

Football betting odds

Yll Mehmeti, Hugo Met and Kevin Turatsinze

ENAC, IENAC15 SITA-PREV

April 7, 2017

1 Introduction

- Context
- Dependant variables
- Explanatory variables

2 First model : simultaneous

- Home Win odd
- Away win odd
- Draw odd
- Tests

3 Second model : hybrid

- Home Win odd using Draw and Away fitted odds

4 Third model : sequential

- Draw odd using Away fitted odds
- Home odd using the new Draw fitted odds

5 Conclusion

How to bet ?

★ Ligue 1, France

LOSC Lille Métropole - OGC Nice

07/04/2017, 20:45

P G G E G	Derniers 5 matchs	G E E G G
E E E E P	Face à face	E E E E G
1,40	Buts/Points (5 matchs)	1,60

Tous **Princi...** Handi... Nomb... Résul... Plus... ▾

1 X 2

★ LOSC Lille Métropole	2.35	X	3.10	★ OGC Nice	3.10
------------------------	------	---	------	------------	------

Table – A game on BWin website

Our sample

Teams

$$p = 7 \text{ teams}$$

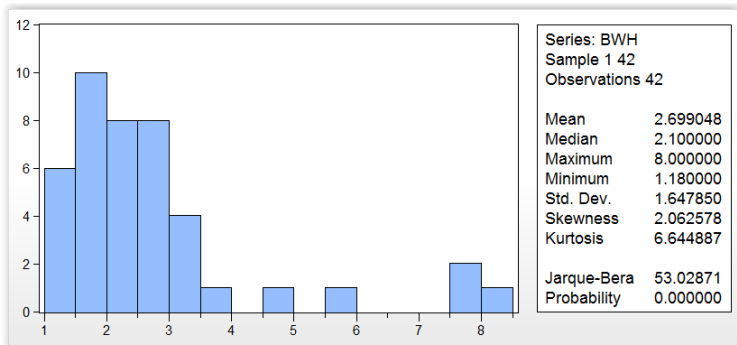
Caen, Lyon, Marseille, Monaco, Paris, Saint-Etienne, Toulouse

Number of observations

$$n = 2 * \times \sum_{k=1}^{p-1} k = 2 \times \frac{p(p-1)}{2} = p(p-1) = 7 \times 6 = 42$$

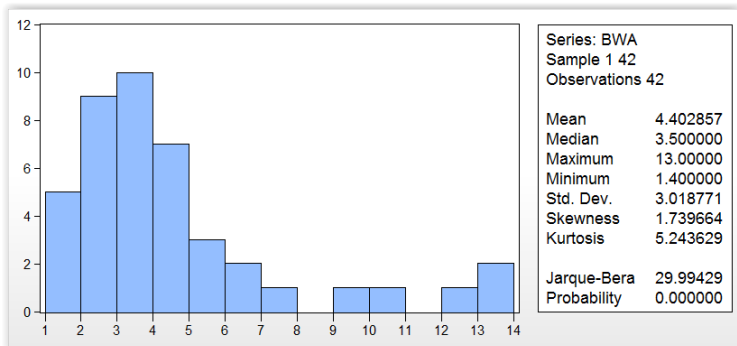
So we focused on 42 games of the 2015-2016 Ligue 1 season.

BWin odds : BWH



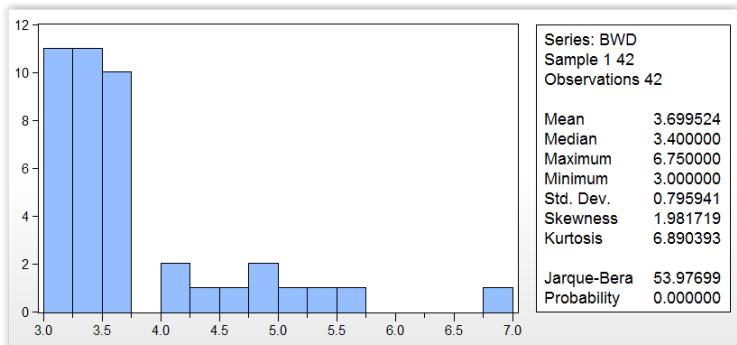
Graph – Histogram and stats for home odds.

BWin odds : BWA



Graph – Histogram and stats for away odds.

BWin odds : BWD



Graph – Histogram and stats for draw odds.

Delta variables

Formula used

$$\Delta = \Delta_{home} - \Delta_{away}$$

Delta variables

DELTA FIFA, DELTA FORM, DELTA POINTS, DELTA RANK, DELTA AFTER, DELTA BEFORE

LAST10HOME / LAST10DRAW / LAST10AWAY

Formula used

$$Last10(Home/Draw/Away) = \sum_{k=1}^{10} \frac{k}{10} \times \delta_{k,outcome}$$

where outcome = win of the current home team / draw / win of the current away team.

Other variables

Strictness of the referee

$$\textit{Strictness} = \frac{N_{\textit{YellowCards}} + 4 \times N_{\textit{RedCards}}}{N_{\textit{games}}}$$

Other variables

FIELDGRADE, TEMPERATURE, PACKING, DISTANCE

Simultaneous model

\widehat{BWH}

|

\widehat{BWD}

|

\widehat{BWA}

Home Win odd : estimation

Number of observations and variables

$n = 42$ and $k = 13$

The linear equation of the model

$$\begin{aligned} \widehat{BWH} = & \widehat{\beta}_0 + \widehat{\beta}_1 DELTAFIFA + \widehat{\beta}_2 DELTAFORM + \widehat{\beta}_3 DELTAPOINTS \\ & + \widehat{\beta}_4 DELTARANK + \widehat{\beta}_5 DELTAAFTER + \widehat{\beta}_6 DELTABEFORE \\ & + \widehat{\beta}_7 LAST10HOME + \widehat{\beta}_8 STRICTNESS + \widehat{\beta}_9 FIELDGRADE \\ & + \widehat{\beta}_{10} TEMPERATURE + \widehat{\beta}_{11} PACKING + \widehat{\beta}_{12} DISTANCE \end{aligned}$$

Home Win odd : marginal effects

Variable	Expected effect	Real effect	Significant
DeltaFifa	-	-	Yes
DeltaForm	-	-	Yes
DeltaPoints	-	-	Yes
DeltaRank	+	-	Yes
DeltaAfter	-	-	Yes
DeltaBefore	-	-	No
Last10Home	-	-	Yes
Strictness	-	-	No
FieldGrade	?	+	Yes
Temperature	?	0	No
Packing	-	0	No
Distance	-	0	No

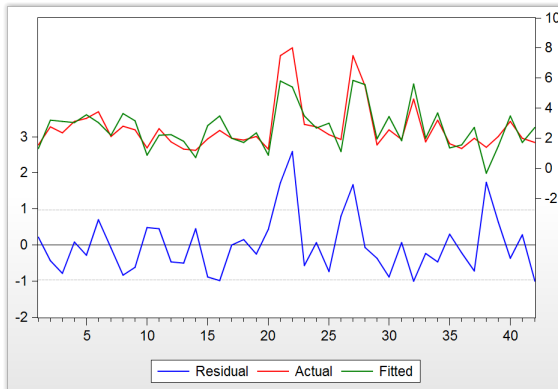
Table – Marginal effects table

Home Win odd : estimation results

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.113895	1.781762	0.063922	0.9495
DELTAFFIFA	-0.077671	0.055243	-1.405973	0.1704
DELTAFORM	-0.075173	0.069053	-1.088626	0.2853
DELTAPOINTS	-0.078102	0.020732	-3.767178	0.0007
DELTA RANK	-0.096871	0.041315	-2.344713	0.0261
DELTA AFTER	-0.110673	0.078749	-1.405382	0.1705
DELTA BEFORE	0.005957	0.083525	0.071314	0.9436
LAST10HOME	-0.335004	0.184321	-1.817499	0.0795
STRICTNESS	-0.027052	0.175437	-0.154196	0.8785
FIELDGRADE	0.189475	0.109869	1.724546	0.0953
TEMPERATURE	-0.000530	0.024664	-0.021484	0.9830
PACKING	0.004017	0.008199	0.489948	0.6279
DISTANCE	2.20E-05	0.000802	0.027423	0.9783
R-squared	0.756080	Mean dependent var	2.699048	
Adjusted R-squared	0.655147	S.D. dependent var	1.647850	
S.E. of regression	0.967686	Akaike info criterion	3.020855	
Sum squared resid	27.15606	Schwarz criterion	3.558705	
Log likelihood	-50.43796	Hannan-Quinn criter.	3.217999	
F-statistic	7.490946	Durbin-Watson stat	1.626779	
Prob(F-statistic)	0.000005			

Table – eqBWH_first_model

Home Win odd : actual and fitted BWH



Graph – eqBWH_first_model

Home Win odd : biggest errors

Toulouse - Paris Saint-Etienne - Paris Paris - Caen

Away Win odd : estimation

Number of observations and variables

$$n = 42 \text{ and } k = 13$$

The linear equation of the model

$$\begin{aligned} \widehat{BWA} = & \widehat{\beta}_0 + \widehat{\beta}_1 \text{DELTA FIFA} + \widehat{\beta}_2 \text{DELTA FORM} + \widehat{\beta}_3 \text{DELTA POINTS} \\ & + \widehat{\beta}_4 \text{DELTA RANK} + \widehat{\beta}_5 \text{DELTA AFTER} + \widehat{\beta}_6 \text{DELTA BEFORE} \\ & + \widehat{\beta}_7 \text{LAST 10 AWAY} + \widehat{\beta}_8 \text{STRICTNESS} + \widehat{\beta}_9 \text{FIELD GRADE} \\ & + \widehat{\beta}_{10} \text{TEMPERATURE} + \widehat{\beta}_{11} \text{PACKING} + \widehat{\beta}_{12} \text{DISTANCE} \end{aligned}$$

Away Win odd : marginal effects

Variable	Expected effect	Real effect	Significant
DeltaFifa	+	+	Yes
DeltaForm	+	+	Yes
DeltaPoints	+	+	No
DeltaRank	-	+	No
DeltaAfter	+	+	No
DeltaBefore	+	+	No
Last10Away	-	-	No
Strictness	?	-	No
FieldGrade	?	+	Yes
Temperature	?	-	No
Packing	?	+	Yes
Distance	?	0	No

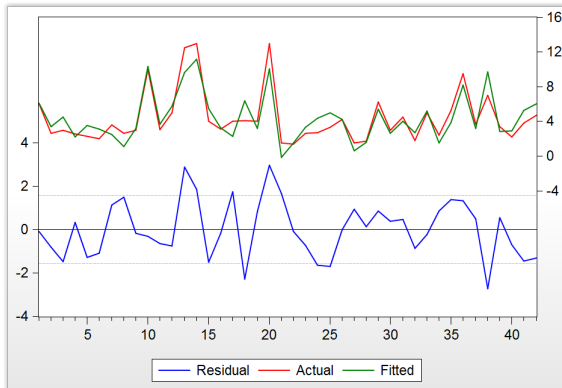
Table – Marginal effects table

Away Win odd : estimation results

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5.088481	2.936266	-1.732977	0.0937
DELTAFFIFA	0.317816	0.081347	3.906935	0.0005
DELTAFORM	0.204098	0.111211	1.835228	0.0767
DELTAPOINTS	0.005793	0.035171	0.164702	0.8703
DELTAARANK	0.065929	0.063985	1.030378	0.3113
DELTAFTER	0.095986	0.127075	0.755355	0.4561
DELTA BEFORE	0.193632	0.136733	1.416130	0.1674
LAST10AWAY	-0.335977	0.268476	-1.251421	0.2208
STRICTNESS	-0.046948	0.272237	-0.172452	0.8643
FIELDGRADE	0.372986	0.173374	2.151337	0.0399
TEMPERATURE	-0.022541	0.038872	-0.579893	0.5665
PACKING	0.047380	0.012897	3.673608	0.0010
DISTANCE	0.002186	0.001290	1.694189	0.1009
R-squared	0.811148	Mean dependent var	4.402857	
Adjusted R-squared	0.733002	S.D. dependent var	3.018771	
S.E. of regression	1.559854	Akaike info criterion	3.975735	
Sum squared resid	70.56115	Schwarz criterion	4.513585	
Log likelihood	-70.49043	Hannan-Quinn criter.	4.172878	
F-statistic	10.37995	Durbin-Watson stat	1.629843	
Prob(F-statistic)	0.000000			

Table – eqBWA_first_model

Away Win odd : actual and fitted BWA



Graph – eqBWA_first_model

Draw odd : estimation

Number of observations and variables

$n = 42$ and $k = 13$

The linear equation of the model

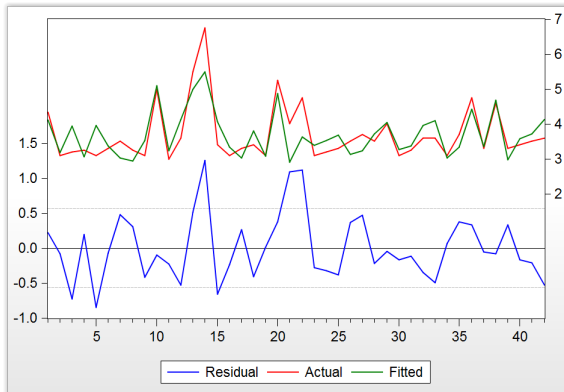
$$\begin{aligned}\widehat{BWD} = & \widehat{\beta}_0 + \widehat{\beta}_1 \text{DELTA FIFA} + \widehat{\beta}_2 \text{DELTA FORM} + \widehat{\beta}_3 \text{DELTA POINTS} \\ & + \widehat{\beta}_4 \text{DELTA RANK} + \widehat{\beta}_5 \text{DELTA AFTER} + \widehat{\beta}_6 \text{DELTA BEFORE} \\ & + \widehat{\beta}_7 \text{LAST 10 DRAW} + \widehat{\beta}_8 \text{STRICTNESS} + \widehat{\beta}_9 \text{FIELD GRADE} \\ & + \widehat{\beta}_{10} \text{TEMPERATURE} + \widehat{\beta}_{11} \text{PACKING} + \widehat{\beta}_{12} \text{DISTANCE}\end{aligned}$$

Draw odd : estimation results

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.112830	1.113255	0.101352	0.9200
DELTA FIFA	0.065161	0.031304	2.081574	0.0463
DELTA FORM	0.052446	0.040267	1.302453	0.2030
DELTA POINTS	-0.013401	0.014083	-0.951558	0.3492
DELTA RANK	0.000465	0.024324	0.019114	0.9849
DELTA AFTER	-0.000606	0.046214	-0.013105	0.9896
DELTA BEFORE	0.011551	0.050700	0.227826	0.8214
LAST10 DRAW	-0.048796	0.114197	-0.427299	0.6723
STRICTNESS	-0.035642	0.097855	-0.364231	0.7183
FIELD GRADE	0.139963	0.063518	2.203506	0.0357
TEMPERATURE	-0.000133	0.013733	-0.009687	0.9923
PACKING	0.016443	0.004736	3.472046	0.0016
DISTANCE	0.000719	0.000469	1.534808	0.1357
R-squared	0.644179	Mean dependent var	3.699524	
Adjusted R-squared	0.496942	S.D. dependent var	0.795941	
S.E. of regression	0.564534	Akaike info criterion	1.943040	
Sum squared resid	9.242245	Schwarz criterion	2.480890	
Log likelihood	-27.80384	Hannan-Quinn criter.	2.140183	
F-statistic	4.375129	Durbin-Watson stat	1.667700	
Prob(F-statistic)	0.000559			

Table – eqBWD01_first_model

Draw odd : actual and fitted BWD



Graph – eqBWD01_first_model

Draw odd : new estimation

Number of observations and variables

$n = 42$ and $k = 13$

The new linear equation of the model

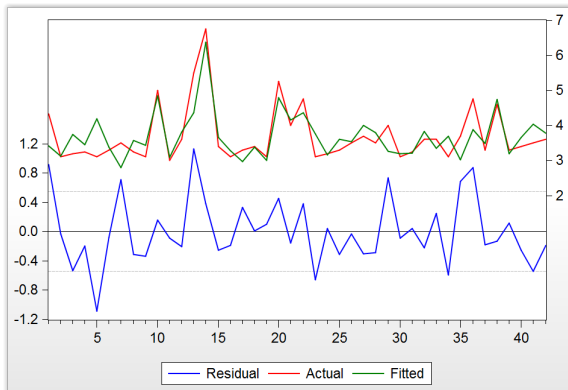
$$\begin{aligned}\widehat{BWD} = & \widehat{\beta}_0 + \widehat{\beta}_1 DELTAFIFA^2 + \widehat{\beta}_2 DELTAFORM^2 + \widehat{\beta}_3 DELTAPOINTS^2 \\ & + \widehat{\beta}_4 DELTARANK^2 + \widehat{\beta}_5 DELTAAFTER^2 + \widehat{\beta}_6 DELTABEFORE^2 \\ & + \widehat{\beta}_7 LAST10DRAW + \widehat{\beta}_8 STRICTNESS + \widehat{\beta}_9 FIELDGRADE \\ & + \widehat{\beta}_{10} TEMPERATURE + \widehat{\beta}_{11} PACKING + \widehat{\beta}_{12} DISTANCE\end{aligned}$$

Draw odd : new estimation results

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.642868	1.034390	0.621495	0.5391
DELTAFFIFA^2	0.004735	0.003529	1.341630	0.1901
DELTAFORM^2	0.010437	0.003965	2.632234	0.0135
DELTAPOINTS^2	-5.30E-05	0.000320	-0.165455	0.8697
DELTA RANK^2	0.000990	0.001327	0.745935	0.4617
DELTA AFTER^2	0.001712	0.007808	0.219293	0.8280
DELTA BEFORE^2	-0.015469	0.013948	-1.108995	0.2765
LAST10DRAW	-0.056438	0.095567	-0.590559	0.5594
STRICTNESS	-0.054343	0.098859	-0.549705	0.5867
FIELDGRADE	0.141825	0.053373	2.657258	0.0127
TEMPERATURE	0.002253	0.012913	0.174460	0.8627
PACKING	0.008103	0.004750	1.706105	0.0987
DISTANCE	0.000326	0.000431	0.756723	0.4553
R-squared	0.669497	Mean dependent var	3.699524	
Adjusted R-squared	0.532738	S.D. dependent var	0.795941	
S.E. of regression	0.544078	Akaike info criterion	1.869226	
Sum squared resid	8.584607	Schwarz criterion	2.407076	
Log likelihood	-26.25374	Hannan-Quinn criter.	2.066369	
F-statistic	4.895427	Durbin-Watson stat	1.952639	
Prob(F-statistic)	0.000228			

Table – eqBWD02_first_model

Draw odd : actual and new fitted BWD



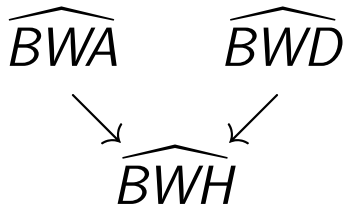
Graph – eqBWD02_first_model

Tests : normality of the errors

	BWH	BWD	BWA
R^2	0.76	0.82	0.81
Adjusted R^2	0.66	0.69	0.73
$JB = \frac{n-k}{n} JB_E$	9.44	1.19	0.37
$\chi^2_{0.95}(2)$	5.99	5.99	5.99
Reject normality	Yes	No	No

Table – normality tests results

Hybrid model



Home Win odd using Draw and Away fitted odds

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.336080	3.318576	-0.101272	0.9201
DELTAIFA	-0.033932	0.159845	-0.212284	0.8335
DELTAFORM	-0.116623	0.121362	-0.960951	0.3451
DELTAPOINTS	-0.045598	0.022004	-2.072266	0.0479
DELTA RANK	-0.045178	0.053247	-0.848471	0.4036
DELTA AFTER	-0.050559	0.085025	-0.594639	0.5570
DELTA BEFORE	0.006335	0.105236	0.060197	0.9524
LAST10HOME	-0.327776	0.189951	-1.725585	0.0959
STRICTNESS	0.012225	0.145124	0.084237	0.9335
FIELDGRADE	-0.008130	0.220232	-0.036918	0.9708
TEMPERATURE	-0.003097	0.022734	-0.136225	0.8927
PACKING	-0.005979	0.025422	-0.235196	0.8158
DISTANCE	-0.000876	0.001329	-0.659361	0.5152
BWAF	-0.156665	0.512008	-0.305982	0.7620
BWDF	1.434914	0.359605	3.990251	0.0005
R-squared	0.849164	Mean dependent var	2.699048	
Adjusted R-squared	0.770952	S.D. dependent var	1.647850	
S.E. of regression	0.788643	Akaike info criterion	2.535448	
Sum squared resid	16.79288	Schwarz criterion	3.256044	
Log likelihood	-40.34441	Hannan-Quinn criter.	2.862921	
F-statistic	10.85728	Durbin-Watson stat	1.989504	
Prob(F-statistic)	0.000000			

Table – eqBWH_third_model

Wald Test			
Equation: BWH_HYBRID			
Test Statistic	Value	df	Probability
F-statistic	8.331091	(2, 27)	0.0015
Chi-square	16.66218	2	0.0002
Null Hypothesis: C(14)=0, C(15)=0			
Null Hypothesis Summary:			
Normalized Restriction (= 0)	Value	Std. Err.	
C(14)	-0.156665	0.512008	
C(15)	1.434914	0.359605	
Restrictions are linear in coefficients.			

Table – Wald_fitted_coeffs

Wald tests on Away win and Draw fitted coefficients

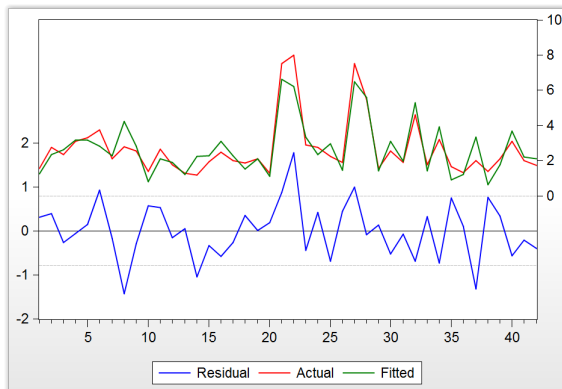
Wald Test: Equation: BWH_HYBRID			
Test Statistic	Value	df	Probability
t-statistic	-0.305982	27	0.7620
F-statistic	0.093625	(1, 27)	0.7620
Chi-square	0.093625	1	0.7596
Null Hypothesis: C(14)=0 Null Hypothesis Summary:			
Normalized Restriction (= 0)	Value	Std. Err.	
C(14)	-0.156665	0.512008	
Restrictions are linear in coefficients.			

Table – Wald_BWAF_coeff

Wald Test: Equation: BWH_HYBRID			
Test Statistic	Value	df	Probability
t-statistic	3.990251	27	0.0005
F-statistic	15.92210	(1, 27)	0.0005
Chi-square	15.92210	1	0.0001
Null Hypothesis: C(15)=0 Null Hypothesis Summary:			
Normalized Restriction (= 0)	Value	Std. Err.	
C(15)	1.434914	0.359605	
Restrictions are linear in coefficients.			

Table – Wald_BWDF_coeff

Home Win odd using Draw and Away fitted odds



Graph – eqBWH_second_model

Sequential model

$$\widehat{BWA} \longrightarrow \widehat{BWD} \longrightarrow \widehat{BWH}$$

Draw odd using Away fitted odds

Dependent Variable: BWD
 Method: Least Squares
 Date: 04/06/17 Time: 14:03
 Sample: 1 42
 Included observations: 42

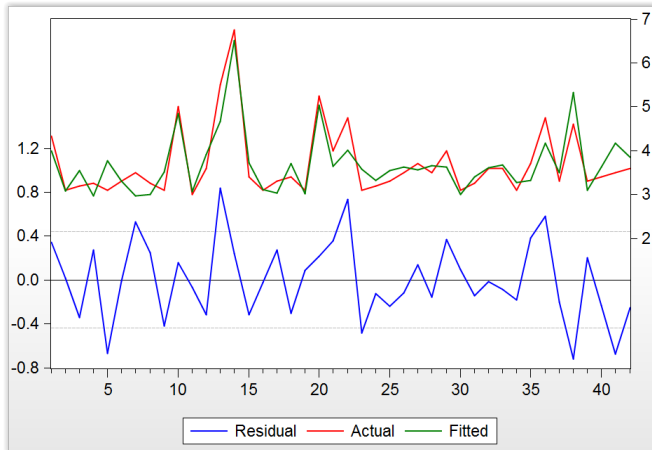
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.134966	0.910711	2.344284	0.0264
DELTAIFA*2	0.003733	0.002856	1.307225	0.2018
DELTAFORM*2	0.006864	0.003315	2.070707	0.0477
DELTAPOINTS*2	0.000232	0.000268	0.865463	0.3941
DELTAIRANK*2	0.000919	0.001070	0.858816	0.3977
DELTAFTER*2	-0.003905	0.006444	-0.605919	0.5494
DELTAFORE*2	-0.002086	0.011714	-0.178089	0.8599
LAST10DRAW	-0.068000	0.077098	-0.881989	0.3853
STRICTNESS	-0.042054	0.079757	-0.527277	0.6022
FIELDGRADE	0.024086	0.051823	0.464767	0.6457
TEMPERATURE	0.005603	0.010443	0.536575	0.5958
PACKING	0.004763	0.003916	1.216279	0.2340
DISTANCE	0.000113	0.000351	0.321723	0.7501
BWAF	0.151465	0.037155	4.076582	0.0003
R-squared	0.792596	Mean dependent var	3.699524	
Adjusted R-squared	0.696301	S.D. dependent var	0.795941	
S.E. of regression	0.438634	Akaike info criterion	1.450900	
Sum squared resid	5.387202	Schwarz criterion	2.030123	
Log likelihood	-16.46890	Hannan-Quinn criter.	1.663208	
F-statistic	8.230922	Durbin-Watson stat	2.029653	
Prob(F-statistic)	0.000002			

Table – eqBWD_third_model

Wald Test			
Equation: EQBWD_SEQ			
Test Statistic	Value	df	Probability
t-statistic	4.076582	28	0.0003
F-statistic	16.61852	(1, 28)	0.0003
Chi-square	16.61852	1	0.0000
Null Hypothesis: C(14)=0			
Null Hypothesis Summary:			
Normalized Restriction (= 0)	Value	Std. Err.	
C(14)	0.151465	0.037155	
Restrictions are linear in coefficients.			

Table – Wald_BWAF_coeff

Draw odd using the Away fitted odds



Home Win odd using new Draw fitted odds

Dependent Variable: BWH
 Method: Least Squares
 Date: 04/06/17 Time: 14:17
 Sample: 1 42
 Included observations: 42

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.401025	1.322267	-0.303286	0.7639
DELTAFFIFA	-0.169117	0.044781	-3.776571	0.0008
DELTAFORM	-0.190031	0.056020	-3.392195	0.0021
DELTAPOINTS	-0.050727	0.016288	-3.114458	0.0042
DELTA RANK	-0.072687	0.030948	-2.348727	0.0261
DELTAFTER	-0.046663	0.059654	-0.782234	0.4406
DELTABEFORE	-0.061448	0.063251	-0.971495	0.3396
LAST10HOME	-0.383320	0.136713	-2.803817	0.0091
STRICTNESS	0.031565	0.130327	0.242196	0.8104
FIELDGRADE	-0.031483	0.092532	-0.340244	0.7362
TEMPERATURE	0.001819	0.018254	0.099634	0.9213
PACKING	-0.024497	0.008328	-2.941667	0.0065
DISTANCE	-0.001672	0.000683	-2.446211	0.0210
BWDF_VIA_BWAF	1.833303	0.366814	4.997903	0.0000
R-squared	0.871086	Mean dependent var	2.699048	
Adjusted R-squared	0.811232	S.D. dependent var	1.647850	
S.E. of regression	0.715948	Akaike info criterion	2.430783	
Sum squared resid	14.35228	Schwarz criterion	3.010006	
Log likelihood	-37.04644	Hannan-Quinn criter.	2.643091	
F-statistic	14.55371	Durbin-Watson stat	2.079829	
Prob(F-statistic)	0.000000			

Table – eqBWH_third_model

Wald Test
 Equation: EQBWH_SEQ

Test Statistic	Value	df	Probability
t-statistic	4.997903	28	0.0000
F-statistic	24.97904	(1, 28)	0.0000
Chi-square	24.97904	1	0.0000

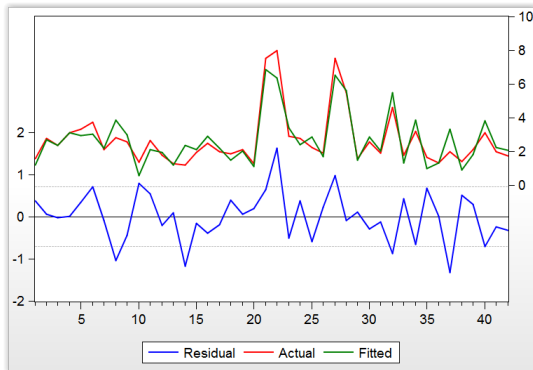
Null Hypothesis: C(14)=0
 Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(14)	1.833303	0.366814

Restrictions are linear in coefficients.

Table – Wald_BWDF_coeff

Home odd using the new Draw fitted odds



Graph – eqBWH_third_model

Tests : normality of the errors

	simultaneous			hybrid			sequential		
	BWH	BWD	BWA	BWH	BWD	BWA	BWH	BWD	BWA
R^2	0.76	0.67	0.81	0.85	0.67	0.81	0.87	0.79	0.81
Adjusted R^2	0.66	0.53	0.73	0.77	0.53	0.73	0.81	0.70	0.73
JB	9.44	1.19	0.37	0.26	1.19	0.37	0.19	0.16	0.37
$\chi^2_{0.95}(2)$	5.99	5.99	5.99	5.99	5.99	5.99	5.99	5.99	5.99
Reject normality	Yes	No	No	No	No	No	No	No	No

Table – normality tests results

-  Football-data. www.football-data.co.uk.
-  LFP. www.lfp.fr.
-  SoFIFA. sofifa.com.
-  Fussbal WettPoint. fussball.wettpoint.com.
-  Soccer Punter. www.soccerpunter.com.
-  Distance between two cities. www.distance2villes.com.
-  Weather records. www.historique-meteo.net.
-  BWin. sports.bwin.fr.