ECO 5120
Econometric theory and application

Parametric Model on Soccer Scoring time

Presented by:

Ko Chiu Yu **05168070**

Ng Shou Zhong **05168560**

Literature Review

Literature Review

Model

Data

Estimation

Conclusion

Game Theorists View: Skill, Strategy and Passion: an empirical Analysis of Soccer (Frederic, Luca, Aldo, Working Paper, CentER, April 2000)

Independent Variable	Probability of scoring
Team skills	↑
Winning	↓
Losing	↑
Home field	↑

Literature Review

Literature Review

Model

Data

Estimation

Conclusion

Statisticians views: Maher(1982), Baxter and Steveson(1988), Ridder et al. (1994), Jackson(1994), Fahrmeir and Tutz (1994), Clarke and Norman (1995), Dixon and Coles(1997), Lee(1997), Pollard and Reep (1997), Rue and Salvesen (1997), and Kuonen (1997a,b).

Dependent Variable	Regression Technique
Scoring of goals	Usual Possion
	Negative Binomial
	Mixed Possion (including effect of home effect, offensive and defensive abilities)

Our Approach

Literature Review

Model

Data

Estimation

Conclusion

Combine views of game theorists and statisticians……

Game Theorist

Probability of Scoring

Team skills
Current Score
Home Field Advantage

Statistican

Number of goal per game

Simple Possion
Mixed Possion
Negative Binomial

Parametric Survival Model

Data Sample

Literature Review

Model

Data

Estimation

Conclusion

Data	**English Barclays Premiership**	**Easy Access**
Period	**Year 2001 to Year 2004**	**Enable Prediction**
Sample	**Liverpool, Chelsea, Manchester United, Arsenal**	**Min. difference in ability**
Main source	**FA Premier League (official) http://www.premierleague.com**	
Other source	**Other football internet source: http://stats.premierleague.com**	

Data Set Problem

Compared with count data in previous studies, our studies are having the following particular problems:

- **Panel data with attrition across club and across players**
- **Censored Scoring Time**
- **Lack of proxy for:**
  - ✓ **Players ability**
  - ✓ **Players strategies**
  - ✓ **Newly promoted clubs effect**
  - ✓ **Overwhelming strong clubs effect**

Variables Collected

There are six categories of instruments:

1. Match specific
2. Field specific
3. Strategy specific
4. Last matches results
5. Own team specific
6. Opponent team specific

Variables Collected

Literature Review

Model

Data

Estimation

Conclusion

Category	Proxy Variable
Match Specific	Climate
	Home field advantage
	Attendance ratio
Field specific	Current score
	Current goal home
	Current goal away
	Current goal conceded home
	Current goal conceded away

Variables Collected

Literature Review

Model

Data

Estimation

Conclusion

Category	Proxy Variable
Strategic Specific	Last year fouls
	Last year yellow/red cards
	Last year goal
	Last year goal conceded
	Last year shoot
Last Matches	Previous Three Matches Results
	Current Position in the league
Opponent Team	Last year position
	Newly promoted team effect
Own Team	Last year position

Estimation Model

Literature Review

Model

Data

Estimation

Conclusion

Parametric (Weilbull) Survival Model:

Hazard function:

$$\lambda(t) = \gamma\alpha t^{\alpha-1}$$

Taking Logarithm of both sides yields:

$$\ln\lambda(t) = \ln(\gamma\alpha) + (\alpha - 1)\ln t$$

where $\gamma = \exp(x'\beta)$

and x is a vector of independent variables

Non-Parametric Estimation

Literature Review

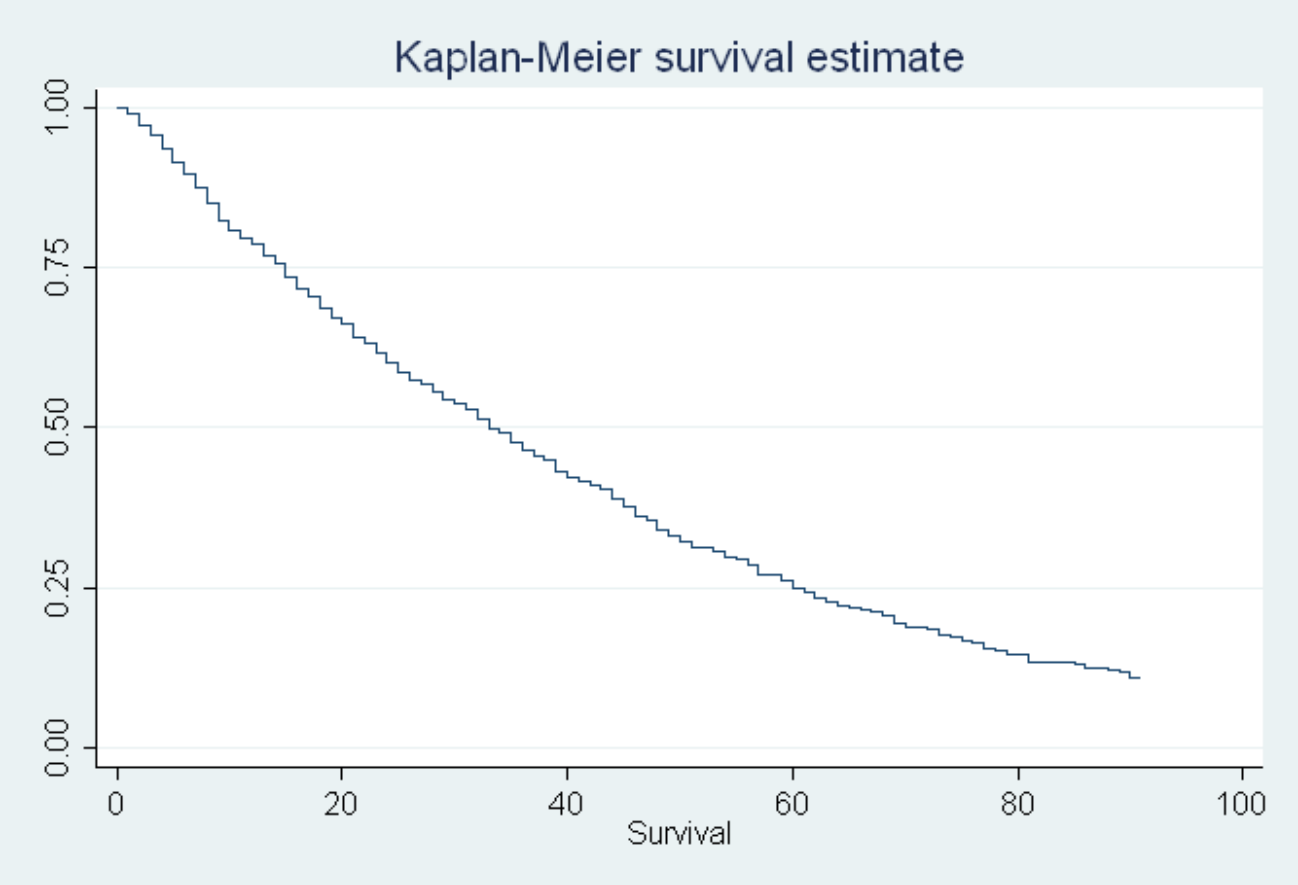
Model

Data

Estimation

Conclusion

Empirical Kaplan-Meier Survival Estimate



Non-Parametric Estimation

Literature Review

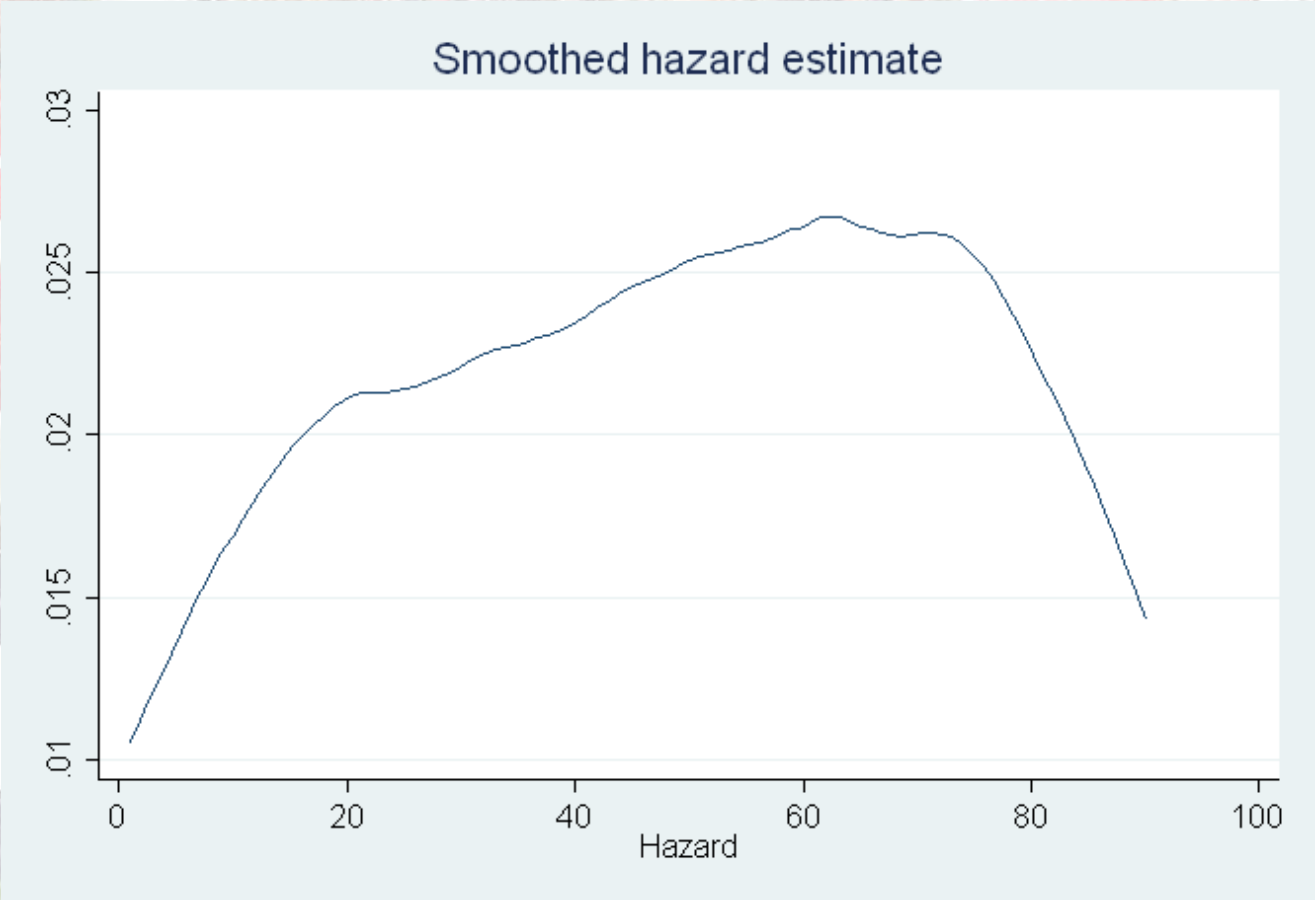
Model

Data

Estimation

Conclusion

Empirical Kaplan-Meier Smoothed Hazard Estimate



Estimation Result

Independent Variables	Coefficient	Hazard Ratio	P-value
Match specific			
Home Field Advantage**	0.147	1.159	0.077
Field specific			
Current Score**	0.119	1.126	0.094
Net Goal (Home)***	0.157	1.17	0.00
Net Goal (Away)***	0.185	1.203	0.002
Opponent Ability			
Opponent Last year position***	0.011	1.012	0.00
Previous Game effects			
Weighted previous three results	0.003	1.003	0.237
Strategic Specific			
Last Year Goal ***	0.006	1.005	0.00
Last Year Goal Conceded	-0.009	0.99	0.301
Constant Term	-5.243	0.00	0.764
Number of observation	909		
Log likelihood	-1095		
Chi-Square Test	0.000		

*** *significant at 0.005*

**significant at 0.01*

Robust standard error Used

Parametric Estimation

Literature Review

Model

Data

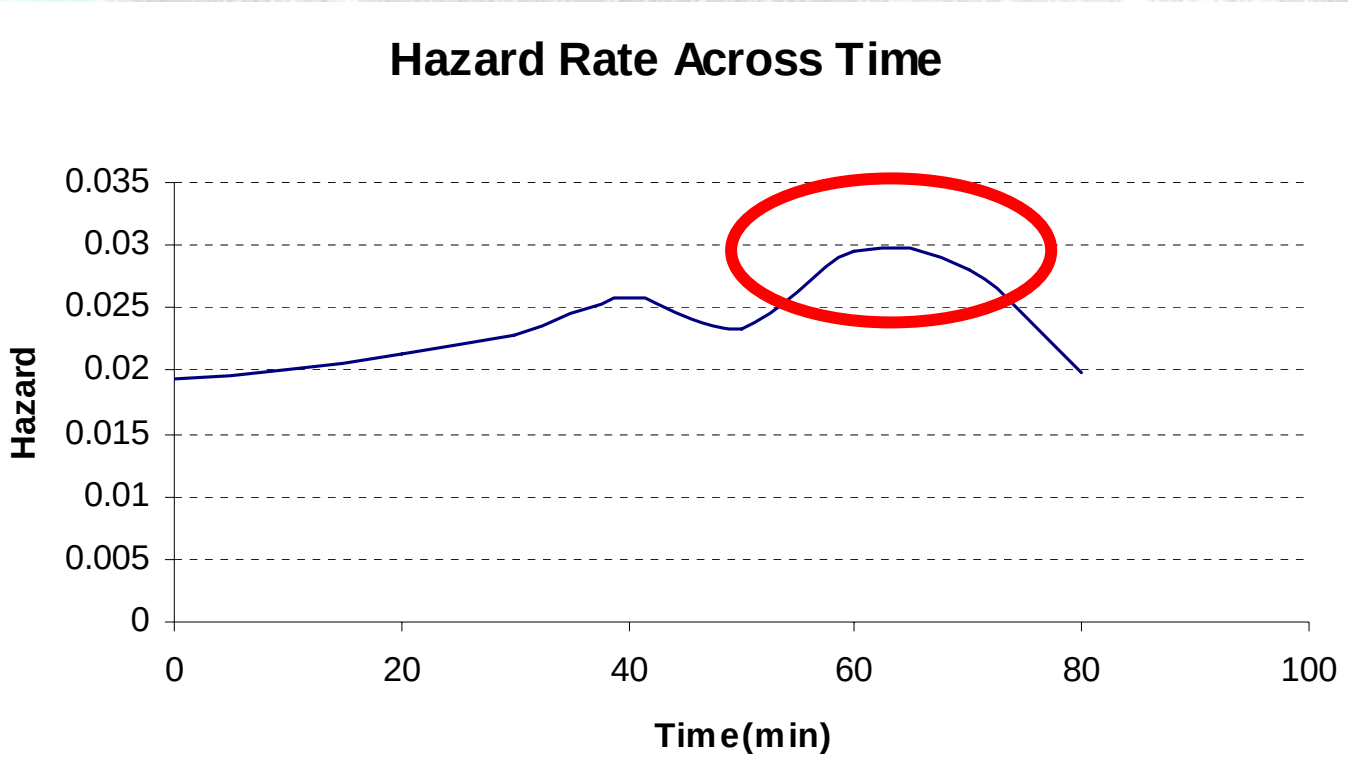
Estimation

Conclusion

Variable	Marginal effect
Opponent ability	-0.305
Home effect	-3.816
Current score	-3.062
Net goal (Home)	-4.045
Net goal (Away)	-4.763
Last year goal	-0.153
Last year goal conceded	0.244
Previous game result	-0.092

Parametric Estimation

- Literature Review
- Model
- Data
- **Estimation**
- Conclusion



Parametric Estimation

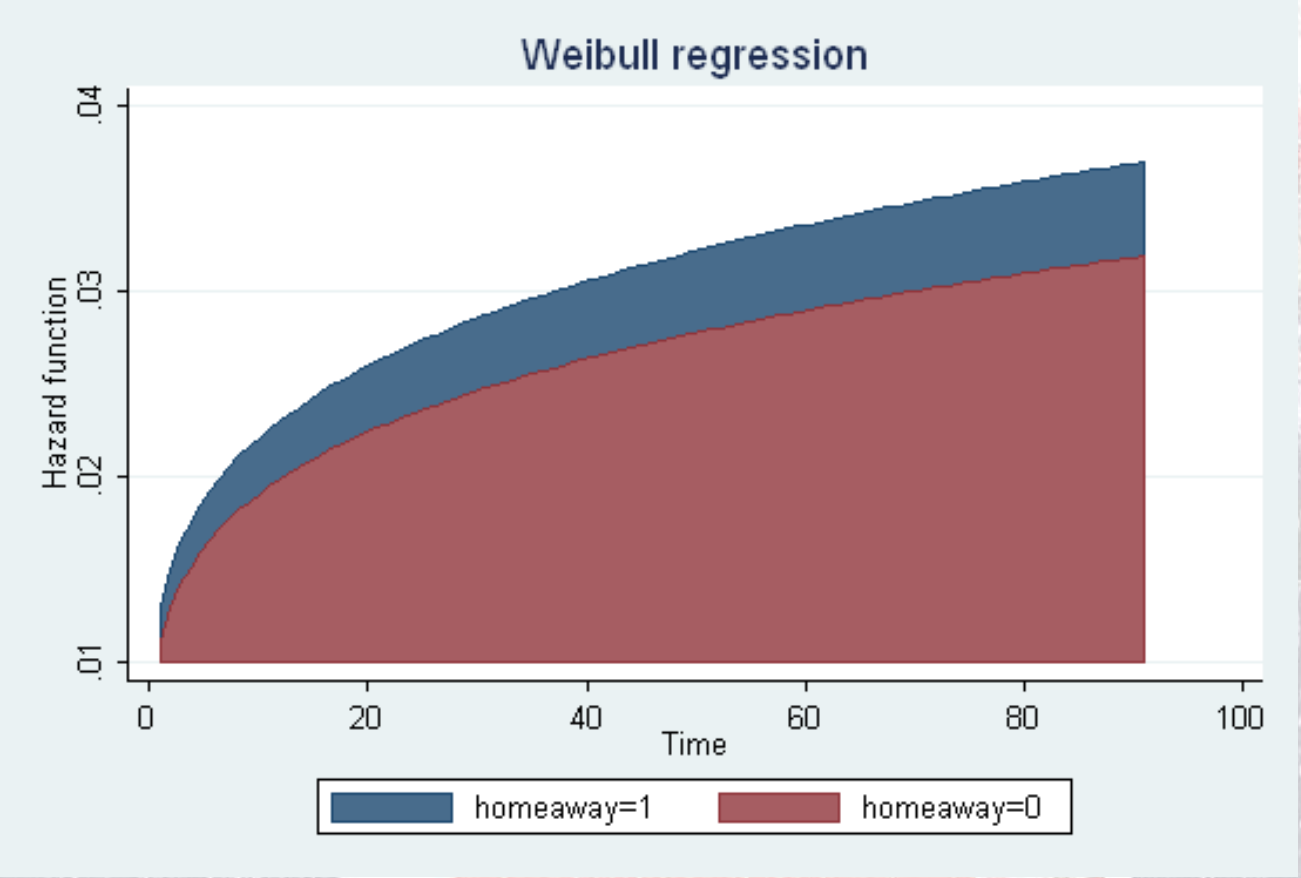
- Literature Review
- Model
- Data
- **Estimation**
- Conclusion

Time interval		Beg.	Cum.	
From	To	Total	Failure	Hazard
0	10	912	0.1768	0.0194
10	20	686	0.3281	0.0202
20	30	520	0.4572	0.0213
30	40	383	0.5683	0.0228
40	50	275	0.6668	0.0258
50	60	191	0.7364	0.0233
60	70	135	0.8043	0.0296
70	80	88	0.8523	0.028
80	90	55	0.8789	0.0198

Parametric Estimation

- Literature Review
- Model
- Data
- **Estimation**
- Conclusion

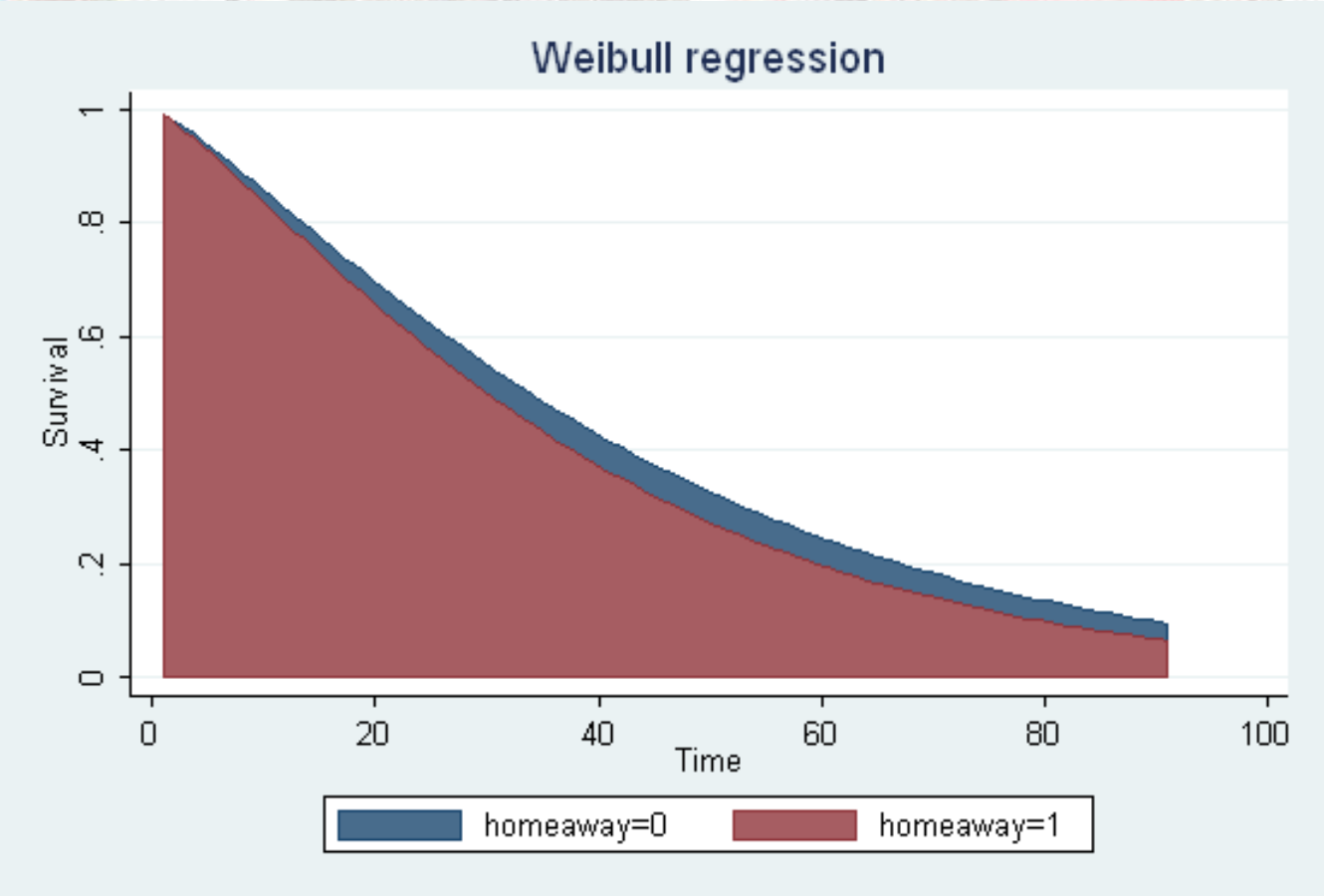
Home effect on Hazard Function



Parametric Estimation

- Literature Review
- Model
- Data
- **Estimation**
- Conclusion

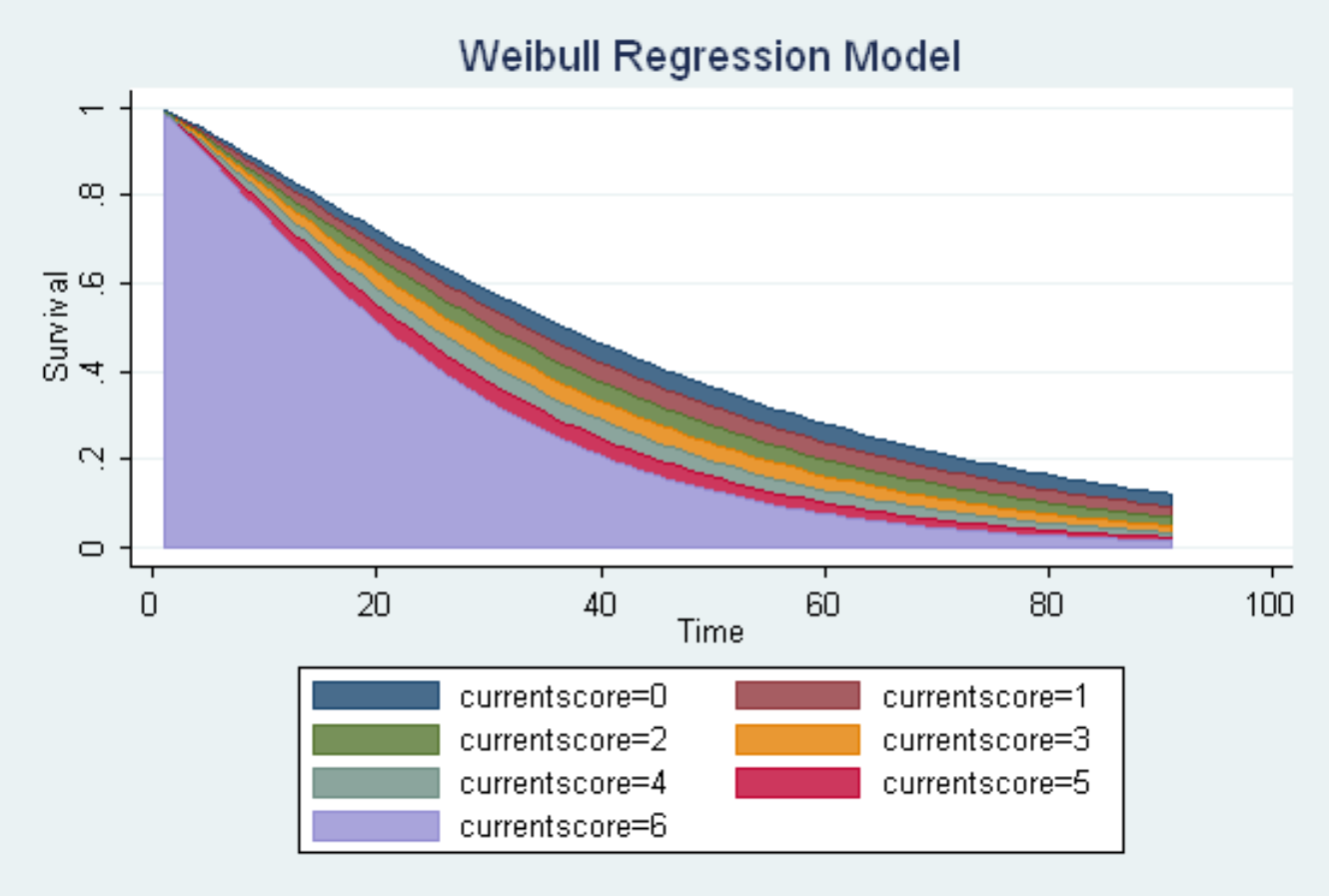
Home effect on Survival Function



Parametric Estimation

- Literature Review
- Model
- Data
- **Estimation**
- Conclusion

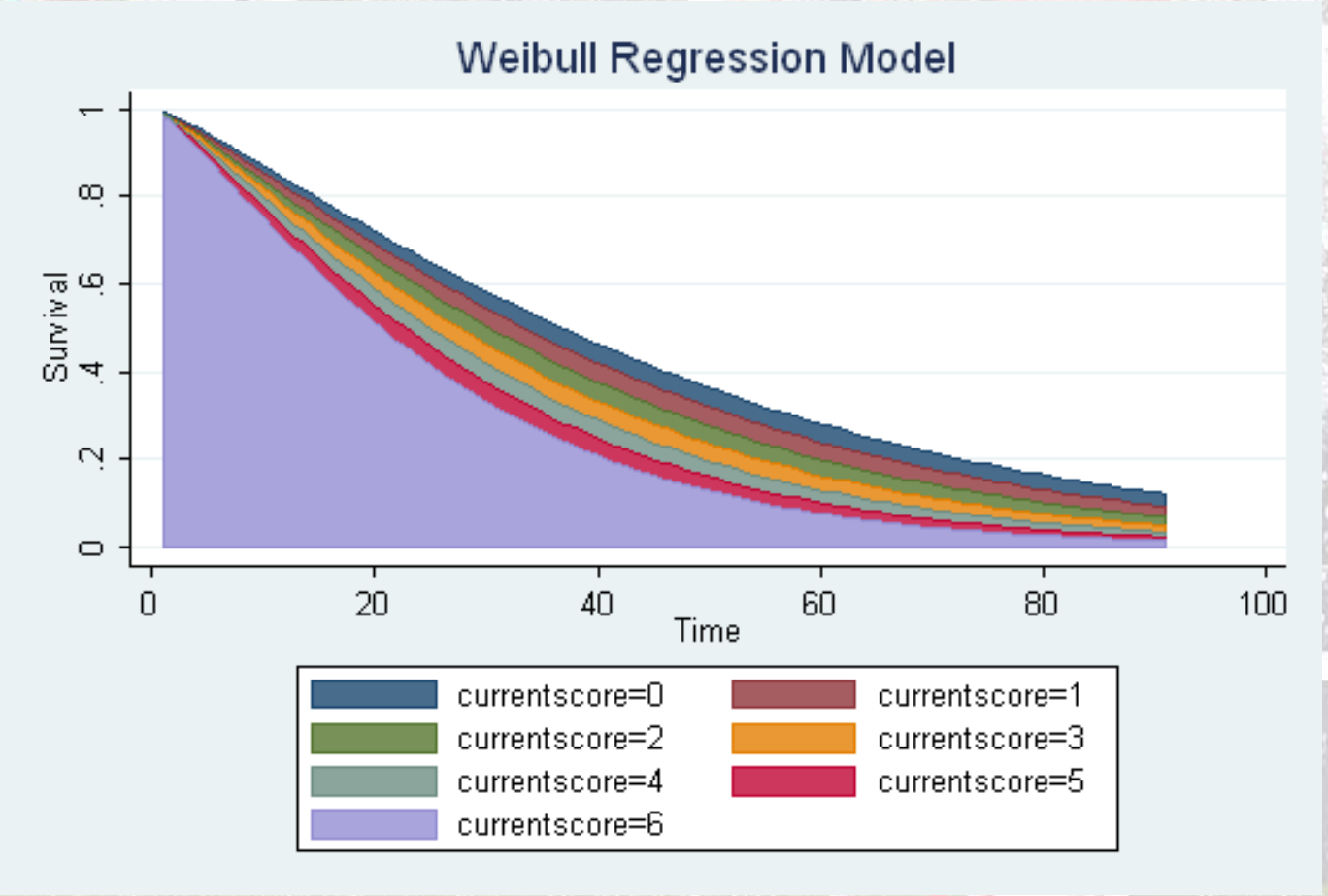
Current Score effect on Survival Function



Parametric Estimation

- Literature Review
- Model
- Data
- **Estimation**
- Conclusion

Current Score effect on Survival Function



Side-story: Beckham's effect

Literature Review

Model

Data

Estimation

Conclusion

We would like to test whether the leave of Beckham would fundamentally change the scoring time of Man United.

From 2001 to 2002

From 2003 to 2004

Side-story: Beckham's effect

Literature Review

Model

Data

Estimation

Conclusion

Assuming the presence of Beckham is time-invariant fixed effect to the scoring probability of Manchester United, we tried to use dummy to proxy his contribution. Beckham's effect is measured by:

$$\ln \lambda(t) = \ln \alpha + \beta_1 x_1 + \ldots\ldots + \beta_{Beckham'seffect}\, x_{beckham'seffect}$$

$$+ (\alpha - 1) \ln t - \mu \ln S(t)$$

where $x_{beckham'seffect} = \begin{cases} 1 & \text{in season } 01/02 - 02/03 \\ 0 & \text{in season } 03/04 - 04/05 \end{cases}$

Side-story: Beckham's effect

Literature Review

Model

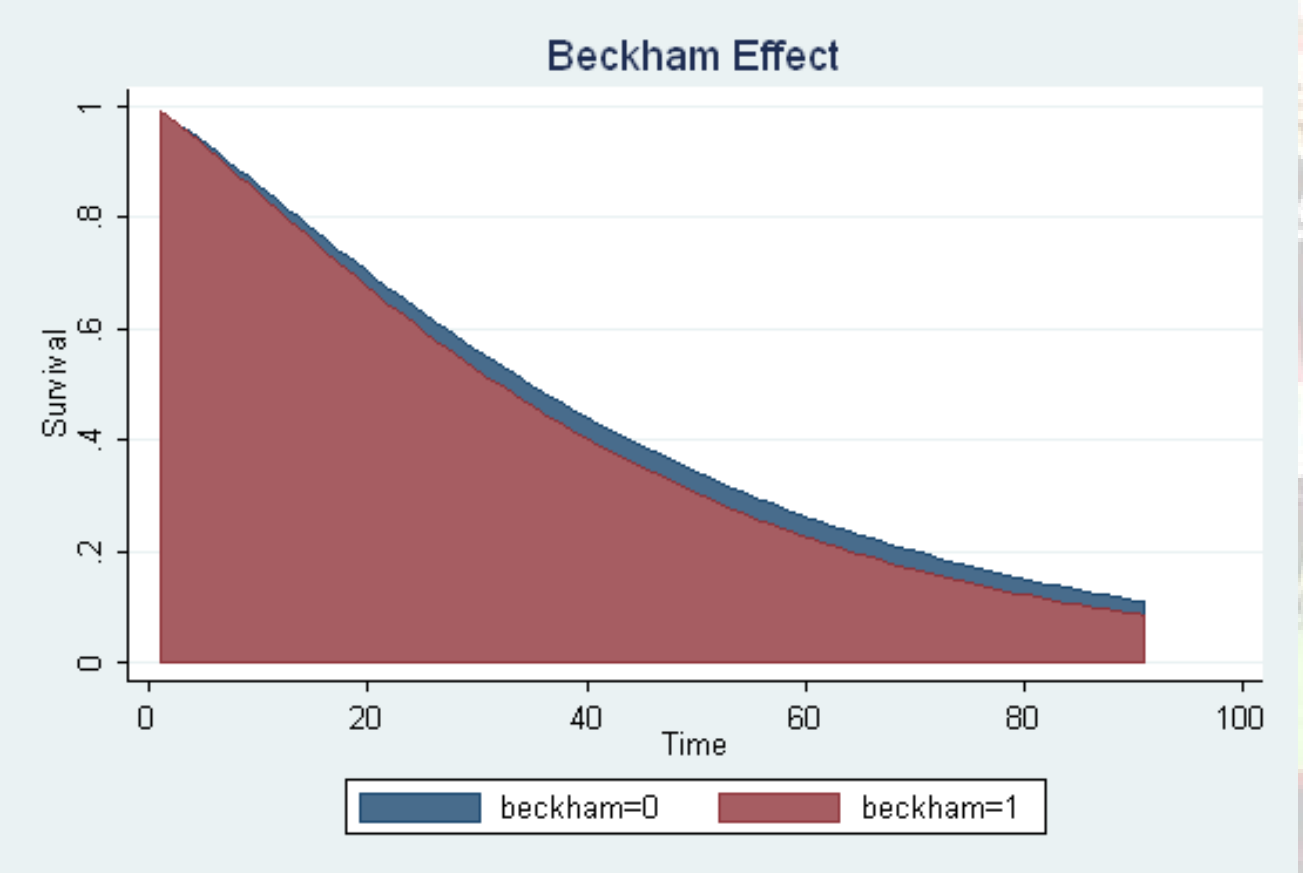
Data

Estimation

Conclusion

Beckham effect is positive to the scoring probability

Yet it is not so significant (p ~ 0.5)



Side-story: Beckham's effect

Literature Review

Model

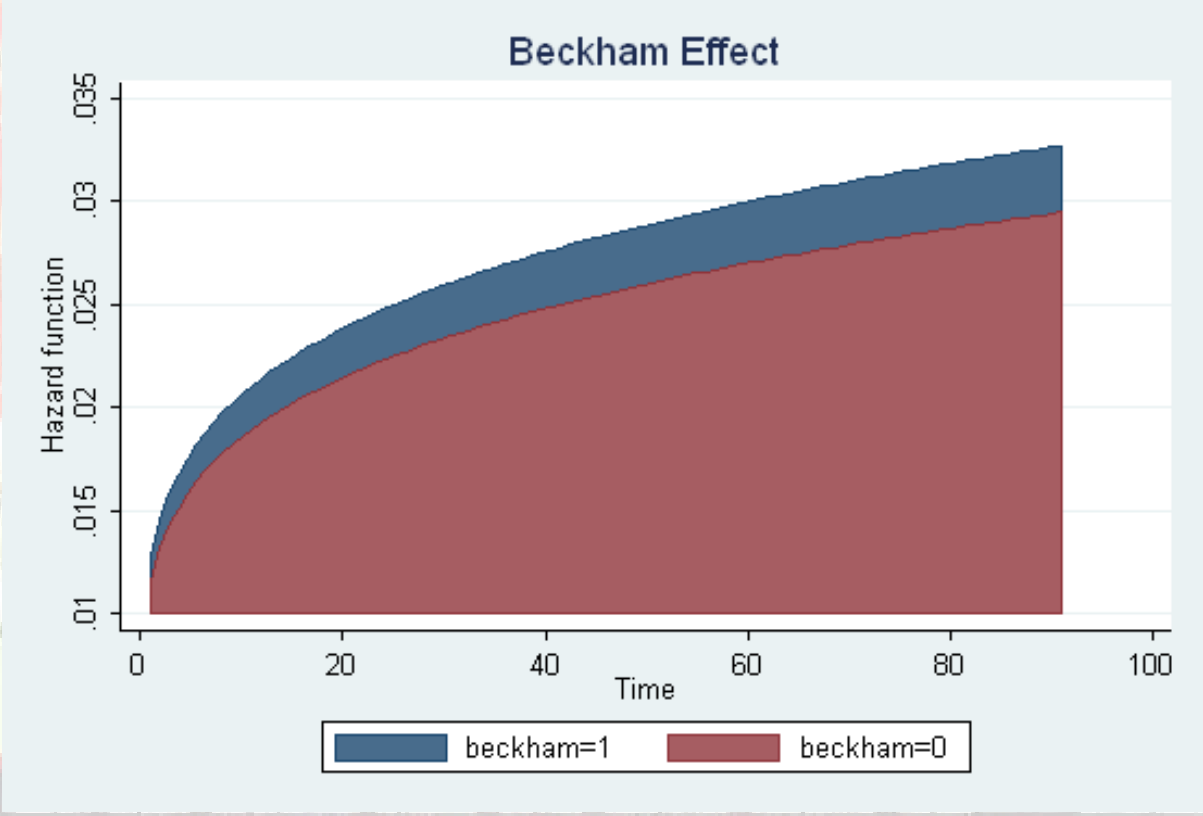
Data

Estimation

Conclusion

Beckham effect is positive to the scoring probability

Yet it is not so significant (p ~ 0.5)



Possible Extension

Literature Review

Model

Data

Estimation

Conclusion

The model derived can be extended by:

1. Including more mid-stream team data
2. Add strategic variables of formation arrangement
3. Observe the effect of pre/post half-time
4. Observe the difference between time to 1st goal and time to 2nd goal
5. Run panel data across different league in different countries
6. Evaluation of each player's marginal contribution

References

Literature Review

Model

Data

Estimation

Conclusion

References

1. AD Fitt, CJ Howls and M Kabelka, [2005]: " Valuation of soccer spread bets", Journal of the Operational Research Society.

C. S. Lam, [2005]: "Survival Analysis of the timing of goals in soccer games", Hong Kong Economic Journal Monthly, pp.68-pp.71.

2. Isabelle Brocas, Juan D Carrillo, [2002]: "Do the 'Three-point victory' and 'Golden goal' rules make soccer game more exciting? A theoretical analysis of a simple game", Centre for Economic Policy Research, Discussion paper series, no. 3266.

3. Myoung-jae Lee, [1996]: "Methods of moments and semiparametric econometrics for limited dependent and variable models", New York: Springer.

4. Palomino, F., Rigotti, L. and Rustichini, A. [2000]: "Skill, Strategy and Passion: An Empirical Analysis of Soccer", CentER.

William H. Greene, [2003]: "Econometric Analysis", Prentice Hall, Fifth edition.

5. Wooldridge, Jeffrey M., [2002]: "Econometric analysis of cross section and panel data", Cambridge, Mass.: MIT Press.


7 CRISTIANO RONALDO
MANCHESTER UNITED PLAYER OF THE SEASON 2003/2004
THANK YOU
Vodafone
AS VOTED FOR BY THE USERS OF MANUTD ONLINE
WWW.MANUTDONLINE.ORG 2003/2004