

Equilibrio competitivo en Liga española de fútbol de Primera División: una aplicación del análisis de datos funcionales

Francisco Montes y Ramón Sala
*Departamentos de Estadística e Investigación Operativa y
Matemáticas para la Economía y la Empresa
Universitat de València*

RESUMEN

La supremacía de unos pocos equipos sobre los otros participantes en la liga de fútbol profesional, es un hecho común en la mayoría de ligas europeas. La Liga española (La Primera División) no es una excepción. Para evaluar este hecho usamos las técnicas del análisis de datos funcionales (fda). Para ello se realizan dos tipos de simulaciones, una simulación uniforme y una ponderada. Esta simulación es la principal aportación de este trabajo, ya que por medio de ella se obtiene una distribución de probabilidad de los tres posibles resultados de un partido a partir de los resultados obtenidos de los 3040 partidos examinados en las 8 temporadas analizadas y distinguiendo las categorías de los equipos. Finalmente, se analiza también desde la temporada 1987-88 hasta la 2009-10. Todos los análisis realizados demuestran la existencia de un desequilibrio importante en la liga española, acentuándose en las últimas temporadas a favor del Real Madrid y del FC Barcelona.

Palabras claves: Análisis de datos funcionales; Fútbol; Equilibrio competitivo.

Área temática: Métodos Cuantitativos y Deporte

ABSTRACT

The supremacy of a few teams over the other participants is a common factor in the major European football leagues. The Spanish first division league (la Primera División) is not an exception. In order to demonstrate this fact, functional data analysis (fda) is used to analyze football league classifications. In doing so, two types of simulations have been performed, a uniform and a weighted simulation. This second simulation is the main contribution of this work, since it provides a probability distribution of the three possible outcomes of a match. To verify the usefulness of this approach, an empirical application is carried out using the results of 3040 matches examined during the last eight seasons (2002-03 to 2009-10) distinguishing differences in category between teams. For all the years examined, the results demonstrate the existence of a significant imbalance in the Spanish league, accentuated in recent seasons for Real Madrid and FC Barcelona.

Keywords: Functional data analysis; Football; Competitive Balance.

1. INTRODUCCIÓN

El equilibrio entre los equipos que participan en una competición deportiva supone la existencia de igualdad de fuerzas entre ellos e implica una incertidumbre máxima en los resultados de sus enfrentamientos. Desde el punto de vista deportivo la incertidumbre de la clasificación final aumenta el interés de los aficionados. Los aspectos económicos que de ello se derivan no son de menor interés, tanto por la asistencia de espectadores a los estadios como por los previsible ingresos de las televisiones.

La teoría del equilibrio competitivo entre los equipos fue desarrollada por Rottenberg (1956) y ha sido objeto de posteriores desarrollos teóricos Scully (1989), Scully (1995), Dobson y Goddard (2001) y de aplicaciones prácticas, entre ellas las de Vrooman (1995) y Eckard (2001) en el campo del beisbol y las de Szymanski (2001) y Neale (1964) en el del fútbol.

Distintos son los métodos empleados por los autores para contrastar el equilibrio competitivo. Dobson y Goddard (2004) emplean un modelo polinomial para establecer las tendencias en las competiciones inglesa y escocesa, detectando que el desequilibrio aparece entre los años 70 y 80 como consecuencia de la movilidad de los jugadores y los cambios en los contratos, que conducen a un desplazamiento de los buenos jugadores hacia los equipos con mayores recursos. Buzzacchi et al. (2003) obtienen la distribución de probabilidad asociada al número de equipos distintos que han ocupado las k primeras posiciones en una liga a lo largo de T temporadas. La distribución se calcula bajo el supuesto de equilibrio perfecto entre los n equipos participantes. Concluyen los autores que las ligas cerradas, sin promociones ni ascensos, son más equilibradas que las ligas abiertas, con ascensos y promociones.

No existe, hasta donde llega nuestro conocimiento, un estudio similar para la Liga española de primera división. Lo que no nos impide prever cuál será el resultado: un patente desequilibrio. A pesar de este previsible resultado, el presente artículo se ocupa de ello. ¿Por qué? Por dos razones: la distinta herramienta empleada para llevar a cabo el estudio, basada en el análisis de datos funcionales, y el uso de una distribución no uniforme sobre las opciones de victoria por parte de los equipos a la hora de contrastar las hipótesis de equilibrio. Estrictamente hablando, y de acuerdo con este segundo aspecto, no se contrasta la hipótesis de equilibrio competitivo, sino la existencia de un desequilibrio mayor que el observado a lo largo de las 8 temporadas analizadas.

El artículo se estructura como sigue. La Sección 2 describe someramente las técnicas de análisis de datos funcionales (*fda*) que vamos a utilizar. La Sección 3 se ocupa de los datos objeto del estudio resumiendo la información que consideramos relevante. La Sección 4 muestra

los resultados de aplicar las técnicas de *(fda)* a la Liga española y, por último, la Sección 5 presenta las conclusiones que de estos resultados se derivan.

2. DETECCIÓN DE OUTLIERS EN DATOS FUNCIONALES

La clasificación al final de una temporada cualquiera de la liga de fútbol puede ser considerada como una función discreta. Si $x=1, 2, \dots, n$, donde n es el número de equipos en competición, la función $y(x)$ representa la puntuación obtenida por el equipo que ha ocupado la x -ésima posición en la tabla. En un reciente trabajo, Shang y Hyndman (2010) presentan métodos gráficos para la representación de datos funcionales mediante la librería *rainbow* de R (Shang y Hyndman, 2010), que permiten la detección de outliers y que serán utilizados posteriormente.

2.1. Diagrama funcional basado en la profundidad (*functional bagplot*)

El primero de ellos, denominado *bivariate bagplot* por los autores, está basado en la noción de bagplot introducido por Rousseeuw et al. (1999). Se parte de la noción de profundidad semiespacial Tukey (1975), que mide el grado de centralidad de un punto $\theta \in \mathbb{R}^2$ con respecto a una nube puntos, $Z = \{z_1, z_2, \dots, z_n\}$. La profundidad semiespacial de θ con respecto a Z , $ldepth(\theta, Z)$, se define como el menor número de puntos de Z que caen en cualquier semiplano cerrado determinado por una recta que pasa por θ , con lo que $0 \leq ldepth(\theta, Z) \leq n$. La región de profundidad k , D_k , es el conjunto de todos los θ con $ldepth(\theta, Z) \geq k$. Las regiones de profundidad son polígonos convexos que verifican $D_{k+1} \subset D_k$. A partir de estos elementos se obtiene el *bagplot* que consta de tres partes:

- la *mediana*, el punto θ con mayor $ldepth(\theta, Z)$ o, en caso de no ser un único punto, el centro de gravedad de la región más profunda.
- El *saco (bag)*, aquella región B comprendida entre dos regiones consecutivas, D_{k+1} y D_k tales que $\#D_k \leq \lceil n/2 \rceil \leq \#D_{k-1}$ (siendo $\#$ el cardinal del conjunto y $\lceil \cdot \rceil$ la parte entera por defecto) y que se obtiene por interpolación lineal tomando como referencia la mediana, y

- La *corteza* (fence), una región obtenida mediante una homotecia de factor ρ del saco, tomando siempre como referencia la mediana.

Shang y Hyndman (2010) aplican la anterior construcción al conjunto Z formado por los puntos cuyas coordenadas son los scores de las dos primeras componentes principales de la familia de funciones. En su trabajo original, Rousseeuw et al. (1999) utilizan $\rho = 3$ pero Shang y Hyndman (2010) proponen tomar $\rho = 2.58$ porque si los scores bivariantes siguen una distribución normal, la corteza contendrá el 99% de las observaciones.

Puesto que cada punto representa una función, las gráficas de las funciones contenidas en el saco y la corteza dan lugar a lo que los autores denominan *functional bagplot*, una región del plano constituida por la curva mediana, la correspondiente al punto más profundo, una región interior definida por las gráficas de las curvas correspondientes a los puntos del saco, y que son el 50% del total, y una región exterior definida por las gráficas de las curvas correspondientes a los puntos de la corteza.

2.2. Diagrama funcional basado en la densidad (*functional HDR boxplot*)

Este nuevo gráfico está basado en el diagrama bivalente de regiones de alta densidad (*bivariate HDR boxplot*) propuesto por Hyndman (1996), aplicado también a los scores de las dos primeras componentes principales de la familia de funciones. Estimada una densidad, $f(\cdot)$ a partir del conjunto de dichos scores, Z , la región de alta densidad (HDR) del $100(1-\alpha)\%$ se define como

$$R_\alpha = \{z; f(z) \geq p_\alpha\}$$

donde $p_\alpha = \int_{R_\alpha} f(z) dz = 1 - \alpha$. Se trata, en definitiva, de la región con probabilidad de cubrimiento $1 - \alpha$ en cuyos puntos la densidad estimada es mayor que en cualquiera de los puntos de su complementario. En el caso bivalente las HDR son contornos que crecen a medida que α disminuye.

En el *bivariate HDR boxplot* suelen representarse los conjuntos correspondientes a $\alpha = 0.5$ (interna) y $\alpha = 0.01$ (externa). El punto de mayor densidad, equivalente al más

profundo en el *bivariate bagplot*, es ahora la moda, definida como $\{z_m; \hat{f}(z) = \sup_{z \in Z} \hat{f}(z)\}$.

Una segunda gráfica, el *functional HDR boxplot*, se obtiene al representar las funciones correspondientes a los puntos contenidos en sus dos conjuntos.

2.3. Detección de outliers

Como señalan Shang y Hyndman (2010), ambos tipos de gráficos permiten identificar las funciones outliers. Serán aquellas que quedan fuera de las regiones exteriores. En su artículo comparan los dos métodos anteriores con los propuestos por otros autores (Febrero et al., 2007; Hyndman y Ullah, 2007; Filzmoser et al., 2008) utilizando datos de mortalidad masculina en Francia en el periodo 1899-2005 y datos de la temperatura media mensual de la superficie del mar en la región de El Niño en el periodo enero 1951 a diciembre de 2007. Concluyen los autores que, teniendo en cuenta el tiempo de computación y el número de outliers correctamente detectados, los dos métodos propuestos se comportan mejor que los restantes.

3. DATOS

Disponemos de una base de datos con los resultados de los 3040 partidos de la Liga española de fútbol de Primera división correspondientes a ocho temporadas, de la 2002-03 a la 2009-10 (8 temporadas x 38 jornadas/temporada x 10 partidos/jornada). Para estudiar el equilibrio competitivo la información más relevante es la clasificación al final de cada temporada, que se recoge en la Tabla 1. En ella no aparecen los equipos porque no importa qué equipo ocupó qué posición sino el número de puntos conseguido y el lugar en la tabla clasificatoria.

Clasificación	2-03	3-04	4-05	5-06	6-07	7-08	8-09	9-10
1	8	7	4	2	6	2	7	0
2	6	2	0	0	6	4	8	6
3	2	1	5	9	1	4	0	1
4	1	0	2	8	6	4	7	3
5	0	6	1	8	2	1	5	2
6	6	5	0	4	0	7	2	8
7	5	5	8	7	0	6	8	6
8	4	4	1	5	4	2	5	4
9	2	2	1	4	2	0	1	7
10	0	1	1	2	0	9	7	7
11	7	1	0	6	9	8	6	4
12	7	8	0	5	9	8	6	3
13	6	8	7	3	7	6	4	2
14	6	7	7	2	6	6	3	1
15	5	6	6	1	2	4	3	0
16	4	3	4	0	0	3	3	9
17	3	2	9	0	0	3	2	7
18	6	1	7	9	9	2	2	6
19	5	9	9	6	5	0	5	6
20	2	6	8	4	8	6	3	4

Tabla 1: Clasificación final de las temporadas 2002-03 a 2009-10.
Fuente: Elaboración propia

Como indicábamos anteriormente, cada clasificación puede ser considerada como una función discreta, $y(x)$, cuyo valor es el número de puntos conseguido por el equipo que finalizó

la temporada en la posición x -ésima, $x=1,2,\dots,20$. Las gráficas de las 8 funciones se muestran en la Figura 1.

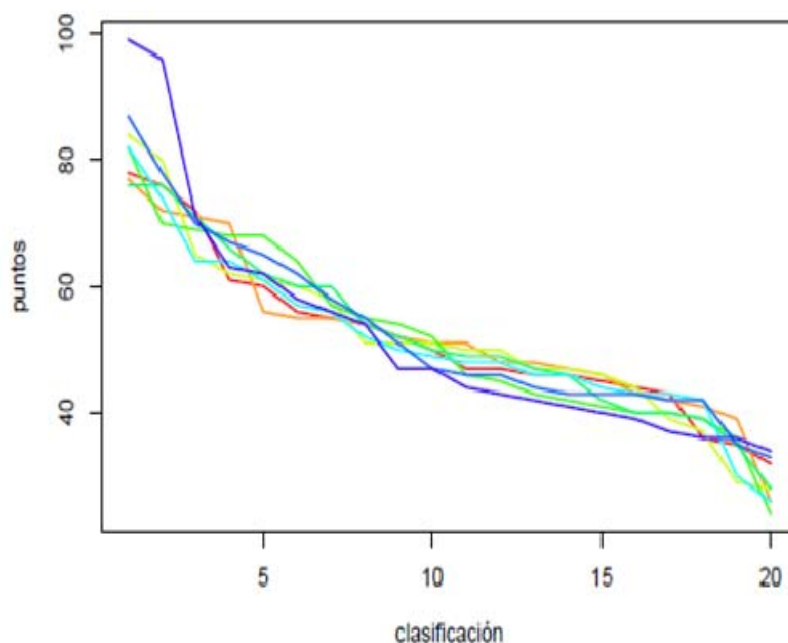


Figura 1: Gráficos de las clasificaciones de las 8 temporadas
Fuente: Elaboración propia

4. RESULTADOS

La aceptación o rechazo de la hipótesis de equilibrio competitivo exige comparar estas clasificaciones con las que obtendríamos bajo dicho supuesto, que implica una distribución uniforme sobre los tres resultados posibles de un partido: victoria local, victoria visitante o empate. Para ello simularemos 200 clasificaciones en estas condiciones y analizaremos las 208 funciones con las herramientas de la sección anterior.

Como ya dijimos en la Introducción, el rechazo de la uniformidad es previsible, como así ocurre. Cualquier aficionado hubiera adelantado este resultado. Se hace necesario un análisis más realista que tome en consideración el desequilibrio entre los equipos y que contraste, no la uniformidad, sino un desequilibrio mayor del observado. Esta nueva aproximación requiere, en primer lugar medir el desequilibrio entre los equipos y, a continuación, estimar una distribución de probabilidad sobre los tres resultados de un partido coherente con este nuevo planteamiento.

La Tabla 2 cumple con el primer requerimiento. Muestra los equipos que han participado en la Primera División en alguna de las 8 temporadas estudiadas. La segunda columna asigna a cada equipo una categoría de 1 (mayor) a 4 (menor). La asignación se llevó

a cabo mediante un análisis cluster en el que hemos hecho intervenir tres variables: el *presupuesto*, los *puntos* y los *ingresos por derechos de TV*, todas ellas acumuladas a lo largo de las 8 temporadas.

Equipo	Categoría	Equipo	Categoría
F.C.Barcelona	1	Osasuna	3
Real Madrid	1	Getafe	3
Atlético de Madrid	2	C.D. Tenerife	4
Deportivo A Coruña	2	Deportivo Alavés	4
Valencia C.F.	2	Rayo Vallecano	4
Villarreal	2	C.D. Numancia	4
Sevilla F.C.	2	Albacete	4
Athletic Club	3	Levante	4
Real Betis	3	Murcia	4
R.C. Celta de Vigo	3	Recreativo de Huelva	4
Espanyol	3	Sporting de Gijón	4
Real Sociedad	3	Almería	4
Real Valladolid	3	Cádiz	4
Real Zaragoza	3	Xerez	4
R.C.D. Mallorca	3	Gimnàstic Tarragona	4
Racing de Santander	3	Hércules	4
Málaga C.F.	3	Las Palmas	4

Tabla 2: Equipos que han jugado en Primera las 8 temporadas y su categoría
Fuente: Elaboración propia

dif. Categorías	visitante	empate	local
-3	0.0400	0.0000	0.9600
-2	0.0954	0.1731	0.7314
-1	0.2036	0.2385	0.5576
0	0.2521	0.2713	0.4766
1	0.3473	0.2873	0.3654
2	0.5053	0.2650	0.2297
3	0.6800	0.2400	0.0800

Tabla 3: Probabilidades estimadas para los distintos resultados en función de la diferencia de categoría entre los equipos que compiten
Fuente: Elaboración propia

La distribución de probabilidad sobre los tres resultados posibles se muestra en la Tabla 3. Con el fin de hacer intervenir también la categoría de los equipos, estimamos una distribución de probabilidad para cada diferencia de categoría entre el equipo local y el visitante. Siendo cuatro las categorías, esta diferencia toma los siete valores comprendidos entre -3 y 3.

4.1. Simulaciones Uniformes

En este primer análisis partimos de una situación de perfecto equilibrio entre todos los equipos. Bajo este supuesto, hemos obtenido 200 clasificaciones correspondientes a otras tantas temporadas, simulando para ello los resultados de los 380 partidos de cada temporada. A las 200 funciones resultantes hemos añadido las clasificaciones de las 8 temporadas reales.

El resultado de cada partido se elige al azar entre los tres posibles, asignado a cada equipo la puntuación que se derive, 3 puntos al vencedor o 1 punto a cada equipo en caso de empate. Obsérvese que bajo el supuesto de uniformidad, no importan ni los equipos que contienen ni sus categorías. En consecuencia, las simulaciones son válidas para cualquiera de las 8 temporadas, razón por la cual podemos analizarlas conjuntamente con las 200 simulaciones.

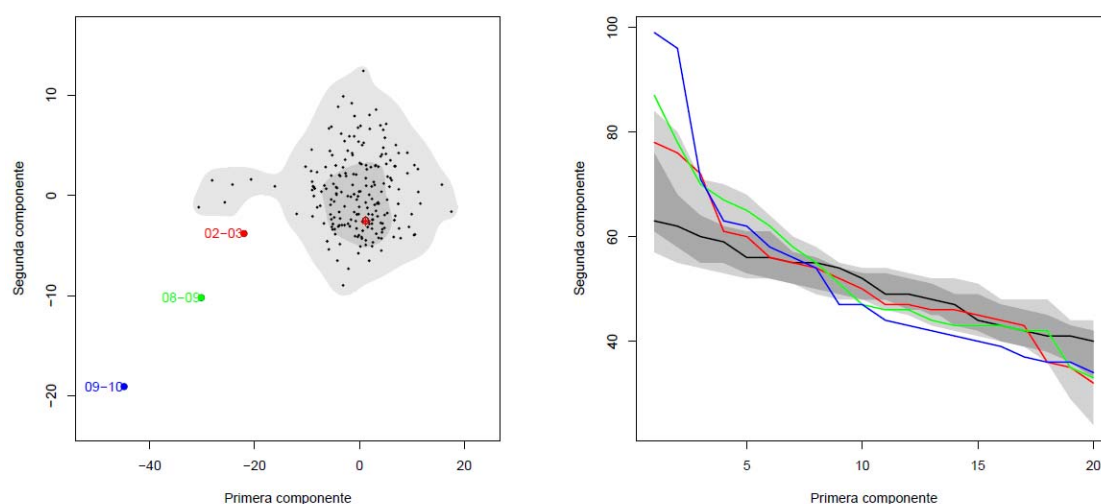


Figura 2: Diagramas bivariantes y funcional HDR bajo el supuesto de uniformidad
Fuente: Elaboración propia

La Figura 2 muestra los diagramas bivalente y funcional basados en la densidad correspondientes a las 208 funciones. De acuerdo con lo explicado en la Sección 2.2, el diagrama bivalente (izquierda) muestra como outliers las temporadas 2002-03, 2008-09 y 2009-10. La gráfica funcional (derecha), con el mismo código de colores, y la Tabla 1 nos ayudan a comprender la razón para ello. La dos últimas temporadas, particularmente la última, sobresalen del resto de funciones en los primeros valores. Se han alcanzado puntuaciones superiores o muy superiores al resto de temporadas. La temporada 2002-03 debe su condición de outlier a las puntuaciones obtenidas por el tercer y decimotavo clasificado. Los 72 puntos de aquél son la máxima puntuación alcanzada en esa posición a lo largo de las 8 temporadas examinadas y también de las 200 simuladas. Por contra, los 36 puntos del decimotavo es un

mínimo para esta posición. Obsérvese también que la temporada 2009-10 muestra puntuaciones muy bajas entre las posiciones novena a decimoctava, con la curva fuera de las zonas sombreadas.

Señalemos por último que la posición de los outliers en el diagrama bivalente señalan una tendencia al desequilibrio que se acentúa con los años. Los resultados de la temporada 2010-11, aunque parciales porque todavía no ha finalizado, parecen corroborarlo.

4.2. Simulaciones Ponderadas

Este segundo análisis asume de entrada que el equilibrio competitivo no es real, haciendo depender el resultado del partido de la diferencia de categoría entre el equipo local y el visitante. Se utilizan ahora las probabilidades recogidas en la Tabla 3. Como consecuencia de este planteamiento las temporadas no son intercambiables, como ocurría anteriormente, porque de una a otra se producen descensos y promociones y cambian los equipos y sus categorías. Cada temporada deberá ser analizada con las 200 simulaciones obtenidas a partir de las concretas diferencias de categorías de sus 380 partidos. En total serán ahora 201 funciones para cada temporada.

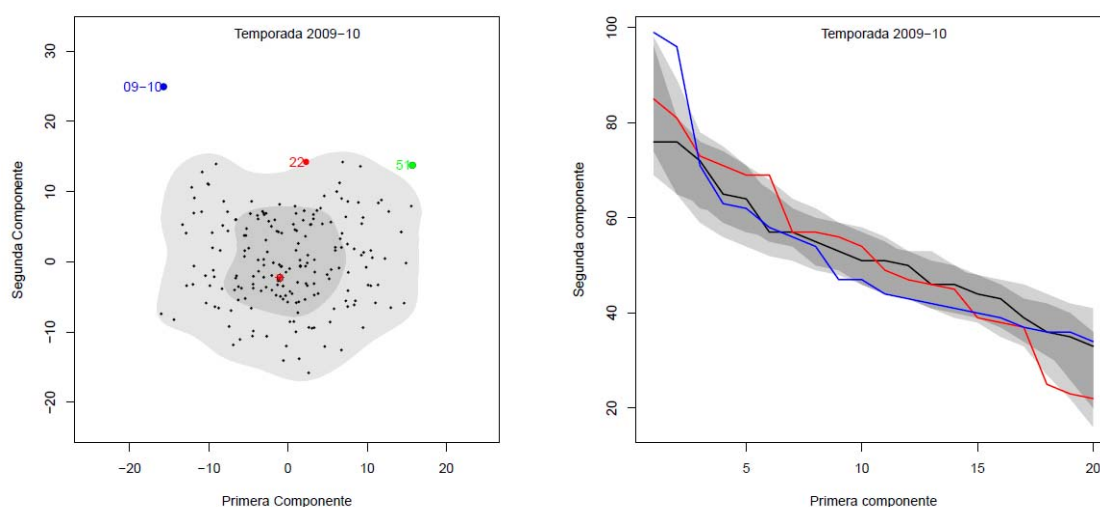


Figura 3: Diagramas bivariantes y funcional HDR bajo el supuesto de ponderación de la temporada 2009-10

Fuente: Elaboración propia

Como era de esperar, los outliers han desaparecido, pero no del todo. La clasificación de la última temporada, 2009-10, continua siendo un outliers y evidenciando un desequilibrio

mayor del que hemos supuesto de partida. La Figura 3 muestra los correspondientes gráficos. Aparecen en ella otros dos outliers correspondientes a la simulaciones 22 y 51 que carecen de interés.

Conviene recordar que en la temporada 2009-10 los dos primeros clasificados obtuvieron 99 y 96 puntos, respectivamente, con una distancia de 28 y 25 sobre el tercero. Se observa en la gráfica (derecha) cómo la correspondiente función sobresale claramente de las zonas sombreadas en sus valores iniciales. El patente desequilibrio que este hecho supone no parece ser anecdótico a tenor de lo que está ocurriendo en la temporada actual 2010-11. Cuando este artículo se escribe han transcurrido 23 jornadas, las tres primeras posiciones están ocupadas por los mismos equipos, y en el mismo orden, que las ocupaban en igual jornada de la temporada 2009-10. Entonces su puntuaciones fueron de 58, 56 y 46 puntos, respectivamente, en la actualidad son 62, 57 y 47. El desequilibrio no sólo se mantiene, sino que parece acrecentarse.

4.3. Análisis de las temporadas 1987-88 a 2009-10

El último es un análisis alternativo a los dos anteriores. Se trata ahora de estudiar las clasificaciones de las temporadas 1987-88 a 2009-10. La razón para ello es que en la temporada 1987-88 jugaron por primera vez el campeonato de Primera División 20 equipos. Desde entonces el número se ha mantenido, excepto en las temporadas 1996-97 y 1995-96 en las que se aumentó a 22. El análisis conjunto de todas ellas exige homogeneizar sus clasificaciones. Para ello hemos suprimido los dos últimos clasificados en las temporadas 1996-97 y 1995-96 y substraído al resto de equipos los puntos que obtuvieron en las confrontaciones con ellos. La Tabla 4 contiene las clasificaciones de las temporadas 1987-88 a 2001-02, que junto con las de la Tabla 1 completan las 23 temporadas analizadas.

La Figura 4 muestra los gráficos de los diagramas bivalente y funcional con un único outlier, la función correspondiente a la última temporada 2009-10. Las razones para ello son las mismas que se han explicado en los apartados precedentes.

Clasi ficación	7-88	8-89	9-90	0-91	1-92	2-93	3-94	4-95	5-96	6-97	7-98	8-99	9-00	0-01	1-02
1	0	7	8	2	8	3	1	8	8	4	4	9	9	0	5
2	3	0	3	4	7	1	8	1	4	2	5	8	4	3	8
3	7	7	4	6	7	6	5	1	2	8	3	6	4	1	6
4	3	5	0	0	7	7	4	4	4	5	3	5	3	3	4
5	5	8	9	0	0	9	9	9	3	9	0	4	2	3	0
6	4	1	1	5	8	9	7	7	0	5	0	3	1	9	9
7	3	7	9	5	6	0	4	1	9	5	0	1	3	7	4
8	1	3	4	3	3	7	3	8	8	4	9	0	3	6	3
9	2	1	6	9	9	6	9	2	2	0	5	7	2	0	3
10	0	0	0	0	9	6	1	1	9	8	3	4	1	9	3
11	7	7	1	7	0	3	8	0	6	7	0	9	0	6	9
12	6	7	8	1	7	7	8	9	5	3	9	8	8	3	8
13	5	8	6	7	6	5	5	7	3	2	8	6	7	3	7
14	3	3	1	9	6	1	0	8	2	2	5	5	7	3	7
15	4	2	2	3	3	4	4	8	2	0	5	2	6	2	3
16	7	5	8	6	4	3	2	5	1	9	5	0	5	2	3
17	0	7	7	4	9	2	0	4	0	8	4	9	5	2	2
18	9	8	4	6	5	8	8	6	0	7	0	6	2	1	0
19	8	3	7	6	4	7	4	3	5	7	9	4	8	9	8
20	1	9	5	1	7	6	4	5	4	6	3	7	8	9	7

Tabla 4: Clasificación de las temporadas 1987-88 a 2009-10
Fuente: Elaboración propia

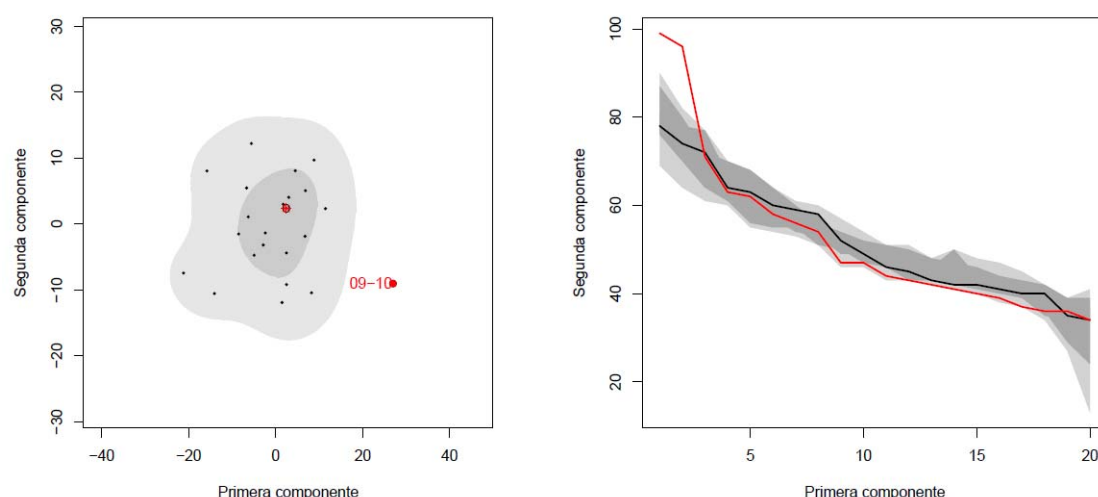


Figura 4: Diagramas bivariantes y funcional HDR con las clasificaciones de la temporada 1987-88 a 2009-10

Fuente: Elaboración propia

5. CONCLUSIONES

Los resultados obtenidos con cualquiera de los tres análisis corroboran el desequilibrio existente entre los equipos de la Liga española de fútbol de Primera División. Del segundo de los análisis, llevado a cabo admitiendo un desequilibrio inicial basado en las probabilidades de la Tabla 3, se deduce un desequilibrio aún mayor para la temporada 2009-10. Esta temporada destaca igualmente entre las 23 últimas, como muestra el resultado del tercer análisis.

No es casual que la última temporada analizada, 2009-10, sea un outlier en todos los análisis. Es la consolidación de una tendencia que muy probablemente se verá confirmada en las temporadas sucesivas. En efecto, lo que destaca en ella es la elevada puntuación alcanzada por los dos primeros equipos, a gran distancia de los restantes, como está ocurriendo ya en la temporada actual, 2010-11, aún no finalizada. No por casualidad dichos equipos son el FC Barcelona y el Real Madrid, equipos que han dominado la Liga española a lo largo de su historia.

Hay además razones económicas para que el desequilibrio continúe y se agrande. Según datos del *Deloitte Football Money League 2010*, ambos equipos encabezan la lista ingresos de todas las ligas europeas para la temporada 2008-09, con 401,4 millones de euros el Real Madrid y 365,9 millones de euros el FC Barcelona. Casi

la mitad de estos ingresos los obtienen de los derechos de televisión. También en la distribución de este dinero el desequilibrio es bien patente, sólo ellos dos se reparten un 35% del total, otros dos equipos, At. Madrid y Valencia CF, se reparten un 11%, distribuyéndose el 4% restante entre los otros 16, lo que les supone apenas un 1,7% de media. Es decir, entre aquellos y estos últimos la proporción es de 10 a 1. Esta desigualdad no se da en el resto de ligas europeas, en las que en el peor de los casos la ratio es de 2 a 1.

Una última consideración acerca del número de simulaciones, N , en los apartados 4.1 y 4.2. Los resultados se muestran robustos frente al valor de N . Con valores mayores, $N=500$, o menores, $N=100$, los únicos cambios que se observan son la aparición de distinto ruido, simulaciones que resultan ser outliers, y la inclusión como outliers de alguna de las 8 temporadas objeto de estudio; pero manteniéndose siempre las señaladas originalmente.

6. REFERENCIAS BIBLIOGRÁFICAS

- BUZZACCHI L., SZYMANSKI, S.y VALLETTI, T. (2003) Equality of Opportunity and Equality of Outcome: Open Leagues, Closed Leagues and Competitive Balance. *Journal of Industry, Competition and Trade*, **3**(3), 167-186.
- DELOITTE (2010) *Spanish Masters. Foot ball Money League*.
<http://www.deloitte.com/assets/Dcom-UnitedKingdom/Local%20Assets/Documents/Industries/Sports%20Business%20Group/UK%20SBG%20DFML2010.pdf>
- DOBSON, S. y GODDARD, J. (2001) *The Economics of football* . Cambridge University Press, Cambridge (UK).
- DOBSON, S.y GODDARD, J. (2004) Revenue divergence and competitive balance in a divisional sports league. *Scottish Journal of Political Economy* , **51** (3), 359-376.
- ECKARD, E. W. (2001) Free Agency, Competitive Balance and Diminishing returns to Pennant Contention *Economic Inquiry* , **39** (3), 430-443.
- FEBRERO, M., GALEANO, P. y GONZALEZ MANTEIGA, W. (2007) A functional analysis of NOx levels: location and scale estimation and outlier detection. *Computational Statistics*, **22**(3), 411-427.

- FILZMOSE, P., MARONNA, R. y WERNER, M. (2008) Outlier identification in high dimensions. *Computational Statistics and Data Analysis*, **52**(3), 1694-1711.
- HYNDMAN, R. J. (1996) Computing and graphing highest density regions. *The American Statistician*, **50**(2), 120-126.
- HYNDMAN, R. J. y ULLAH, M. S. (2007) Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics and Data Analysis*, **51**(10), 4942-4956.
- NEALE, W. (1964) The peculiar economics of professional sports. *Quarterly Journal of Economics*, **78**, 1-14.
- ROTTENBERG, S. (1956) The baseball players labor market. *Journal of Political Economy*, **64**(3), 242-258.
- ROUSSEUW, P. J., Ida Ruts, I. y Tukey. J. W. (1999) The Bagplot: A Bivariate Boxplot. *The American Statistician*, **53**(4), 382-387.
- SCULLY, G. W. (1989) *The Business of Major League Baseball*. The Chicago University Press, Chicago.
- SCULLY G. W. (1995) *The Market Structure of Sports*. The Chicago University Press, Chicago.
- SHANG, H. L. y HYNDMAN, R. J. (2010) *rainbow: Rainbow plots, bagplots and boxplots for functional data*. R package version 2.3.1. <http://CRAN.R-project.org/package=rainbow>
- SZYMANSKI, S. (2001) Income inequality, competitive imbalance and the attractiveness of team sports: some evidences and a natural experiment from English soccer. *Economic Journal*, **111**, F59-F84.
- TUCKEY J. W. (1975) Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians*, Vancouver, . 523-531.
- VROOMAN, J. (1995) A general theory of professional sports league. *Southern Economic Journal*, **61**,971-990.