



Les déterminants de la demande de tickets de football : l'exemple de la ligue féminine anglaise (saison 2022-2023)

Sarah Mehiyddine

Contents

1	Introduction	1
2	Présentation des données	2
2.1	Description des variables	2
2.2	Statistiques descriptives	2
3	Modèle économétrique	3
3.1	Interprétation des résultats des estimations	3
3.2	Test	5
3.2.1	Test d'existence des effets spécifiques	5
3.2.2	Test d'indépendance d'Hausman	5
4	Conclusion	5

1 Introduction

Cette étude se concentre sur l'analyse des déterminants de la demande de billets pour assister à des matchs de football féminin, en s'appuyant sur la FA Women Super League pour la saison 2022-2023. Cet intérêt pour la discipline et l'analyse de sa demande provient d'un double enjeu : d'une part, les clubs opèrent dans un cadre économique et doivent donc comprendre les variables influençant l'affluence dans les stades pour atteindre leurs objectifs, et d'autre part, la compréhension de la demande est cruciale pour le développement de la discipline, en particulier dans le contexte de l'émergence des sections féminines de football.

Une multitude d'études similaires ont été menées dans la littérature. Mettre en avant les déterminants de la demande a été un enjeu majeur dans plusieurs pays durant la deuxième moitié du 20ème siècle suite au déclin de l'affluence dans les stades (Jennett [1984], Borland [1987]). Dans leur étude utilisant des données de panel, Borland and Lye [1992] ont montré que la demande pour les événements sportifs était inélastique au prix et que les billets pour assister aux matchs étaient des biens inférieurs. D'autres variables, davantage liées à la qualité du jeu, étaient en revanche déterminantes dans la demande : la taille des stades ou encore l'incertitude. García and Rodríguez [2002] ont effectué une analyse similaire en se concentrant sur la ligue espagnole entre 1992 et 1996. En utilisant des données de panel, ils ont montré que les variables liées à la qualité ex-ante des matchs disposaient du plus grand pouvoir explicatif.

Nous nous appuyerons sur le modèle et les variables utilisées dans ce dernier article pour conduire notre étude sur la ligue féminine anglaise pour la saison 2022-2023. Le modèle considéré est donc :

$$\log(Y_{i,t}) = \beta_0 + \beta_1 \text{prix}_{i,t} + \beta_2 X'_{i,t} + \beta_3 Z'_{i,t} + \beta_4 C'_{i,t} + u_{i,t} \quad (1)$$

où la variable endogène $Y_{i,t}$ représente la proportion du stade rempli pour l'équipe i à la période t . $X'_{i,t}$ est un vecteur de variables relative à la qualité ex-ante, $Z'_{i,t}$ un vecteur de variables sur les performances récentes de l'équipe i et $C'_{i,t}$ un vecteur des variables relatives au coût d'opportunité d'assister au match de l'équipe i à la date t .

La première section sera consacrée à la présentation des données et des variables utilisées. La deuxième section présentera les résultats ainsi que les interprétations économiques des estimations. Le modèle présenté sera estimé de 3 façons différentes : avec un estimateur OLS, Within et GLS. Les deux derniers estimateurs considéreront un effet individuel, fixe pour l'estimateur Within et aléatoire pour l'estimateur GLS. Enfin, une dernière section sera consacrée aux tests statistiques qui permettront de justifier le choix de l'estimateur GLS avec effets individuels aléatoires comme étant le plus pertinent dans le cadre de cette étude.

2 Présentation des données

2.1 Description des variables

On utilise ici des données de panel où chaque individu i représente un club et où chaque période t est un match joué à domicile. Les données sur l’affluence dans les stades sont issues du site SkySport qui récapitule l’ensemble des statistiques pour chaque match joué durant la saison. Une particularité de la ligue féminine à prendre en considération est que l’ensemble des matchs d’une même équipe n’est pas joué dans le même stade. L’attractivité de cette ligue étant plus faible que la ligue masculine, de plus petits stades sont mis à disposition des équipes par les clubs. Cependant, il peut arriver de manière exceptionnelle que certains matchs, plus attractifs, soient joués dans le stade officiel du club. Les capacités d’accueil des stades étant différentes, la variable expliquée sera la proportion remplie du stade et non pas simplement l’affluence afin de ne pas biaiser l’analyse. Les variables explicatives utilisées sont similaires à celles de García and Rodríguez [2002], et peuvent être classées en 4 catégories : variable économique, qualité ex-ante du match, performances récentes de l’équipe et coût d’opportunité.

Variable économique. Il s’agit du prix le plus faible d’un billet pour assister au match de l’équipe i . Les données utilisées sont issues de Statistica¹. Les prix des matchs ayant été joués dans un autre stade que le stade habituel ont été récoltés directement sur le site des clubs.

Qualité ex-ante. Cela correspond à la qualité attendue du match, elle est mesurée par le nombre de joueuses jouant en équipe nationale dans la composition de départ. À cela, s’ajoutent 2 variables binaires qui prennent la valeur 1 lorsque l’équipe extérieure est Chelsea ou Arsenal, qui sont les deux équipes les plus titrées de la ligue. Enfin, une variable mesurant la différence de classement entre les deux équipes est prise en compte. On s’attend à ce que les 3 premières variables exercent une influence positive sur la demande de billets en augmentant l’attractivité des matchs. La dernière variable, quant à elle, suit la logique inverse.

Performances récentes. Elles seront mesurées par le nombre de victoires au cours des 3 derniers matchs disputés par le club dans la ligue ainsi que par la possibilité de relégation pour l’équipe.

Coût d’opportunité. 2 variables sont considérées : une variable binaire qui prend la valeur 1 si le match est joué un week-end et une variable qui mesure la distance en kilomètres entre le stade des 2 équipes.

2.2 Statistiques descriptives

La table 1 présente diverses statistiques descriptives relatives aux différentes équipes. Assister à un match coûte en moyenne £8.96 et engendre un déplacement d’en moyenne 160km. Les statistiques relatives à la variable "Nb. internationales" indiquent qu’il y a autant d’équipes qui disposent d’un nombre d’internationales inférieur à 8 que d’équipe avec un nombre supérieur à 8. La valeur minimale étant de 5 et la valeur maximale étant est de 11, ces statistiques décrivent les disparités des compositions des équipes. La Skewness négative suggère par ailleurs que les valeurs extrêmes sont à gauche de la moyenne. Les équipes avec un nombre d’internationales inférieures à la moyenne sont donc légèrement plus fréquentes que les équipes où cette variable est supérieure à la moyenne. La qualité de l’équipe étant mesurée entre autres par le nombre d’internationales dans l’équipe, elle est variante d’une équipe à une autre dans la ligue. Paradoxalement le prix n’est pas nécessairement corrélé à cette variable, par exemple, Aston Villa qui présente en $t = 3$ 5 joueuses internationales dans son équipe de départ propose des billets à £10 minimum, soit un prix compris dans les 25% les plus chers.

¹<https://www.statista.com/statistics/1421104/womens-super-league-ticket-prices/>

	Prix	Nb. internationales	Distance (km)
nobs	121	121	121
NAs	0	0	0
Minimum	6.000	5.000	0.000
Maximum	15.000	11.000	321.670
1. Quartile	7.000	7.000	56.077
3. Quartile	10.000	11.000	267.841
Mean	8.959	8.562	160.966
Median	9.000	8.000	151.985
Sum	1084.000	1036.000	19476.837
SE Mean	0.200	0.170	9.694
LCL Mean	8.563	8.226	141.772
UCL Mean	9.354	8.898	180.159
Variance	4.823	3.482	11370.606
Stdev	2.196	1.866	106.633
Skewness	0.459	-0.066	0.013
Kurtosis	-0.618	-1.088	-1.598

Table 1: Statistiques descriptives

3 Modèle économétrique

3.1 Interprétation des résultats des estimations

La Table 2 présente les résultats des estimations OLS, Within et GLS. Les estimations Within et GLS prennent en compte des effets individuels, dont la pertinence sera justifiée dans la prochaine section. Les résultats d'estimation du modèle OLS correspondent aux attentes théoriques évoquées. L'ensemble des variables relatives à la qualité ex-ante du match exercent une influence positive sur le taux de remplissage du stade. Par exemple, toutes choses égales par ailleurs, le fait que l'équipe extérieure soit Chelsea permet d'augmenter la proportion remplie du stade de 48.1%. Les 4 variables de cette catégorie sont significatives à un seuil d'au moins 5%. À l'inverse, le prix ainsi que la différence de classement entre les deux équipes exercent une influence négative sur la proportion remplie du stade, ces résultats étant toujours en accord avec la théorie.

En revanche, avec les estimations Within et GLS le coefficient associé au prix n'est pas significatif. Cependant, dans ces deux estimations le coefficient associé à la variable binaire week-end est significatif au seuil de 5%. Dans les deux cas, ce coefficient est positif, ce qui signifie que toutes choses égales par ailleurs le fait que le match ait lieu un week-end permet d'augmenter le nombre de places occupées. Ce résultat témoigne du fait que lorsqu'un match a lieu le week-end le coût d'opportunité pour y assister est moindre. La significativité de l'ensemble des variables sur la qualité ex-ante dans l'estimation GLS, révèle que c'est davantage l'aspect spectaculaire et divertissant des matchs qui exerce une influence sur le remplissage des stades. Ces résultats suggèrent qu'une meilleure qualité ex-ante engendre une plus grande incertitude sur le résultat du match. Ces résultats sont dans la continuité de la littérature, et plus particulièrement en accord avec le paradoxe Louis-Schmelling (Neale [1964]) : si les spectateurs peuvent anticiper l'issue du match, la demande de billets sera plus faible. Les résultats suggèrent également que la différence de classement exerce une influence négative sur la proportion du stade remplie : l'augmentation de cette différence de 1 engendre une diminution de 7.7% de la proportion rempli du stade, toutes choses égales par ailleurs. Une grande différence de classement indique qu'une équipe en haut du tableau affronte une équipe en bas du tableau, le résultat est plus facile à prédire ce qui réduit l'attractivité du match. Enfin, la venue des 2 équipes extérieures considérées Chelsea et Arsenal a une forte influence sur le remplissage du stade. En effet, toutes choses égales par ailleurs, un match où l'équipe extérieure est Chelsea aura un stade rempli de 33.4% de plus qu'un autre match.

	Variable dépendante : log(Proportion stade)		
	OLS	WITHIN	GLS
	(1)	(2)	(3)
Prix	-0.080*** (0.026)	-0.034 (0.030)	-0.036 (0.029)
Qualité ex-ante			
Nombre d'internationaux	0.315*** (0.035)	0.061 (0.055)	0.093* (0.051)
Équipe extérieure : Chelsea	0.481*** (0.156)	0.315** (0.147)	0.334** (0.143)
Équipe extérieure : Arsenal	0.323** (0.162)	0.328** (0.146)	0.347** (0.142)
Différence de classement	-0.060*** (0.020)	-0.074*** (0.015)	-0.072*** (0.015)
Coût d'opportunité			
Week-end	0.129 (0.163)	0.253** (0.118)	0.244** (0.116)
Distance en km	-0.0003 (0.001)	-0.0004 (0.0004)	-0.0004 (0.0004)
Performances récentes			
Nb. victoires 3 derniers matchs	0.291* (0.149)	-0.017 (0.133)	0.007 (0.130)
Possible relégation	-0.260* (0.154)	-0.025 (0.134)	-0.032 (0.131)
Constante	-3.328*** (0.377)		-1.880*** (0.586)
Observations	121	121	121
R ²	0.607	0.360	0.353
Adjusted R ²	0.575	0.240	0.301
Residual Std. Error	0.592 (df = 111)		
F Statistic	19.010*** (df = 9; 111)	6.318*** (df = 9; 101)	60.572***

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 2: Résultats

3.2 Test

3.2.1 Test d'existence des effets spécifiques

Afin de justifier le choix d'estimation avec l'effet individuel, le F-test ainsi que le LM test sont effectués pour tester l'existence de l'effet individuel μ_i , de l'effet temporel ϵ_t ou des effets combinés. Ceci revient à effectuer les tests suivants :

$$H_0 : \sigma_{\mu}^2 = 0 \text{ contre } H_A : \sigma_{\mu}^2 \geq 0 \text{ pour l'effet individuel}$$

$$H_0 : \sigma_{\epsilon}^2 = 0 \text{ contre } H_A : \sigma_{\epsilon}^2 \geq 0 \text{ pour l'effet temporel}$$

La Table 3 et la Table 4 confirment l'existence de l'effet individuel au seuil de 1%, ce qui permet de justifier les estimations Within et GLS avec l'effet individuel. Par ailleurs, l'existence de cet effet rend le modèle OLS impertinent, ce dernier ne prenant pas en compte ces effets.

Test	χ^2	df	p-value	Décision
Individuels	67.034	1	$2.669 \times 10^{-16***}$	Rejet de H_0
Temporels	0.098147	1	0.7541	Non rejet de H_0
Combinés	67.132	2	$2.646 \times 10^{-15***}$	Rejet de H_0
Note:			*p<0.1; **p<0.05; ***p<0.01	

Table 3: Tests des multiplicateurs de Lagrange

Test	F	df1	df2	p-value	Décision
Individuels	12.625	10	101	$5.465 \times 10^{-14***}$	Rejet de H_0
Temporels	0.90263	10	101	0.5338	Non rejet de H_0
Combinés	7.3446	20	91	$8.827 \times 10^{-12***}$	Rejet de H_0
Note:			*p<0.1; **p<0.05; ***p<0.01		

Table 4: F-Test

3.2.2 Test d'indépendance d'Hausman

Le test d'Hausman sert à présent à déterminer l'estimateur le plus pertinent entre l'estimateur GLS et Within. Le test est le suivant :

$$H_0 : \mu_i \text{ est indépendant des variables explicatives}$$

$$H_M : \mu_i \text{ est corrélé aux variables explicatives}$$

La Table 5 présente le résultat du test d'indépendance d'Hausman. La p-value n'étant pas significative, l'hypothèse H_0 ne peut pas être rejetée, donc les effets individuels μ_i sont indépendants des variables explicatives. L'estimateur GLS est alors le plus pertinent dans notre analyse.

Test	χ^2	df	p-value	Décision
Hausman	3.2703	9	0.9526	Non rejet de H_0

Table 5: Test de Hausman

4 Conclusion

Ainsi, les résultats obtenus avec l'estimation GLS sont retenus, car c'est l'estimateur le plus pertinent selon les tests effectués. Ces résultats, cohérents avec la littérature existante, suggèrent que les variables permettant d'améliorer la qualité ex-ante des matchs exercent une influence positive sur la demande. Ces résultats sont particulièrement intéressants et il est nécessaire de les prendre en considération pour poursuivre le développement de la discipline. En effet, ils montrent que poursuivre les investissements dans les équipes pour qu'elles deviennent plus compétitives,

notamment en recrutant plus de joueuses internationales, permettrait d'augmenter la demande pour les matchs. Les variables Chelsea et Arsenal ont une influence positive élevée. Ces deux équipes, qui sont les deux équipes les plus titrées du championnat sont les plus compétitives. Mais ce sont aussi deux équipes qui rassemblent les plus grandes stars internationales du football féminin à l'image de Samantha Kerr pour Chelsea ou de Leah Williamson pour Arsenal. La forte influence de ces 2 variables dans les résultats témoigne de l'importance de la réputation de l'effectif, les individus se déplaçant davantage pour voir ces équipes jouer. Ceci témoigne à nouveau de l'efficacité des investissements pour renforcer l'attractivité et la compétitivité des équipes. Généraliser de tels financements à l'ensemble des équipes du championnat permettrait alors d'augmenter la fréquentation des stades.

Quelques pistes d'amélioration peuvent être soulignées. Premièrement, les données utilisées ne prennent en considération que la saison 2022 - 2023. Élargir la période d'étude pourrait permettre d'affiner les résultats. Par ailleurs, le choix de mesurer l'incertitude d'un match par la différence au classement entre les deux équipes qui s'affrontent peut être discutée. Une incertitude à long-terme relative à la domination d'un club dans le temps et une incertitude à court-terme (Sloane [1976]) auraient également pu être prises en compte. L'ajout de variables de contrôle aurait également pu permettre d'affiner l'étude, notamment la prise en compte de la météo lors des jours de match ou encore le montant des budgets des clubs sont des pistes intéressantes.

References

- J. Borland. The demand for australian rules football. *Economic Record*, 63(3):220–230, 1987.
- J. Borland and J. Lye. Attendance at australian rules football: A panel study. *Applied Economics*, 24(9):1053–1058, 1992.
- J. García and P. Rodríguez. The determinants of football match attendance revisited: Empirical evidence from the spanish football league. *Journal of Sports Economics*, 3(1):18–38, 2002.
- N. Jennett. Attendances, uncertainty of outcome and policy in scottish league football. *Scottish Journal of Political Economy*, 31(2):176–198, 1984.
- W. C. Neale. The peculiar economics of professional sports. *The quarterly journal of economics*, 78(1):1–14, 1964.
- P. Sloane. Sporting equality: Labour market versus product market control—a comment. *Journal of Industrial Relations*, 18(1):79–84, 1976.