las Comunidades. Para ello partimos de una matriz de datos que contiene tantas filas como Comunidades Autónomas (19) y tantas columnas como dimensiones de precariedad consideradas (7). Cada dato representa la media de los individuos pertenecientes a dicha Comunidad. El valor para cada individuo es la puntuación sobre la dimensión latente correspondiente.

Recordamos que se obtuvieron, tras el Análisis Factorial, 7 dimensiones latentes, que recogen diferentes aspectos de la precariedad, concretamente: Salud1 (estado general de salud, problemas crónicos y limitaciones en la actividad diaria), Problemas Financieros, Vecindad1 (crimen, polución y ruido), Vecindad2 (suciedad en el barrio y mal estado de los servicios públicos), Salud2 (motivos por la no asistencia al médico o dentista), Socio-Económico, Condiciones (goteras, falta de calor en el hogar y demasiada oscuridad).

La técnica utilizada es el HJ-BIPLLOT (Galindo, 1986). Los análisis se realizaron con el programa informático MULTBIPLLOT (Vicente-Villardón ,2010).

Tal como señalamos en el apartado de metodología, el HJ-Biplot para la matriz de datos $X_{19 \times 7}$, se define como una representación gráfica multivariante mediante marcadores j_1, j_2, \dots, j_{19} para las filas y h_1, h_2, \dots, h_7 para las columnas de X , elegidos de forma que ambos marcadores puedan superponerse en el mismo sistema de referencia, con máxima calidad de representación. Los marcadores se eligen a partir de la descomposición en valores singulares de la matriz de media X .

Para diferenciarlos representaremos a los marcadores de los índices de precariedad con un vector y a los marcadores de las Comunidades Autónomas con un punto y la sigla identificativa.

La calidad de representación para las filas y para las columnas de la matriz de datos X es la misma y las filas y columnas vienen expresadas en coordenadas principales.

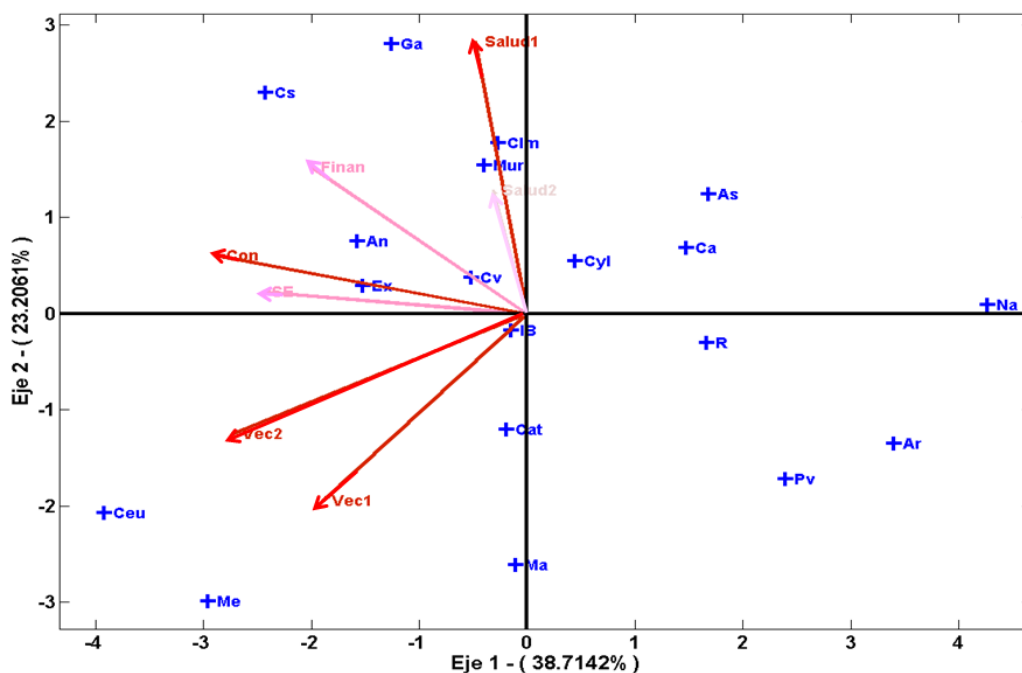
Esta representación multivariante nos permite interpretar las relaciones fila-columna a través de los ejes factoriales.

Debido a que tanto las filas como las columnas tienen la misma calidad de representación, se pueden interpretar las posiciones de las filas, de las columnas y las relaciones fila-columna a través de las contribuciones relativas del factor al elemento y del elemento al factor (Galindo y Cuadras, 1986).

La absorción de inercia en el plano 1-2 es del 62% y supera en el espacio tridimensional el 80%. Limitaremos la interpretación, pues, a los planos principales generados por los tres primeros vectores propios.

El gráfico siguiente contiene la representación HJ-Biplot en el primer plano principal.

Gráfico 1. Representación HJ-Biplot de plano principal (ejes 1-2).



Podemos ver que ya el primer plano factorial (eje 1 y eje 2) absorbe casi el 62% de varianza explicada, y con tres ejes se absorbe un 80% de varianza total. Es por ello por lo que continuamos con el análisis centrándonos en los tres primeros ejes. A continuación estudiamos las calidades de representación tanto de las comunidades autónomas como de los índices de precariedad.

Una absorción de inercia alta no garantiza calidad de representación alta para todas las Comunidades ni para todos los índices de precariedad. Analizaremos, por tanto, cuáles de las Comunidades están bien representadas y pueden ser

interpretadas sus posiciones en este plano, y cuáles no. Análogamente haremos el estudio para las variables (índices)

La tabla siguiente (Tabla 6) recoge la contribución relativa de cada uno de los tres primeros Factores a cada una de las Comunidades Autónomas. La suma de las contribuciones de cada eje que conforma un plano factorial nos permite estimar la calidad de representación de esa Comunidad Autónoma, en ese plano factorial. Por ejemplo, **Melilla** tiene una alta calidad de representación en el plano 1-2 (456 +464 = 920, sobre mil). **Navarra** también tiene una alta calidad de representación, 884 (884 +0=884). Análogo razonamiento para las demás.

Analizando la información del Eje 1 podemos afirmar que **Andalucía, Aragón, Cantabria, Ceuta, Navarra, La Rioja, el País Vasco**, son Comunidades de Eje 1.

Madrid, Castilla La Mancha y Galicia son de Eje 2.

Melilla es de plano (recibe contribuciones altas de Eje 1 y de Eje 2) y también **Asturias**, aunque las contribuciones son más bajas.

Cataluña, Comunidad Valenciana y Baleares son de eje 3.

Cataluña, Madrid y Galicia están mejor representadas en el plano 2-3.

Tabla 6. Contribuciones relativas de los Ejes Factoriales a las CCAA

CC AA	Eje 1	Eje 2	Eje 3
An (Andalucía)	464	107	65
Ar (Aragón)	685	108	121
Cs (Canarias)	322	289	163
Ca (Cantabria)	409	88	321
CyL (Castilla y León)	49	73	183
Clm (Castilla La Mancha)	12	540	75
Cat (Cataluña)	4	151	820
Ceu (Ceuta)	575	160	235
Me (Melilla)	456	464	13
Ma (Madrid)	1	579	273
Na (Navarra)	884	0	0
Cv (Comunidad Valenciana)	58	30	481
Ex (Extremadura)	343	12	194
Ga (Galicia)	94	467	241
IB (Baleares)	8	10	595
R (Rioja)	595	20	19
Pv (País Vasco)	574	297	58
As (Asturias)	382	209	3
Mur (Murcia)	15	223	0

De lo anterior se deduce que excepto **Cataluña, Comunidad Valenciana, Baleares, Murcia y Castilla y León**, todas las demás pueden ser interpretadas en

este primer plano principal. Con respecto a Castilla y León y Murcia no se pueden establecer sus perfiles de precariedad porque no se obtiene una buena calidad ni en el plano ni en el espacio tridimensional.

Con respecto a los índices de precariedad obtenemos que en el primer plano principal están todos bien representados, excepto Salud2, que recibe su mayor contribución de Eje 3.

Vecindad2, Precariedad Socio- Económica y Condiciones son variables de Eje 1; Salud1 y Vecindad1 es de Eje 2 y Salud2 es de Eje 3 (Ver tabla 7).

Tabla 7. Contribuciones relativas de los Ejes Factoriales a los índices

Índices	Eje 1	Eje 2	Eje 3
Salud1	21	733	96
Finanzas	365	212	186
Vecindad1	327	347	144
Vecindad2	682	143	12
Salud2	9	152	610
Prec. Socio-Económica (SE)	550	5	138
Condiciones	756	34	26

A la vista de la tabla anterior podemos decir que el indicador de precariedad Salud1 es, principalmente, de eje 2. Por otro lado los indicadores problemas financieros, vecindad1 y vecindad2 son de plano 1-2. También tenemos que el índice Salud2 es de plano 2-3 mientras que la precariedad socio-económica es de plano 1-3. Finalmente comentar la precariedad de condiciones es de eje 1.

Vamos a concluir este apartado analizando el mapa de dos planos factoriales (1-2, y 1-3). No analizaremos el plano 2-3 debido a la escasa información adicional que nos aporta.

Una vez identificadas las Comunidades Autónomas y los índices de Precariedad que están bien representados en el plano factorial 1-2, pasamos a su interpretación.

Cuatro propiedades del HJ- Biplot son centrales en este momento de la interpretación:

- 1) La proximidad entre los marcadores Comunidad se interpreta en términos de similitud en sus perfiles de precariedad.
- 2) Cuando más distantes aparecen los puntos que representan a los marcadores columna (índices de precariedad) del centro de gravedad, mayor variabilidad habrá en el estudio.

- 3) Cuando menor sea el ángulo que forman los vectores que unen los puntos que representan a dos variables con el centro de gravedad, más correlacionadas estarán las variables.
- 4) El ángulo que forma una variable con un eje principal, representa la importancia que tiene esa variable en el poder discriminante del eje.

Teniendo en cuenta estas propiedades y la proyección de las nubes sobre el primer plano principal (Gráfico 1), podemos afirmar:

En este plano biplot se detectan dos gradientes multivariantes en el perfil de precariedad de las Comunidades Autónomas Españolas: uno horizontal y otro vertical. Las dimensiones de precariedad más correlacionadas con el primer gradiente son **Condiciones** y **Prec. Socio-Económica (SE)**, ya que son los que forman ángulos más agudos con el eje factorial 1 (el horizontal). Por tanto estos dos aspectos de la precariedad son los más importantes a la hora de discriminar entre las Comunidades que presentan más y menos niveles de precariedad; es decir, los índices más importantes a la hora de discriminar entre uno y otro extremo del gradiente horizontal de precariedad son Con y SE.

Los extremos del gradiente horizontal, vienen marcados por las Comunidades Na(Navarra) con el menor nivel de precariedad de España en sus dimensiones Con y SE, y Ceu (Ceuta) que presenta el mayor índice de precariedad. Visualizando las proyecciones de todas y cada una de las Comunidades, a excepción de las mal representadas en este plano (Cataluña, Comunidad Valenciana, Baleares, Murcia y Castilla y León), podemos establecer un orden de precariedad de las Comunidades que yendo de menor a mayor sería: Na, Ar, Pv, As, R, Ca, Ex, An, Cs, Me y Ceu.

Navarra y el País Vasco (Pv) son las comunidades que en numerosos estudios se han considerado como menos precarios mientras que Canarias, Ceuta y Melilla son siempre los que peores resultados ofrecen. Veremos, pues, a lo largo de este apartado como se cumple esta teoría sociológica.

Analizando ahora el gradiente vertical vemos que el aspecto de la precariedad más correlacionado con ese gradiente (segundo eje principal) es **Salud1** ya que forma el ángulo más pequeño con ese eje. En menor medida, también tiene importancia **Vecindad2 (Vec2)**.

Interpretando ahora las proyecciones de las Comunidades Autónomas sobre Eje 2 vemos que **Galicia** marca el extremo del gradiente vertical indicando que tienen los niveles más altos de precariedad en **Salud1**, mientras que, Melilla y Madrid marcan

el extremo opuesto indicando que tienen los niveles más bajos de precariedad en Salud1, seguidas de **Ceuta**, País Vasco (**Pv**) y **Aragón**.

Aunque este aspecto de la precariedad es el más importante a la hora de discriminar entre las Comunidades Autónomas y por eso hemos centrado más la atención en él, también merece resaltarse el papel de **Vec1**, que como puede observarse en el gráfico HJ-Biplot forma un ángulo agudo con el eje factorial 2. Salud1 y Vecindad1 (Vec1) forman un ángulo muy obtuso, lo que significa que están correlacionadas de manera inversa. Por esta razón las Comunidades Autónomas **Ceuta**, **Melilla**, **Madrid** son las que presentan mayores valores de precariedad en **Vec1** y las Comunidades **Galicia**, **Canarias (Cs)**, **Asturias**, **Cantabria** son las que presentan menores valores de precariedad en **Vecindad**.

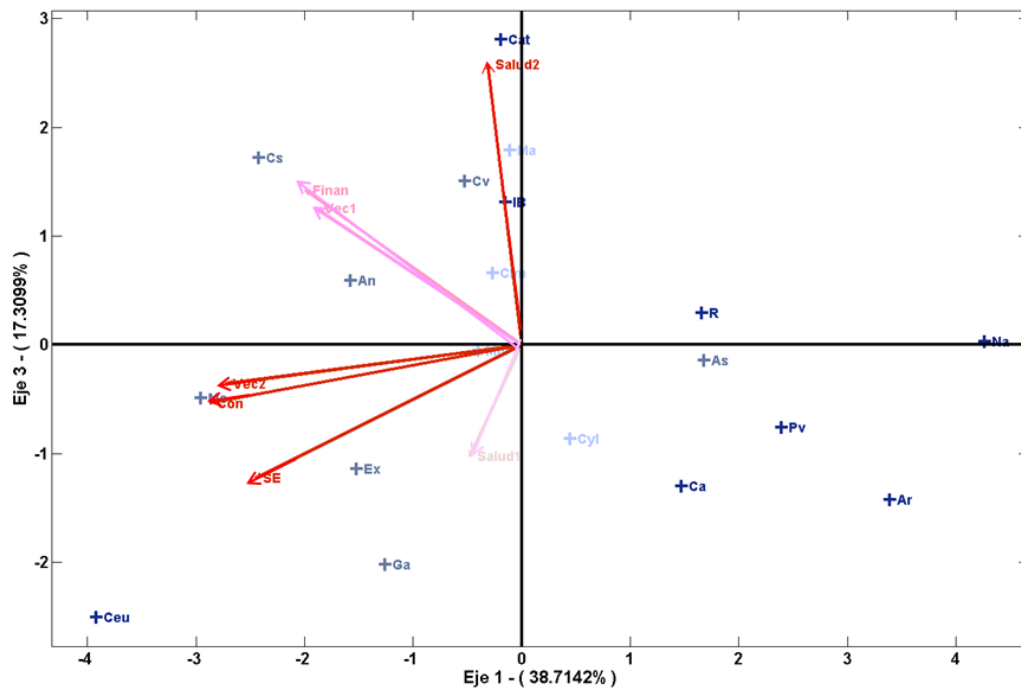
Para conocer el orden de las Comunidades con respecto a cualquiera de los demás índices de precariedad bastaría con imaginar cada Comunidad proyectada sobre el vector que identifica ese índice de precariedad. Las Comunidades que proyectan en el sentido de la punta del vector tienen valores más altos de precariedad y las que proyectan en sentido contrario, los valores más bajos.

Si quisiéramos conocer el perfil de precariedad de una Comunidad concreta podríamos visualizarlo en el gráfico HJ- Biplot proyectando la Comunidad sobre todos y cada uno de los índices (vectores). Fijemos la atención, por ejemplo en **Madrid**. Proyecta en el sentido de la punta de flecha de Vecindad1 y Vecindad2, luego los valores de precariedad en esos dos índices son altos; sin embargo, proyecta en el sentido contrario a la punta de flecha en todos los demás, luego en los demás aspectos de la precariedad toma valores bajos. Análogo razonamiento para cualquier otra Comunidad.

Aun podemos sacar más información explorando los ángulos que forman en el plano 1-2 del HJ- Biplot los vectores marcadores de los distintos índices. Cuando los ángulos son agudos indican relación directa, cuando son obtusos inversa. Si son rectos, independencia. Comentaremos alguno. Por ejemplo, Vecindad1 y Vecindad2 covarían (obsérvese que sin embargo que **Vec1** es importante para discriminar entre Comunidades y **Vec2** no). **Salud1** y **Vec2** tienen un comportamiento prácticamente independiente, sin embargo **Salud 1** y **Vec1** covarían de manera inversa, **ES** está más relacionado con **Con** que con **Finan**, etc.

Para estudiar el comportamiento de las Comunidades **Cataluña**, **Comunidad Valenciana** y **Baleares** es necesario recurrir a un plano en el que intervenga el Eje 3. También es necesario analizar este plano para interpretar el indicador **Salud2** tal y como muestra la tabla 7. Analicemos pues el plano 1-3.

Gráfico 2. . Representación HJ-Biplot del plano 1-3.



Podemos ver que el plano 1-3 absorbe alrededor de un 56% de varianza explicada. En este gráfico vamos a describir aquellas Comunidades Autónomas que son de Eje 3.

Tal y como se comentó en el apartado anterior el indicador Salud2 (motivos por los que no se asiste al médico o dentista) es de eje 3 debido al ángulo agudo que forma con este eje.

Interpretando las proyecciones de las Comunidades Autónomas (**Cataluña, Comunidad Valenciana y Baleares**) sobre **Salud2** vemos que **Cataluña** marca el extremo del indicador por lo que tienen los niveles más altos de precariedad en Salud2 seguidos de Comunidad Valenciana (**Cv**) y Baleares (**IB**). Las Comunidades que proyectan en el sentido de la punta del vector tienen valores más altos de precariedad y las que proyectan en sentido contrario, los valores más bajos. Por lo que concluimos que las tres tienen valores altos de precariedad en este indicador.

Aunque el HJ-Biplot tenga un carácter descriptivo ha sido posible obtener resultados muy similares que en el caso del MLG con la ventaja de poder ver en un plano y, por tanto, de forma más sencilla y visual tanto los indicadores como las comunidades autónomas. Esta técnica nos ha permitido establecer relaciones entre indicadores, entre comunidades y entre comunidades e indicadores perfilándose como una técnica mucho más completa al estudiarse de forma conjunta todas las

dimensiones (es por ello por lo que es una técnica multivariante) a diferencia del MLG que se estudia cada indicador por separado.

5.5. Estudio de las estructuras multifactoriales en España, Francia, Italia, Bélgica y Portugal

En este caso disponemos de 5 matrices de datos que recogen la información de los diferentes aspectos de la precariedad en los individuos de los 5 países.

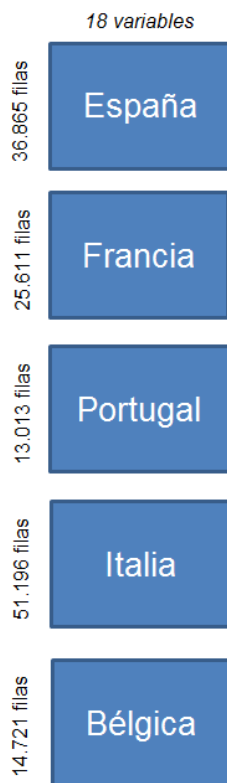
Comparamos estos cinco países pues todos ellos están dentro del modelo conservador y de tradición católica (Esping Andersen, 2000), pero con una clara diferencia entre los más familiaristas (España, Portugal e Italia) y los que tienen un régimen de bienestar más desarrollado (Bélgica y Francia).

En este caso trabajamos con las variables de la propia encuesta y las dos que fueron creadas (labour_stat_full\$ y Hhincome_CLASS15)²:

- i. Carga financiera debido al pago de compras a plazos o préstamos (Db)
- ii. Carga financiera debido al pago de la hipoteca (Fb).
- iii. Capacidad para llegar a fin de mes (Em).
- iv. Demasiado oscuro, poca luz (Dk).
- v. Posibilidad de tener la casa caliente (Wr).
- vi. Existencia de goteras (ro).
- vii. Demasiado ruido (N).
- viii. Polución (P).
- ix. Crimen (C).
- x. Vecindario sucio (ll).
- xi. Servicios públicos en mal estado (da).
- xii. Autopercepción del estado general de salud (Hs).
- xiii. Motivos por los que no se ha visitado un médico (Rd).
- xiv. Sufre de una enfermedad crónica (Chr).
- xv. Limitación en las actividades debido a tener problemas de salud (LH)
- xvi. Motivos por los que no se ha visitado un dentista (Rd)
- xvii. Precariedad Laboral (lab)
- xviii. Precariedad de Ingresos (Inc)

² Para más información ver figura 1, segunda columna. Se eliminaron las variables: Váter de uso privado (HH091) y Baño o ducha (HH081) que tampoco fueron usadas en el análisis factorial.

Figura 6. Estructura de las 5 matrices



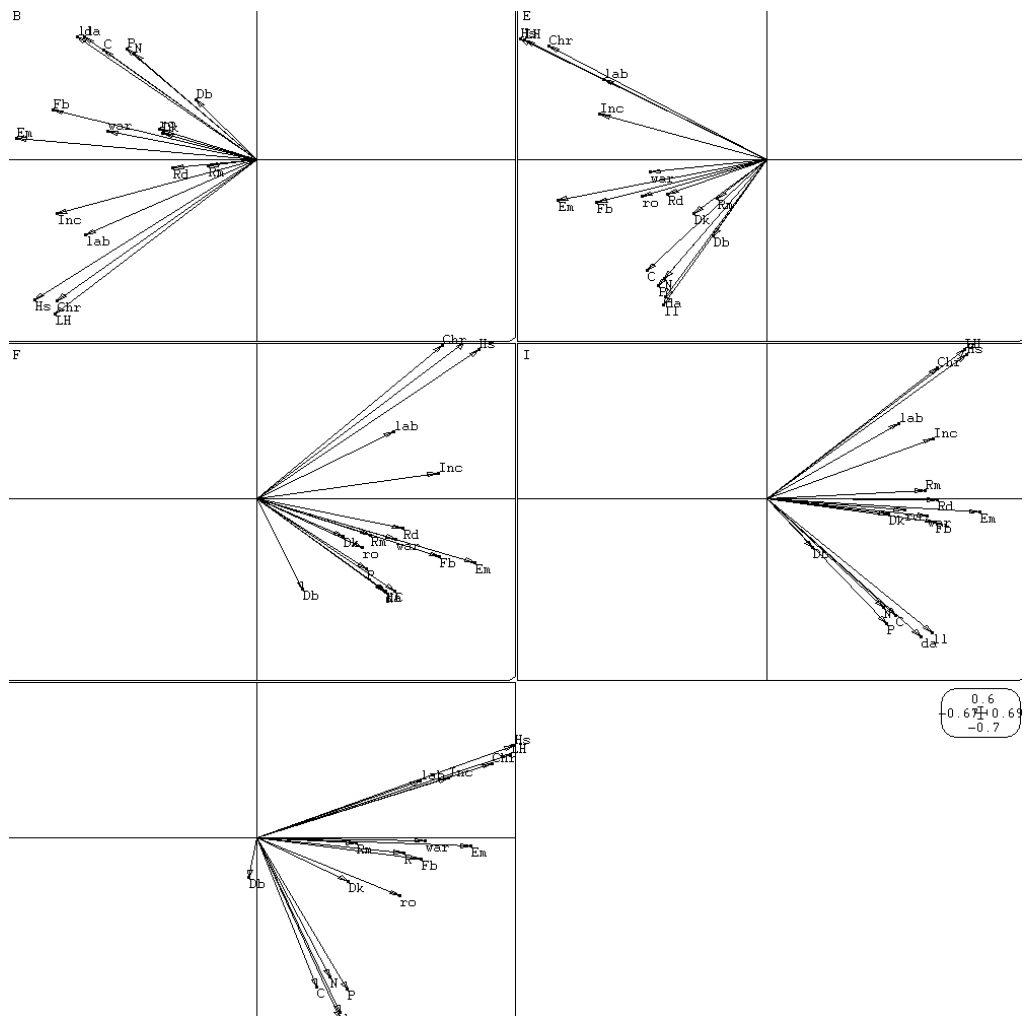
Pretendemos conocer si las estructuras de covariación entre los distintos índices de precariedad son similares en los distintos países y encontrar la estructura consenso que recoja el comportamiento de los cinco países. Para ello aplicamos el STATIS DUAL

Cuando estudiamos España, disponíamos de una sola matriz de datos X que contenía la información de I ciudadanos y J índices de precariedad. En ese caso era posible representar o columnas de la matriz de partida en un subespacio en el cual era posible interpretar la proximidad de las filas en términos de similitud entre ellas. Para poder realizar ese paso era necesario calcular una matriz de covarianzas (correlaciones) entre variables, buscar sus valores y vectores propios, proyectar y tener definida una distancia en el espacio de representación. Si dos puntos que representaban a dos filas estaban próximos en el plano eso significaba que tomaban valores similares para las variables en estudio.

Ahora pretendemos algo similar, solo que en lugar de comparar filas, queremos comparar matrices. Para ello es necesario, tal como ya vimos al explicar el método, extender a la nueva situación los conceptos de covarianza (producto escalar), correlación, y distancia.

La representación gráfica de la estructura de variables en los 5 países es la siguiente:

Gráfico 3. Estructura de los cinco países



El primer gráfico, corresponde a Bélgica (B). En este caso se distinguen 4 grupos de variables que serían:

1. Llegar a fin de mes(Em), Carga financiera debido a préstamos(Db), Pago hipoteca(fb), Posibilidad de tener la casa caliente(War), Motivos por los que no se ha visitado un dentista(Rd) y Motivos por los que no se ha visitado un médico (Rm).
2. Precariedad de ingresos (inc) y Precariedad laboral (lab) con los que se construye el indicador socio-económico.
3. Otro grupo de variables relacionadas sería el Autopercepción del estado general de salud (hs), Sufre de una enfermedad crónica(Chr) y Limitación en las actividades debido a tener problemas de salud (LH).
4. Resto.

En el caso de España, Italia y Francia, tal y como era de esperar a tenor de las evidencias señaladas anteriormente, la agrupación es muy similar. En ellos se distinguen, más o menos, tres grupos:

1. Un primer grupo sobre la precariedad de salud: Autopercepción del estado general de salud (hs), Sufre de una enfermedad crónica (Chr) y Limitación en las actividades debido a tener problemas de salud (LH).
2. Un segundo grupo formado por las variables Llegar a fin de mes (Em), Carga financiera debido a préstamos (Db), Pago hipoteca(fb)y Posibilidad de tener la casa caliente (War).
3. Y, finalmente, un tercer grupo que contendría las demás variables. Sin embargo es importante señalar el comportamiento tanto de hhincome_class15 (inc) como el de labour_stat_full\$ (lab). Podría considerarse la variable inc como un grupo a parte aunque a veces esté próxima al primer grupo. Al igual que la variable lab que en el caso de España estaría totalmente integrada en el primer grupo, pero no así en los otros dos países.

Finalmente el caso de Portugal acaba por dibujar unos tres grupos:

1. En el primer grupo se observan las tres variables sobre la precariedad en salud pero, además, en este caso las variables lab e inc están incluidas en ella también. Esto sugiere que en el caso de Portugal existe una estrecha relación entre precariedad socio-económica y precariedad en salud.
2. Un segundo grupo con las variables Motivos por los que no se ha visitado un dentista, Posibilidad de tener la casa caliente, Llegar a fin de mes, Pago hipoteca, Existencia de goteras, Demasiado oscuro, poca luz, Motivos por los que no se ha visitado un médico.
3. Mientras que el resto de variables conformarían el tercer grupo.

De la misma forma que calculamos productos escalares o distancias entre individuos o variables, podemos calcular ahora productos escalares o distancias entre países a través de sus objetos representativos (operadores) $X_t^T X_t$. El producto escalar es el producto escalar de Hilbert-Schmidt. El producto escalar permite definir una norma y una distancia entre operadores.

Una vez realizados los dos primeros pasos, es decir, una vez buscadas las configuraciones respectivas y definida una métrica trataremos de conseguir una imagen euclídea de las correspondientes configuraciones.

Para ello basta calcular la matriz de productos escalares entre objetos; es decir, entre configuraciones representativas, y/o la matriz de correlaciones vectoriales y una vez obtenida esa matriz realizar su descomposición espectral y a partir de ella, proyectar en un espacio de baja dimensión donde cada punto va a representar ahora a un país; es decir, a la matriz que cruza individuos por variables en cada país.

La matriz de correlaciones vectoriales encontrada es la siguiente:

Tabla 8. Matriz de correlaciones vectoriales.

	Belgica	España	Francia	Italia	Portugal
Belgica	1000				
España	987	1000			
Francia	987	991	1000		
Italia	984	987	990	1000	
Portugal	979	988	984	989	1000

Todas las correlaciones de HS son superiores a 0.97, lo cual traduce gran similitud entre las estructuras de precariedad de los 5 países

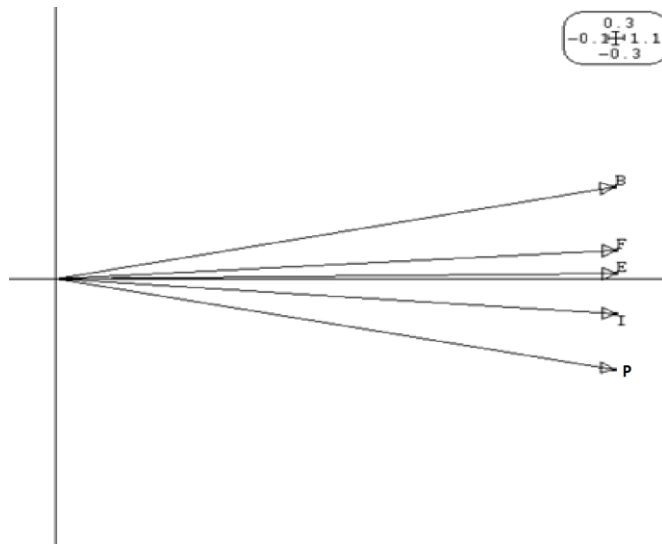
La Etapa de estudio de la interestructura tiene como objetivo realizar la comparación global de la estructura de las **5** matrices de datos. Para ello se ha definido la distancia de Hilbert-Schmidt entre “configuraciones” y se busca la representación euclídea en la cual se verifique que la proximidad entre dos puntos en el plano se corresponde con un parecido entre las tablas en el sentido de la distancia de HS. O dicho de otra manera: la proximidad entre dos puntos del plano traduce una estructura de covariación común a las tablas de datos (matrices) correspondientes.

Por el teorema de Frobenius sabemos que una matriz simétrica con todos los términos positivos admite un primer vector propio en el que todas las componentes son del mismo signo.

Los cosenos de los ángulos entre los vectores que representan a los países aproximan la correlación vectorial entre los mismos, por lo que si hay una estructura común, los ángulos serán pequeños y la mayor parte de la variabilidad estará recogida por el primer eje de la representación.

La imagen euclídea en al que se representan los cinco países es la siguiente:

Gráfico 4. Imagen euclídea que respresenta a los 5 países.



Al retener dos ejes factoriales se consigue una absorción de inercia del 29,3%.

Un aspecto importante que nos puede dar pistas de la buena o mala representación de la inter-estructura es el análisis de las correlaciones entre los países: si existe poca correlación entre alguno de ellos puede suceder que la estructura consenso no represente bien a todos los países.

La representación anterior pone de manifiesto como España es la que tendrá mayor peso en la configuración consenso ya que es la que forma el ángulo más pequeño con el eje X (horizontal). También pone de manifiesto que el país con estructura de precariedad más similar a España es Francia y el siguiente más similar es Italia. Bélgica y Portugal son los menos similares aunque dado que los ángulos son en todos los casos agudos, la estructura es bastante similar en todos los países. Italia es más similar a Portugal que a Francia y a Bélgica. Obsérvese que Bélgica y Portugal marcan los extremos precisamente porque Portugal es el país más precario y Bélgica es el que tiene más baja precariedad. Tomando Bélgica como referencia, por ser el país con más baja precariedad, podemos afirmar que el siguiente país con menor precariedad es Francia, le sigue España, después Italia y el más precario es Portugal, tal como ya habíamos señalado.

La información algebraica más detallada puede encontrarse en la tabla siguiente:

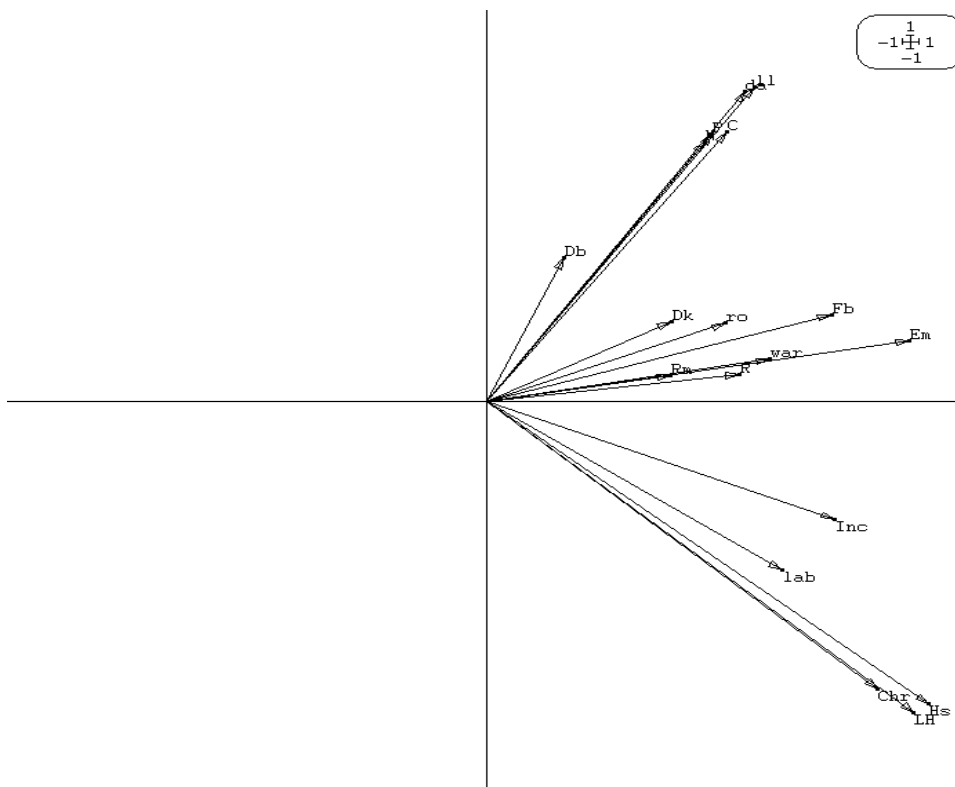
Tabla 9. Información algebraica detallada

Pais	Filas	Pesos	HS	Cos2
Belgica	11551	4,46E+02	8,55E+01	0,571
España	27464	4,48E+02	7,94E+01	0,508
Francia	19211	4,48E+02	7,84E+01	0,5
Italia	40308	4,48E+02	8,41E+01	0,535
Portugal	10716	4,47E+02	8,64E+01	0,552

El coseno al cuadrado de Bélgica y Portugal evidencia el mayor ángulo, es decir, la mayor distancia, ergo la mayor diferencia de pesos.

La representación gráfica de la estructura compromiso es la siguiente:

Gráfico 5. Estructura compromiso



Emergen tres grandes grupos, tal vez cuatro:

1. En primer lugar, las variables relacionadas con la precariedad en salud: Chr, Hs y LH. Es debatible la inclusión de las variables lab e inc, o bien la alternativa, separarlas en un cuarto grupo.
2. Por otro lado, tenemos las variables Dk, ro, Rm, R, war ,Em, Fb
3. Y un último grupo con el resto de variables.

En la publicación: "European comparision on precariousness indicators " (Simó C., Wagener M, 2013), sus conclusiones observan que España, Italia y Francia obtienen resultados muy similares sin embargo observan grandes diferencias entre Portugal y Bélgica. Viendo el gráfico 4 nosotros extraemos las mismas conclusiones. Portugal es el país con más precariedad de los cinco mientras que Bélgica es el país con menos precariedad. Debemos recordar que el documento base aplicaba el análisis factorial a todos los países a la vez y aunque vemos que las distancias entre los países (tabla 9) son muy similares y cabría, por tanto, una justificación al uso de esta técnica, se ha preferido aquí emplear el Statis puesto que éste tiene en cuenta la variabilidad entre los países al aplicarles un peso diferente a la hora de confeccionarse la matriz consenso.

6. Discusión

En este apartado pretendemos sistematizar y discutir los resultados obtenidos en cada apartado con fin de ligarlos a la teoría sociológica. Veamos la potencialidad de cada técnica utilizada. En el caso del análisis factorial cabe aquí aportar varios matices. Por un lado, la técnica utilizada no fue diseñada para variables cualitativas sino cuantitativas, sin embargo, fue usada con el objeto de seguir el mismo esquema que el artículo central. Se realizó un HOMALS con el ánimo de valorar la pertinencia de los resultados. Los resultados obtenidos fueron, sin embargo, muy similares y por ello se prosiguió con el análisis factorial.

Los resultados obtenidos concuerdan con la bibliografía analizada pues tal y como menciona Tejerina (2011) “La precarización como proceso toca diferentes dimensiones que tienen que ver con pérdidas o entradas en zonas de riesgo y que hacen referencia a limitaciones en recursos y capacidades de los individuos: trabajo, remuneración, consumo, residencia, cualificación educativa, entorno, vida familiar y afectiva, relaciones sociales, salud y participación cívica.” Si bien en ningún se ha realizado un análisis dinámico. En el análisis factorial fueron encontradas las dimensiones de trabajo y remuneración bajo la variable latente: Socio-económica; la dimensión residencia, bajo el indicador Condiciones; la dimensión entorno, la cual fue dividida en dos (Vecindad1 y Vecindad2); y, finalmente, la dimensión de precariedad salud fue también dividida en dos (Salud1 y Salud2). Los resultados refuerzan la teoría sociológica en la que se basa este trabajo. Ahora bien, las dimensiones de salud y entorno no son unidimensionales con ello queremos decir que es muy pertinente subdividir las en dos. O dicho de otra manera, ciertas subdimensiones no son unidimensionales sino multidimensionales. Un ejemplo de esto es el hecho de que la precariedad en salud esté dividida en dos: la salud percibida general y los motivos por los que no se puede ir al médico cuando se necesita. En este último caso se puede entender que está muy relacionada con la precariedad socio-económica debido a que muchos de los motivos tienen que ver con falta de tiempo o falta de recursos económicos para poder asistir. Cabe pues entender la precariedad no como un constructo multidimensional en el que cada dimensión está separada sino, tal y como era previsible, las múltiples dimensiones de la precariedad están relacionadas unas con otras. Tal y como se comentó en el marco teórico, la precariedad aparece vinculada al Estado de Bienestar y sus políticas de integración y exclusión. Siguiendo como ejemplo el caso francés, la expansión del término se produce primero en relación a la pobreza (Pitrou, 1978). Esta autora es la primera en asociarlo a las familias vulnerables. Posteriormente se asocia con el estatus laboral (Schnapper y Villac, 1989) y en la década de 1990 en relación con el trabajo (Paugam, 2007). Mientras, en el ámbito anglosajón, a partir de la década de 1990 la precariedad se

pone en relación con la flexibilidad y la corrosión (del Bienestar) (Sennett, 2000). Sin embargo, los resultados ilustran que en el caso Español la precariedad general vendría muy determinada, principalmente, por la falta de salud general y, en segundo lugar, por la precariedad financiera. Puesto que nuestro Estado de Bienestar no alcanza las cuotas de países como Francia y Bélgica en la protección contra la precariedad y puesto que el bienestar de las personas está prácticamente determinado por su posición en el mercado de trabajo es, en cierta medida esperable, que en la precariedad globalmente entendida, tanto la precariedad en salud (salud general) como la precariedad financiera obtienen un mayor protagonismo.

En el caso de la realización del HJ-Biplot, los resultados se antojan también interesantes. Sin embargo, cabe enunciar una sugerencia: es pertinente continuar con este estudio mediante la realización de análisis de clusters dentro del mapa factorial resultante y ahondar en este tipo de técnicas muy poco aplicadas en el campo de la Sociología. Debe remarcarse que no se encontró la aplicación de esta técnica al campo de la investigación de la precariedad con lo que, todo evidencia que es interesante seguir por este camino. En la publicación: "European comparison on precariousness indicators" (Simó C., Wagener M, 2013), encontramos una aproximación descriptiva en la que se obtuvieron resultados muy similares aun con una metodología diferente. En el caso de la dimensión socio-económica los autores muestran que las comunidades más precarias son Galicia y Extremadura y, sin embargo, no lo son tanto Ceuta y Melilla. En nuestro caso coincidimos con este artículo en lo referente a Galicia y Extremadura pero nuestros resultados difieren en el caso de Ceuta y Melilla pues, para nosotros, son las comunidades más precarias en este indicador. Por otro lado, en el indicador de vecindario coincidimos completamente debido a que ellos, al igual que nosotros, obtienen que Ceuta y Melilla son las más precarias. En definitiva, se llega a un mismo puerto pues los resultados son muy similares aunque se echa en falta la posibilidad de compararlos para la variable salud general dado que en nuestro estudio ha sido decisiva en el ordenamiento de la precariedad.

En el caso del Statis dual nos remitimos también al mismo trabajo del párrafo anterior para poder comparar resultados. En sus conclusiones observan que España, Italia y Francia obtienen resultados muy similares sin embargo observan grandes diferencias entre Portugal y Bélgica. Viendo el gráfico 5 nosotros extraemos las mismas conclusiones. Portugal es el país con más precariedad de los cinco mientras que Bélgica es el país con menos precariedad. Debemos recordar que el artículo central aplicaba el análisis factorial a todos los países a la vez y aunque vemos que las distancias entre los países (tabla 23) son muy similares y cabría, por tanto, una

justificación al uso de esta técnica, se ha preferido aquí emplear el Stata puesto que éste tiene en cuenta la variabilidad entre los países al aplicarles un peso diferente a la hora de confeccionarse la matriz consenso.

Finalmente se propone ampliar la línea de investigación en la aplicación de técnicas de tres vías al campo del estudio de la precariedad. Se han encontrado algunas aplicaciones en España a las ciencias sociales como, por ejemplo: Herrán JIM., Aguirre CF., Calvo MIL. (2004); Fernandez EA., Lautre IG., Calvo MIL. (2004); y Lopez Vizcaíno ME., Sánchez Fernández P. (2003) .Tan solo hemos encontrado estos tres artículos y son todos posteriores al año 2000 con lo que pensamos que el estudio de datos de tres vías se encuentra todavía entre las técnicas de aplicación emergente. En una futura línea de investigación se aplicarán este tipo de técnicas. Éstas se revelan de gran importancia en el abordaje de análisis dinámicos y de procesos de precarización puesto que la Encuesta de Condiciones de Vida cuenta con una muestra longitudinal. Esta aproximación nos permite una visión alternativa de los modelos lineales con medidas repetidas que han sido muy utilizados (casi en exclusiva) en los estudios longitudinales.

7. Conclusiones

1. El Modelo Factorial, con solución en Componentes Principales, nos ha permitido inspeccionar el constructo *precariedad* desde una perspectiva multivariante, y descubrir que ciertas dimensiones de las hipotetizadas por los expertos, deben ser revisadas.
 - a. El indicador de *precariedad en salud* recoge dos aspectos que, según el modelo factorial, deben ser estudiados de forma separada: Los motivos de precariedad en salud por un lado y los aspectos generales de precariedad en salud, por otro.
 - b. El indicador de vecindad también engloba dos dimensiones que deben ser diferenciadas: ruido, polución e inseguridad por un lado y suciedad y malos servicios públicos por otro.
 - c. Los aspectos socioeconómicos, sin embargo, que los expertos separaban en dos variables, conforman realmente, según el modelo, una única dimensión latente del constructo precariedad.
2. El estudio de precariedad, en las Comunidades Autónomas españolas, mediante un HJ-Biplot, ha puesto de manifiesto tres gradientes bien diferenciados y un ordenamiento de las Comunidades según esos gradientes.
3. El gradiente que absorbe mayor información es el conformado por Suciedad en el vecindario y Servicios públicos en mal estado, Motivos por los que no acude al médico o al dentista, y las Condiciones de la vivienda (poca luz, goteras, etc.).
 - a. El segundo gradiente con mayor capacidad informativa está conformado por los aspectos generales de Salud y la polución, ruido e inseguridad en el barrio.
 - b. El tercer gradiente se asocia con los motivos por los que no accede a la sanidad.
4. Las Condiciones de la vivienda y la precariedad laboral y bajos ingresos son los aspectos más importantes a la hora de discriminar entre las Comunidades. Ceuta es la que presenta mayor precariedad en estos aspectos y Navarra presenta los valores más bajos. No solo hemos identificado las Comunidades que marcan los extremos del gradiente, sino que hemos conseguido ordenar a todas las comunidades según sus niveles de precariedad en esta dimensión: Navarra, Aragón, País

Vasco, Asturias, La Rioja, Cataluña, Extremadura, Andalucía, Canarias, Melilla y Ceuta.

5. Galicia marca el extremo del gradiente vertical indicando que tienen los niveles más altos de precariedad en los aspectos generales de Salud, mientras que, Melilla y Madrid marcan el extremo opuesto indicando que tienen los niveles más bajos de precariedad en Salud¹, seguidas de Ceuta, País Vasco y Aragón.
6. Cataluña, Comunidad Valenciana y Baleares forman un cluster bien diferenciado del resto de la Comunidades, fundamentalmente por los motivos por los que no acceden a la sanidad.
7. El STATIS DUAL nos ha permitido descubrir la parte común de las estructuras de covariación entre los distintos índices de precariedad, en los 5 países estudiados, los cuales fueron elegidos porque, aunque todos ellos están dentro del modelo conservador y de tradición católica, existe una clara diferencia entre los más familiristas (España, Portugal e Italia) y los que tienen un régimen de bienestar más desarrollado (Bélgica y Francia).
8. Asimismo, el STATIS DUAL nos ha permitido encontrar la estructura consenso que recoge el comportamiento de los cinco países y demostrar que Italia Francia y España tiene un patrón de precariedad similar, bien diferenciado del de Bélgica (el país con menores niveles de precariedad) y del de Portugal (el país con mayores niveles de precariedad).
9. España es el País que presenta mayor peso en la configuración consenso por lo cual puede ser considerado como el país de referencia. El país con estructura de precariedad más similar a España es Francia.
10. El orden, según niveles de precariedad, es (de menor a mayor precariedad): Bélgica, Francia, España, Italia y Portugal.

8. Bibliografía

- Ayala L. (2008). *Desigualdad, pobreza y privación*. Madrid: Fundación Foessa.
- Castel, R. (1997) *Trabajo y utilidad para el mundo*. Revista Internacional del Trabajo, 115: 671-678.
- Escouffier B. (1980). *L'analyse conjointe de plusieurs matrices de données*. Biometrie et Temps. Paris: Société Française de Biometrie, 59–76.
- Esping Andresen, G. (2000). *Fundamentos sociales de las economías preindustriales*. Barcelona: Ariel.
- Greenacre M. (1986). *Theory and Applications of Correspondence Analysis*. . London: Academic Press.
- Fernandez EA., Lautre IG., Calvo MIL. (2004). *Análisis de la evolución a través de encuestas: trayectoria electoral de las comunidades autónomas españolas en el período 1977-2004*. Metodología de Encuestas.
- Gabriel, K.R. (1971). *The biplot-graphic display of matrices with application to principal component analysis*. Biometrika, 58: 453-467.
- Galindo, M.P. (1986). *Una alternativa de representación simultánea: HJ-Biplot*. Questíio, 10:13-23.
- Galindo, M.P., Cuadras, C.M. (1986). *Una extensión del método biplot a su relación con otras técnicas*. Bioestadística y Biomatemática. Universidad de Barcelona, 17.
- Gower, J. C. (1971). *Statistical methods of comparing different multivariate analyses of the same data*. Mathematics in the archaeological and historical sciences, 138-149.
- Herrán JIM., Aguirre CF., Calvo MIL. (2004). *ACM y Statis dual ponderado: dos técnicas complementarias para analizar una visión de la cultura de la Universidad*. Estadística española - dialnet.unirioja.es

Hervé, A., J. Williams, L., Valentin, D. Bennani-Dosse, M. (2012). *STATIS and DISTATIS: optimum multitable principal component analysis and three way metric multidimensional scaling*.

http://www.utdallas.edu/~herve/abdi_Wires_AWVB2012_Final.pdf

Hotelling, H. (1933). *Analysis of a Complex of Statistical Variables Into Principal Components*. Journal of Educational Psychology, 24:417-441 y 498-520.

Korth B., Tucker LR. (1976). *Procrustes matching by congruence coefficients*. Psychometrika, 41:531–535.

Lopez Vizcaíno ME., Sánchez Fernández P. (2003). *Tipoloxía dos fogares galegos a través da aplicación de técnicas factoriais de análise de táboas múltiples*. En sgapeio.es. http://www.sgapeio.es/INFORMEST/VI Congreso/artigos/sesion7_23.pdf

L'Hermier des Plantes H. (1976). *Structuration des Tableaux a Trois Indices de la Statistique*. Thèse de troisième cycle. Université de Montpellier.

Lavit, C. (1988). *Analyse Conjointe de Tableaux Quantitatifs*. Masson Editorial

Paugam, S. (2007). *Las formas elementales de la pobreza*. Madrid: Alianza Editorial

Pearson, K. (1901). *On lines and planes of closest fit to systems of points in space*. Philosophical Magazine, 2:559-572.

Pitrou, A. (1978). *La vie précaire. Les familles face à leurs difficultés*. Paris: Études CNAF

Salvador Figueras, M., Gargallo Valero, P. (2006). *Análisis Factorial*. En ciberconta.unizar.es. <http://www.5campus.com/leccion/factorial>

Sanchis, E. (2011). *Trabajo y paro en la sociedad postindustrial*. Valencia: Tirant Lo Blanch.

Simó, C., Sanchez, I., Llopis, R., Castelló, R. (2011). *La política en la red*. Murcia, X CONGRESO DE AECPA.

Schanpper, D., D. Villac, M. (1989). *Rapport à l'emploi, protection sociale et status sociaux*. Revue Française de Sociologie, 30:3-29

Sennett, R. (2000). *La corrosión del carácter. Las consecuencias personales del trabajo en el nuevo capitalismo*, Barcelona: Anagrama

Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. The American Journal of Psychology, 15:201–292.

Subirats, J., Gallego, R. (2002). *Veinte años de autonomías en España: leyes, políticas públicas, instituciones y opinión pública*. Madrid: CIS

Subirats, J. (2004). *Pobreza y exclusión social. Un análisis de la realidad española y europea*. Fundación La Caixa. http://obrasocial.lacaixa.es/StaticFiles/StaticFiles/a84f7102892ef010VgnVCM1000000e8cf10aRCRD/es/es16_esp.pdf

Subirats, J. (2008) *¿La fiesta del 'preariado'?* El País edición Cataluña de 1 de Mayo de 2008. http://elpais.com/diario/2008/05/01/catalunya/1209604039_850215.html

Tejerina, B (2011). *La precariedad vital. Los procesos de precarización de la vida social y de la identidad en la sociedad española contemporánea. Dossier teórico*. parte de la Investigación I+D+I del MICINN convocatoria 2008

Thurstone, L. L. (1947). *Multiple-factor analysis*. Chicago: University of Chicago Press.

Villardón, J. (2010). *Análisis de Componentes Principales*. Departamento de Estadística de la Universidad de Salamanca. <http://biplot.usal.es/doctorado/3ciclo/bienio-04-06/ACP/ACP.pdf>

Programas usados:

SPSS Version 20: <http://www-01.ibm.com/software/es/stats20/>

Infostat: <http://www.infostat.com.ar/>

ADE 4: <http://biomol.univ-lyon1.fr/ADE-4-old/ADE-4.html>

Vicente Villardon, J.L (2010). MULTBILOT: A package for Multivariate analysis using Biplots. Departamento de Estadística. Universidad de Salamanca. (<http://biplot.usal.es/ClassicalBiplot/index.html>)