

Información Importante

La Universidad de La Sabana informa que el(los) autor(es) ha(n) autorizado a usuarios internos y externos de la institución a consultar el contenido de este documento a través del Catálogo en línea de la Biblioteca y el Repositorio Institucional en la página Web de la Biblioteca, así como en las redes de información del país y del exterior con las cuales tenga convenio la Universidad de La Sabana.

Se permite la consulta a los usuarios interesados en el contenido de este documento para todos los usos que tengan finalidad académica, nunca para usos comerciales, siempre y cuando mediante la correspondiente cita bibliográfica se le de crédito al documento y a su autor.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, La Universidad de La Sabana informa que los derechos sobre los documentos son propiedad de los autores y tienen sobre su obra, entre otros, los derechos morales a que hacen referencia los mencionados artículos.

BIBLIOTECA OCTAVIO ARIZMENDI POSADA
UNIVERSIDAD DE LA SABANA
Chía - Cundinamarca

ESTRATEGIAS EN LA PREMIER LEAGUE: EVIDENCIA EMPIRICA

Econometría y decisiones tácticas en el futbol inglés

ABSTRACT

En este trabajo se presenta evidencia empírica del impacto de variables futbolísticas sobre el desempeño de los equipos para la Premier League. Variables como posesión, faltas, porcentaje de juego por bandas y por centro entre otras, se utilizan como centro de estudio para proporcionar información estadística que permita facilitar la toma de decisiones tácticas en el fútbol inglés.

José Fernando Contreras Yáñez

Cristian Muñoz Ramírez

Proyecto de grado

Universidad de La Sabana

Economía y Finanzas Internacionales

Contenido

Introducción.....	2
Datos.....	3
Metodología	5
Pruebas	7
Significancia de las variables fecha	7
La prueba F.....	7
Errores robustos a heterocedasticidad.	7
Prueba de Hausman.	7
Selección del modelo	9
Resultados	12
Conclusiones.....	17
Referencias	18

Introducción

Desde su introducción, por Henry Ludwell Moore durante la segunda década del siglo XX, la econometría moderna se ha basado en la estadística para intentar explicar fenómenos que en un análisis simple no son sencillos de comprender. Esta disciplina se ha usado en diferentes áreas como economía, política, sociología, biología, medicina y deportes entre otros, para predecir, explicar y observar efectos causales desde variables independientes hacia una variable dependiente.

Puntualmente, en el deporte, autores como Alonso, Ruesga y Vicens (1991), Andreff y Weber (1995) y Heinemann (1998), entre otros, han tocado temas como: El impacto del deporte en la economía, impacto de la economía en el deporte, mercado de bienes y servicios deportivos, práctica deportiva como potenciador de la productividad de los trabajadores y éxito en diferentes eventos deportivos.

En el fútbol masculino existen 5 elites de ligas nacionales, estas son: España (Liga BBVA), Inglaterra (Premier League), Alemania (Bundesliga), Italia (Serie A) y Francia (Ligue 1). La Premier League es la liga más competitiva entre estas. En los últimos seis años, sólo dos equipos han logrado ganar la Liga BBVA, en Italia y Alemania sucede lo mismo, por otra parte, en Francia tres equipos han logrado el título al igual que en la Premier League. No obstante, para el final de la temporada 2012-2013, la diferencia en la Bundesliga entre el campeón y el subcampeón fue de 25 puntos mientras que para la Premier League fue de 11; entre campeón y mitad de tabla, para la Liga BBVA fue de 54 puntos mientras que para la premier fue de 45; y entre campeón y colista, para la Bundesliga fue de 70 puntos y para la Premier League fue de 64 (Siendo la Bundesliga una liga en la cual se juegan cuatro partidos menos por temporada que en las demás ligas). Adicionalmente, es común ver que en la Premier League un equipo de la cima de la tabla (p.e Manchester United) pierda o empate con un colero (p.e Leicester City), mientras que en las demás ligas un evento como este tiene poca ocurrencia.

Siendo la Premier League la liga más competitiva, no existe una estrategia o táctica a simple vista que lleve a un equipo a marcar goles o ganar. De esta manera, no es claro si tener la pelota tenga un efecto positivo o negativo al marcar goles, al igual que jugar al contra golpe, por las bandas, el centro o pausando más el juego por medio de faltas. A esto se le suma la incertidumbre frente a cómo impacta el ser local, hacer pases cortos, pases al vacío, intercepciones y tackles en el resultado de un partido. Consideramos necesario saber que variables repercuten en el resultado y en que magnitud lo afectan. Esto con el fin de brindar un soporte econométrico para la toma de decisiones tácticas en un equipo.

En la actualidad, no existe literatura que abarque con profundidad el comportamiento de la Premier League ni la estrategia más común a implementar a lo largo de cada temporada.

Datos

La base de datos a utilizar comprende datos de la primera división del fútbol inglés (Premier League) para la temporada 2013-2014. Esta liga está conformada por 20 equipos y cada uno de estos juega 38 partidos, de la fecha 1 a la fecha 19 es la primera ronda y de la 20 a la 38 se juega la segunda ronda. Es decir, cada equipo juega 2 veces contra el mismo oponente, uno en condición de local y otro en condición de visitante. El orden de los encuentros se establece de manera aleatoria desde antes que comience la temporada.

Para cada partido de la temporada y para cada equipo se tiene información referente a los goles a favor, goles en contra, puntos conseguidos, tiros, tiros al arco, pases totales, pases cortos, pases al vacío, juego por izquierda-derecha-centro (medido en porcentaje), fueros de lugar, faltas, asistencia, local-visitante, posesión (medida en porcentaje), tackles, intercepciones y valor de mercado. La recolección de estos datos fue a través de varias fuentes estadísticas de fútbol¹. Consideramos que esta temporada es de gran importancia por dos eventos técnicos ocurridos en dos equipos históricos de esta liga. El primer evento a destacar es el regreso al Chelsea de José Mourinho quien ganó títulos en la temporada 2004-2005 y 2005-2006. El segundo evento es la partida de Sir Alex Ferguson tras 27 años a cargo del Manchester United. El cambio técnico de los dos equipos más ganadores de la Premier League hace que diferentes tácticas surgen y permitan un análisis sobre una base de datos que elimina la hegemonía de estos dos equipos.

Se usa una base de datos panel balanceada con 760 observaciones (38 fechas por cada uno de los 20 equipos), además se tienen 37 variables dummy para aislar el efecto fecha que se da por clima, desgaste o presión psicológica. Por ejemplo, en el periodo diciembre-enero, Inglaterra está en temporada de invierno y los juegos en premier son más consecutivos hasta el punto de jugar dos partidos por semana. Cada equipo se le asigna un código del 1 al 20. Con respecto a las variables juego por izquierda-centro-derecha (medido en porcentaje) sólo se pueden incluir dos ya que al incluir las tres se generaría un problema de correlación.

Para aclarar qué comprende cada una de las variables, se van a definir a continuación.

- Gf: Son los goles a favor de cada equipo, en cada partido, esta variable no puede tomar valores negativos y es una variable discreta.
- Ag: Son los goles en contra de cada equipo, en cada partido, esta variable no puede tomar valores negativos y es una variable discreta.
- Gd: Es la diferencia entre Gf y Ag, esta variable puede tomar valores negativos y es discreta.
- Shots: Son los tiros con intención de gol ejecutados por cada equipo en un partido que no necesariamente tuvieron dirección al arco, esta variable no puede tomar valores negativos

¹Las fuentes utilizadas fueron www.whoscored.com y www.transfermarket.com.

y es una variable discreta.

- **ShotsOT:** Son los tiros con intención de gol y efectivamente llegaron al arco, esta variable no puede tener valores negativos y es una variable discreta.
- **Shortpasses:** Son los pases cortos efectuados por un equipo en un partido determinado, esta variable no puede tomar valores negativos y es una variable discreta.
- **Throughballs:** Son pases al vacío, esto quiere decir que son pases con intención de ataque a un espacio donde no existía un jugador.
- **Leftside:** Dada la posesión, esta variable mide que tanto un equipo jugó por su banda izquierda en un partido. Esta variable tiene un rango de 0 a 1, no puede tomar valores negativos y es continua.
- **Middleside:** Dada la posesión, esta variable mide que tanto un equipo jugó por su carril central en un partido. Esta variable tiene un rango de 0 a 1, no puede tomar valores negativos y es continua.
- **Rightside:** Dada la posesión, esta variable mide que tanto un equipo jugó por su carril central en un partido. Esta variable tiene un rango de 0 a 1, no puede tomar valores negativos y es continua.
- **Offsides:** Esta variable mide las veces que un equipo incurrió en fuera de lugar por partido. Esta variable no puede tomar valores negativos y es discreta.
- **Fouls:** Esta variable mide la cantidad de faltas cometidas por un equipo en cada partido. Es válido aclarar que esta variable no mide las faltas que el equipo recibió si no por el contrario las que cometió.
- **Local:** Es una variable dummy, que toma valor de 1 si el equipo jugó de local ese partido y valor de 0 si el equipo jugó de visitante ese partido.
- **Posesión:** Esta variable mide la posesión de un equipo en un partido, tiene rango de 0 a 1 y es una variable continua.
- **Tackles:** Esta variable mide la cantidad de veces que un equipo le quitó el balón a su contrincante implicando un contacto físico directo sin causar falta. Es una variable discreta.
- **Interception:** Esta variable mide la cantidad de veces que un equipo le quito el balón a su contrincante evitando el contacto físico. Es una variable discreta.
- **Valordemercado:** Esta variable esta medida en millones de libras, es constante para cada equipo a lo largo de la temporada.

Los datos no fueron modificados ni procesados ya que tenían la estructura deseada y necesaria para ser ingresados en el programa estadístico.

Metodología

Ecuaciones y modelos.

Las ecuaciones a estimar son:

$$(1) gf = \alpha_0 + \sum_{i=1}^{37} \alpha_i fecha_i + \beta_1 ShotsOT_{it} + \beta_2 Shortpasses_{it} + \beta_3 Throughballs_{it} + \beta_4 Leftside_{it} + \beta_5 Middleside_{it} + \beta_6 Offsides_{it} + \beta_7 Fouls_{it} + \beta_8 Local_{it} + \beta_9 Posesion_{it} + \beta_{10} Tackles_{it} + \beta_{11} Interceptions_{it} + \beta_{12} Valordemercado_{it} + u_{it}$$

$$(2) gd = \alpha_0 + \sum_{i=1}^{37} \alpha_i fecha_i + \beta_1 ShotsOT_{it} + \beta_2 Shortpasses_{it} + \beta_3 Throughballs_{it} + \beta_4 Leftside_{it} + \beta_5 Middleside_{it} + \beta_6 Offsides_{it} + \beta_7 Fouls_{it} + \beta_8 Local_{it} + \beta_9 Posesion_{it} + \beta_{10} Tackles_{it} + \beta_{11} Interceptions_{it} + \beta_{12} Valordemercado_{it} + u_{it}$$

Para ser más precisos en la estimación del modelo, no se incluyen las variables totalpasses y shots ya que se tienen variables más precisas como shortpasses, shotsOT y Throughballs. Adicionalmente, si se incluyera Totalpasses y shots, existiría correlación entre las variables. En cuanto a la elección entre las variables leftside, middleside y rightside, se corrió una regresión por MCO y se probó que leftside y middleside si son estadísticamente significativas mientras que rightside no (Ver anexo o cuadro tal).

En el análisis de nuestra base de datos (panel), se consideraron tres modelos que se ajustan a las necesidades de este trabajo. A continuación se explicaran cada uno de estos métodos.

- Mínimos cuadrados agrupados: Este método analiza los datos tipo panel omitiendo las dimensiones de espacio y tiempo en los datos agrupados. De esta manera, la regresión estimada es una de MCO usual.

- Efectos fijos: En los modelos para datos panel, se considera un factor a_i el cual se conoce como efecto inobservable. Esta variable captura todos los factores que son inobservables, constantes en el tiempo y que influyen en la variable dependiente. El método de efectos fijos, permite que este error inobservable se correlacione con las variables explicativas.

- Efectos aleatorios: Este modelo tiene como objetivo eliminar a_i porque se considera que está correlacionado con una o más de las variables explicativas.

Estos tres modelos se van a aplicar en las dos ecuaciones mencionadas anteriormente, es de resaltar que factor diferenciador de las dos ecuaciones es su variable dependiente. Para la primera ecuación, se estima el éxito ofensivo del equipo representado por los goles anotados por este equipo. En la segunda ecuación se busca estimar el desempeño general del equipo, si la variable gd es positiva el equipo ganó, por el contrario si es negativa el equipo perdió y si toma

el valor de 0 se observó un empate en el partido. Por lo tanto esta última ecuación busca estimar que factores inciden más a la hora de ganar un partido.

Pruebas

Significancia de las variables fecha

Anteriormente se mencionó que se usaron 37 dummies de fecha con el fin de aislar posibles efectos causados por las fechas. Se realiza una prueba F por MCO representada en la siguiente ecuación:

$$gf = \alpha_0 + \sum_{i=1}^{37} \alpha_i fecha_i + u_{it}$$
$$gd = \alpha_0 + \sum_{i=1}^{37} \alpha_i fecha_i + u_{it}$$

La prueba F

Tiene como H_0 que todos los α_i son iguales a 0, es decir, no son conjuntamente significativos. Por lo tanto H_1 es que si son conjuntamente significativas. Como medida adicional se analizará cada uno de los estadísticos t de cada una de las dummies para probar su significancia individual.

Errores robustos a heterocedasticidad.

La heterocedasticidad consiste en que la varianza no es constante en la muestra, esto genera un problema en el cálculo de los errores estándar para cada uno de los estimadores. Como se sabe, el estadístico t ayuda a medir la significancia de los estimadores (Betas) por medio de la siguiente ecuación:

$$t = \frac{\text{Estimación}}{\text{error estándar}}$$

Si el error estándar no es preciso, entonces el estadístico t tampoco lo será, lo cual genera problemas en el análisis de la significancia de los estimadores. Por esta razón, ambas ecuaciones para cada uno de los modelos será corrida con errores estándar robustos a heterocedasticidad.

Prueba de Hausman.

La prueba de Hausman permite determinar si hay diferencias sustanciales o sistemáticas entre los estimadores de dos modelos diferentes. La H_0 es que no existe diferencias sistemáticas entre los coeficientes, es decir, se es indiferente en usar cualquiera de los dos modelos. La H_1 es que si

existe diferencias sistemáticas entre los coeficientes estimados por los dos modelos, por ejemplo si se usa la prueba de Hausman entre efectos aleatorios y efectos fijos y la H_0 se rechaza, entonces es aconsejable usar efectos fijos.

Esta prueba se utilizará dos veces, la primera es para medir diferencia entre efectos fijos y mínimos cuadrados agrupados y la segunda para medir diferencias entre efectos fijos y efectos aleatorios.

Selección del modelo

A continuación se muestra el resultado de la prueba de hausman entre efectos fijos y MCO agrupados:

	— Coefficients —			
	(b) FIXEDEFFECTS	(B) MCOA	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
shotsot	.2524999	.2635671	-.0110672	.0027605
shortpasses	.001311	.002782	-.0014711	.0003567
throughballs	.0791005	.0941548	-.0150543	.0124455
middle	-2.288587	-2.405638	.1170508	.3287943
leftside	1.351343	1.432533	-.0811902	.2939717
offsides	-.0253984	-.0137841	-.0116143	.0031199
fouls	-.001527	.0037709	-.0052978	.0024756
local	.1590365	.1154181	.0436184	.
possession	-2.731064	-2.932955	.2018907	.1767137
tackles	.003981	.0094804	-.0054993	.001813
intercepti~s	-.0043107	-.0053803	.0010696	.0031678

b = consistent under Ho and Ha; obtained from xtreg
 B = inconsistent under Ha, efficient under Ho; obtained from regress

Test: Ho: difference in coefficients not systematic

```
chi2(48) = (b-B)'[(V_b-V_B)^(-1)](b-B)
          =      15.95
Prob>chi2 =      1.0000
(V_b-V_B is not positive definite)
```

Como H_0 es que no hay diferencias sistemáticas entre los coeficientes y que la probabilidad es de 100%, entonces no se rechaza la nula. Es decir, que los coeficientes estimados entre efectos fijos y MCO agrupados no tienen diferencias sistemáticas de tal manera que es indiferente escoger entre los dos modelos mencionados.

Posteriormente se hace la misma prueba pero para efectos fijos y efectos aleatorios, el resultado se muestra a continuación.

	— Coefficients —			sqrt(diag(V_b-V_B)) S.E.
	(b) FIXEDEFFECTS	(B) RANDOMEFFE~S	(b-B) Difference	
shotsot	.2524999	.259188	-.006688	.0027196
shortpasses	.001311	.0022447	-.0009338	.0002913
throughballs	.0791005	.0917567	-.0126562	.0098743
middle	-2.288587	-2.367463	.0788751	.2511804
leftside	1.351343	1.387636	-.0362927	.2260101
offsides	-.0253984	-.0193456	-.0060528	.0031515
fouls	-.001527	.0015844	-.0031114	.0021539
local	.1590365	.1307625	.0282741	.0070224
possession	-2.731064	-2.849726	.1186624	.1461189
tackles	.003981	.0074663	-.0034852	.0015571
interceptions	-.0043107	-.0044222	.0001115	.0024482

b = consistent under Ho and Ha; obtained from xtreg
 B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(48) = (b-B)'[(V_b-V_B)^(-1)](b-B)
 = 15.67
 Prob>chi2 = 1.0000
 (V_b-V_B is not positive definite)

Al igual que la prueba anterior, la probabilidad es de 100%, lo que significa que no se rechaza la hipótesis nula. Al ser indiferente los coeficientes entre efectos fijos y efectos aleatorios se decide es coger efectos fijos.

Una vez seleccionado efectos fijos como modelo, ahora se probará la significancia conjunta de las dummies de fecha:

Fixed-effects (within) regression	Number of obs	=	760
Group variable: teamcod	Number of groups	=	20
R-sq: within = 0.0397	Obs per group: min	=	38
between = .	avg	=	38.0
overall = 0.0332	max	=	38
	F(37, 703)	=	0.79
corr(u_i, Xb) = 0.0000	Prob > F	=	0.8165

La probabilidad es de 0.83, esto quiere decir que no se rechaza la H_0 . Según la prueba F, todos los betas estimados para las fechas son iguales a cero, aun así, se incluirán en el modelo porque se considera que las fechas son un control importante para las regresiones a estimar. Al cambiar la variable dependiente gf por gd (diferencia de goles), la conclusión es la misma.

Fixed-effects (within) regression	Number of obs	=	760
Group variable: teamcod	Number of groups	=	20
R-sq: within = 0.0000	Obs per group: min =		38
between = .	avg =		38.0
overall = 0.0000	max =		38
	F(37,703)	=	0.00
corr(u_i, Xb) = -0.0000	Prob > F	=	1.0000

Resultados

A continuación se muestra los resultados de la regresión sobre gf (goles a favor):

```

Fixed-effects (within) regression      Number of obs   =    760
Group variable: teamcod                Number of groups =    20

R-sq:  within = 0.3224                  Obs per group:  min =    38
      between = 0.7781                    avg   =   38.0
      overall = 0.3838                    max   =    38

                                          F(19,19)       =
corr(u_i, Xb) = 0.2232                  Prob > F        =

```

(Std. Err. adjusted for 20 clusters in teamcod)

gf	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
shotsot	.2524999	.0326351	7.74	0.000	.1841939	.320806
shortpasses	.001311	.0007801	1.68	0.109	-.0003218	.0029437
throughballs	.0791005	.0260945	3.03	0.007	.0244841	.1337169
middle	-2.288587	.8490157	-2.70	0.014	-4.065598	-.5115772
leftside	1.351343	.6353755	2.13	0.047	.0214868	2.681199
offsides	-.0253984	.0250326	-1.01	0.323	-.0777923	.0269954
fouls	-.001527	.0114753	-0.13	0.896	-.0255451	.0224912
local	.1590365	.0808785	1.97	0.064	-.0102441	.3283172
possession	-2.731064	.6735574	-4.05	0.001	-4.140836	-1.321292
tackles	.003981	.0094473	0.42	0.678	-.0157924	.0237545
interceptions	-.0043107	.0085475	-0.50	0.620	-.0222008	.0135793
sigma_u	.30480371					
sigma_e	1.0302224					
rho	.08048881	(fraction of variance due to u_i)				

Tomando en cuenta los coeficientes y la significancia de las variables al 10% tenemos que:

- Fouls: No es significativa. Si un equipo aumenta o disminuye su número de faltas cometidas no afectaría la cantidad de goles que marca dicho equipo.

- Offsides: El valor p es de 0,323 y por lo tanto no tiene significancia. A nuestro criterio, esta variable mide que tan sobre la línea defensiva del oponente juega el equipo. Ya que esta variable no tiene significancia, un equipo que juega sobre el límite del oponente no tiene ningún efecto sobre el marcador propio de goles.
- Tackles: Al no ser significativa esta variable para la muestra que tenemos, quitar el balón habiendo contacto físico no tiene incidencia a la hora de marcar goles. Si a esto se suma que las faltas tampoco son significativas entonces encontramos que un equipo violento o pasivo no afecta el marcador, aun así, esto excluye eventos de tarjeta amarilla o roja.
- Interceptions: Al igual que las anteriores, no hay significancia de esta variable. Esto quiere decir que robar la pelota del equipo contrario sin contacto físico no tiene efecto a la hora de marcar goles. De esta manera variables que tienen en cuenta la recuperación de la pelota no juegan un papel importante en los goles a favor.
- Shotsot: Esta variable es significativa y su coeficiente es positivo, lo que quiere decir que, además de la relación clara entre tiros al arco y goles marcados, un equipo necesita en promedio 4 tiros al arco para marcar un gol en la Premier. Además, esta variable tiene una fuerza de significancia bastante alta pues su valor p es de cero.
- Throughballs: Hay una relación estadística positiva entre esta variable y los goles a favor. Lo que indica que un equipo que tiende a jugar verticalmente o con pases a las espaldas de los jugadores defensivos tiene más oportunidad de marcar gol aunque no es muy grande.
- Middle: A la hora de analizar porque lado es más efectivo marcar goles, jugar por la mitad es una variable significativa con un coeficiente negativo, es decir, jugar por el centro no es una manera adecuada para marcar goles.
- Leftside: Jugar por izquierda resulta estadísticamente importante. Por lo tanto, jugar por las bandas o desbordar por los costados es una herramienta eficiente para marcar goles teniendo en cuenta que la liga Premier es una liga rápida y que tiende a jugar a contragolpe.
- Local: Si bien es cierto que esta variable es significativa, contrario a lo esperado, su coeficiente no es relevante porque es 0,15, es decir, el hecho de que el equipo juegue de local tendría una ventaja prepartido de 0,15 goles. En términos prácticos tendría poca incidencia en marcar goles. Pero, de nuevo, es relevante.
- Posesion: El valor P es de 0.001 y el signo del coeficiente es negativo. Esta variable es significativa pero el valor negativo muestra que una alta posesión del balón incide negativamente en los goles marcados por un equipo.
- Shortpasses: La significancia de esta variable es inferior al 90% pero en un partido de fútbol es claro que debe existir una relación entre pases cortos y el desempeño de un

equipo. Así que, suponiendo que esta variable es significativa, la incidencia de los pases cortos sobre los goles a favor es positiva pero pequeña.

La siguiente tabla muestra los resultados de la regresión sobre gd (goles diferencia).

```

Fixed-effects (within) regression           Number of obs   =       760
Group variable: teamcod                    Number of groups =        20

R-sq:  within = 0.2250                     Obs per group:  min =       38
        between = 0.6721                    avg =           38.0
        overall = 0.2805                    max =           38

                                           F(19,19)       =       .
corr(u_i, Xb) = 0.2028                     Prob > F        =       .

```

(Std. Err. adjusted for 20 clusters in teamcod)

gd	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
shotsot	.3161987	.0445289	7.10	0.000	.2229986	.4093988
shortpasses	-.0003793	.0017161	-0.22	0.827	-.0039712	.0032126
throughballs	.1123846	.0390714	2.88	0.010	.0306071	.1941621
middle	-5.464064	1.344118	-4.07	0.001	-8.277335	-2.650794
leftside	-.2768182	.9026105	-0.31	0.762	-2.166004	1.612367
offsides	-.0152243	.0404312	-0.38	0.711	-.0998478	.0693991
fouls	.0015722	.020638	0.08	0.940	-.0416238	.0447681
local	.4353195	.1323434	3.29	0.004	.1583216	.7123174
possession	-.6506459	1.619349	-0.40	0.692	-4.039983	2.738692
tackles	.0002305	.0134435	0.02	0.986	-.027907	.028368
interceptions	-.0075045	.013948	-0.54	0.597	-.036698	.0216891
sigma_u	.5244509					
sigma_e	1.6572171					
rho	.09103302	(fraction of variance due to u_i)				

Aunque los resultados son similares, el comportamiento de algunos coeficientes difiere ya que la variable dependiente se cambia. El principal de estos es el R², para el modelo anterior era de 0.38 y para este modelo es de 0.28, por lo tanto el poder explicativo es menor. La interpretación de las variables independientes es:

- Shortpasses: Esta variable pierde significancia. El valor p es de 0.827. La cantidad de pases cortos que un equipo hace durante un partido no tiene relación con la diferencia de goles.
- Leftside: Al pasar de un modelo a otro, esta variable pierde su significancia. En este caso quiere decir que el juego por izquierda no afecta el ganar o perder.
- Offsides: Al igual que en la tabla anterior, offsides, no afecta la variable dependiente.
- Fouls: Esta variable no es significativa. La cantidad de faltas cometidas por un equipo no cambia el resultado de un equipo.
- Tackles: Robar o recuperar el balón por medio de tackles no tiene significancia.
- Intercepciones: No hay significancia en esta variable. Si tenemos en cuenta este resultado, el de fouls y el de tackles, al igual que en el modelo anterior, un equipo violento o pacífico no afecta el resultado del partido.
- ShotsOT: Hasta en un 1%, esta variable sigue siendo significativa. A diferencia del modelo anterior, el coeficiente es mayor, aumenta de 0.25 a 0.31. Los tiros al arco tienen un efecto mayor en los goles de diferencia que en los goles a favor.
- Posesion: Otra diferencia importante entre modelos es la pérdida de la significancia en esta variable. Para este modelo, la posesión no tiene importancia. Tener más tiempo el balón o tenerlo por menos tiempo no afecta el marcador final de un encuentro.
- Throughballs: El coeficiente es positivo y es significativo. Si hay más pases al vacío, entonces habrá mayor diferencia de goles.
- Local: Cuando la variable dependiente es gd, el coeficiente de local es más alto. Por lo tanto, el ser local afecta más el marcador final del partido que el marcar goles por parte de un equipo.
- Middle: Sigue siendo significativo y su coeficiente sigue siendo negativo, es decir, jugar por el centro o por la banda central afecta negativamente el resultado final de un partido.

Cuadro comparativo:

Dependiente	Gf	Gd	
Independiente			¿Cambia?
Shotsot	Significativo (+)	Significativo (+)	No
Shortpasses	Significativo (+)	No significativo	Si
Throughballs	Significativo (+)	Significativo (+)	No
Middle	Significativo (-)	Significativo (-)	No
Leftside	Significativo (+)	No significativo	Si
Offsides	No significativo	No significativo	No
Fouls	No significativo	No significativo	No
Local	Significativo (+)	Significativo (+)	No
Possesion	Significativo (-)	No significativo	Si
Tackles	No significativo	No significativo	No
Interceptions	No significativo	No significativo	No

Conclusiones

Para brindar un soporte econométrico a las decisiones tácticas de un equipo, este trabajo se centró en la econometría como herramienta para buscar una respuesta. Si bien es cierto que obtener una muestra lo suficientemente grande como para generalizar el impacto de las variables utilizadas en este trabajo sobre el fútbol a nivel mundial es realmente un trabajo de difícil realización. Buscar una respuesta para una liga de fútbol profesional en particular puede facilitar esta tarea.

En cada país y en cada liga las tácticas y la formación de los jugadores es diferente, por lo tanto la diversidad del fútbol es apasionante, encontrar patrones absolutos sería arbitrario e impreciso. Aunque teóricamente no existe literatura que soporte matemáticamente estrategias tácticas de fútbol ya que pueden ser complicadas de realizar y poco precisas. Este trabajo aporta evidencia empírica que permite obtener resultados útiles para un grupo poblacional en particular con una base de datos actual.

Abarcando información sobre la Premier League, a groso modo se puede obtener que las variables que este trabajo comprende, no tienen el mismo impacto cuando se trata de marcar goles que cuando se habla de ganar o perder el partido. Si suponemos que los resultados son constantes en el tiempo, entonces hay resultados de gran importancia que deberían ser tenidos en cuenta en las decisiones tácticas de un equipo.

Las estimaciones obtenidas muestran que: para aumentar los goles a favor los tiros al arco, los pases cortos, los pases al vacío y el juego por los costados deben aumentar. Mientras que el juego por la banda central y la posesión deben disminuir para aumentar los goles a favor. Cuando cambiamos la variable dependiente por diferencia de goles (ganar, perder o empatar), los tiros al arco y los balones al vacío deben aumentar para que los goles de diferencia aumenten y el juego por la banda central debería disminuir. Hay que tener en cuenta que cuando se cambia de variable dependiente los pases cortos, el juego por el costado izquierdo y la posesión dejan de tener relevancia. Jugar de local sí da una ventaja considerable.

Sin dejar a un lado la complejidad del fútbol, se encuentra que si hay variables generales que expliquen en cierta parte el comportamiento de un partido. Las variables utilizadas en los modelos explican entre un 28% y 38% el resultado en goles a favor y en diferencia de goles.

Referencias

Giraud, R. (2014), Favoritism and social pressure: Evidence from italian soccer referees, Pontificia Universidad Católica de Chile, Instituto de Economía

Tovar, J. (2014), Números Redondos: Leyendas y estadísticas del futbol mundial, GrijalboPenguin Random House Grupo Editorial, S.A.S

Jan, Hatius. Jari, Stehn (2014). A statistical model of the 2014 world cup, Goldman Sachs global investment research.

Recopilación de datos: www.whoscored.com [consultado del 16 de Febrero al 11 de Marzo de 2015]

Recopilación de datos: www.transfermarket.com [consultado del 16 de Febrero al 11 de Marzo de 2015]