Journal of the Operational Research Society (2009) 60, 147–155 © 2009 Operational Research Society Ltd. All rights reserved. 0160-5682/09

www.palgrave-journals.com/jors/

Insights for clubs from modelling match attendance in football

B Buraimo[1], D Forrest[2]* and R Simmons[3]

[1]University of Central Lancashire, Preston, UK; [2]University of Salford, Salford, UK; and [3]Lancaster University, Lancaster, UK

The paper employs data from 2884 matches in the English Football League Championship. It builds a model of determinants of attendance designed to yield results relevant to decision-taking at individual clubs. The model has two innovatory features. It controls for the market size of home and away teams precisely by including local population measures constructed from the application of GIS software and information on competition from other clubs. It incorporates these time-invariant covariates in a Hausman–Taylor random effects estimator to take explicit account of variables typically excluded in earlier studies based on fixed effects models. Unlike fixed effects results, Hausman–Taylor estimates permit assessment of the role of market size and quality of the playing squad in determining attendance. Results also quantify the reduction in attendance from televising a match and show that attendance diminishes when a match is played simultaneously with a televised game in a higher status competition.

Journal of the Operational Research Society (2009) **60,** 147–155. doi:10.1057/palgrave.jors.2602549
Published online 9 January 2008

Keywords: sports; management; statistics; regression

1. Introduction

The application of statistical modelling to data from football has increased in recent years and much of the research aims to provide direct guidance to practitioners within the professional game. For example, Hope (2003) asked 'when should you sack the manager?' while Hirotsu and Wright (2002) addressed the question of the optimal timing of substitution of players and other tactical decisions during a game.

Those who take business and marketing decisions at football clubs have generally been less well served by management scientists. It is true that a large pre-existing literature seeks to measure the importance of key determinants of attendances at matches played in professional sports leagues (Borland and Macdonald (2003) provide a comprehensive survey). But, commonly, the focus of studies is on issues concerning public policy. For example, the role of outcome uncertainty is examined to illuminate the debate over whether restrictive practices such as collusive selling of television rights might be justified (leagues defend such revenue sharing practices by claiming that, if matches are played between teams with relatively equal financial, and therefore playing resources, attendances will increase).

Our focus is different as it is designed explicitly to model match attendance in a way that makes the results potentially useful for decision-taking at individual clubs. We argue that the estimator conventionally employed in statistical modelling of attendances is inherently unsuited to the needs of these users and propose an alternative, illustrating its usefulness in the context of a case study of crowd sizes in the English Football League Championship.

The archetypical study in the recent attendance literature (eg García and Rodríguez, 2002, Forrest *et al*, 2004) applies a fixed effects model to panel data describing attendances at each club's sequence of home games during one or more seasons. The dependent variable is crowd size at club i's home game number t. Categorical variables representing each home club control for influences such as varying market size, historical tradition and ticket pricing policy. Additional regressors include variables particular to each match such as the distance between the home and away stadia (to allow for the effect of travel cost on attendance by away fans), indicators of team quality and form (such as the current league positions of the home and away teams) and the day of week and time of year a match takes place. Naturally, a variety of other variables may be included in the specification according to the hypotheses particular authors wish to test.

This fixed effects model is now almost always the technique of choice in match attendance studies. It is able to provide answers to a number of potentially interesting questions. But it has a serious limitation in that, if fixed effects are modelled to capture unobserved heterogeneity in club attributes, one cannot then separate out the impact of those time-invariant club characteristics, such as size of local population and club

*Correspondence: *D Forrest, Centre for the Study of Gambling, Salford Business School, University of Salford, Salford M5 4WT, UK.*
E-mail: d.k.forrest@salford.ac.uk

wage bill (a proxy for team quality) that are in fact observed. Results are therefore less rich than they might be since the bulk of the variation in attendance across matches is then invariably simply attributed to which team happens to host the fixture. The underlying reasons for some clubs having higher intercept terms than others cannot be explored at all given the structure of the model.

This limitation of the fixed effects model is particularly significant from the perspective of individual clubs. For example, they may wish to benchmark their attendances by asking what crowd size should be expected for a club serving their particular size of local market and facing competition from a given number of clubs in the area; or they may wish to predict how many extra tickets they would sell if they increased the budget for player wages. The fixed effects model can give no guidance on questions such as these because the influence of factors like local population, number of rival clubs nearby and level of the club wage bill are simply collected together in the fixed effects, which are like a black box which the standard approach to match attendance modelling cannot open. Nor could such issues be addressed by a season-level analysis across clubs which regressed, for example, club mean attendance on local population, the number of other clubs nearby and wage bill. Wage bill is potentially endogenous because what the club can afford is related to the dependent variable and estimates of the impact of wage bill would then be inconsistent.

We propose the application of the Hausman–Taylor random effects estimator, described in Section 2, to permit isolation of the effects of time-invariant variables such as lcal population, degree of competition from other clubs and club wage bill. We illustrate its use in the context of English football, exploiting GIS software to facilitate very precise measures of market size and competition from other clubs. Further, we also employ the results of estimation to address other issues likely to be of interest to clubs, such as whether and how much attendance is reduced if a match is televised and whether and how much attendance suffers if matches are scheduled in competition with televised international games. Appropriate statistical modelling should be able to inform decisions such as what minimum fee should be required from television rights to compensate for lost ticket revenue and to capture trade-offs, for example between scheduling midweek fixtures on Wednesday (when there might be lost revenue because Champions League football is on television) or Thursday (which leaves less time for player recovery before weekend fixtures).

2. The Hausman–Taylor estimator

Consider a general model in which the dependent variable $\ln y_{it}$ is determined by:

$$\ln y_{it} = \boldsymbol{Z}_i\boldsymbol{\alpha} + \boldsymbol{X}_{it}\boldsymbol{\beta} + \varepsilon_{it} \qquad (1)$$

where the subscript 'i' denotes the cross-sectional unit ($i = 1, 2, \ldots N$), the subscript 't' denotes the time period ($t = 1, 2, \ldots T$), $\boldsymbol{Z}_i$ is a vector of fixed covariates and $\boldsymbol{X}_{it}$ is a vector of time-varying covariates. A *random effects* model of match attendances would assume that the unobserved team-specific effects, $\boldsymbol{Z}_i$, are uncorrelated with the explanatory variables, $\boldsymbol{X}_{it}$, and this is a strong assumption to make. A *fixed effects* model would introduce a set of unit-specific effects (team-specific in our case), θ_i that are time-invariant by construction. But then the roles of $\boldsymbol{Z}_i$ and θ_i will be conflated. The fixed effects model cannot separate impacts of known time-invariant covariates (such as market size or population in our case) from unknown team-specific effects. The Hausman–Taylor estimator will facilitate this separation of effects.

In our case, market size can be proxied by population residing within five miles of the team's stadium and this is time-invariant (since the measure is derived from the Census of Population published once every 10 years). As such, variations in local population across teams will be included in $\boldsymbol{Z}_i$. Now suppose (plausibly) that whether or not a team's match is broadcast live on television depends on its market size. A dummy variable for live broadcast of a match will be included in $\boldsymbol{X}_{it}$ and is most definitely time-varying. This, and any other, correlation between a time-varying covariate and team-specific fixed effect will undermine inferences drawn from the model.

The Hausman–Taylor (1981) estimator of the random effects model proceeds by assuming that some of the covariates are correlated with the unobserved cross-section unit-level *random* effect and uses an instrumental variable method. Here we offer just a brief summary, based upon the fuller explanations given by Baltagi (2005) and Greene (2003).

The format of the Hausman–Taylor model is:

$$\ln y_{it} = \boldsymbol{Z}_{1i}\boldsymbol{\alpha}_1 + \boldsymbol{Z}_{2i}\boldsymbol{\alpha}_2 + \boldsymbol{X}_{1it}\boldsymbol{\beta}_1 + \boldsymbol{X}_{2it}\boldsymbol{\beta}_2 + \theta_i + \varepsilon_{it} \qquad (2)$$

where all team-specific effects in $\boldsymbol{Z}_{1i}$ and $\boldsymbol{Z}_{2i}$ are observed. The team-specific random term, θ_i, contains the unobserved team-specific effects that are included in $\boldsymbol{Z}_i\boldsymbol{\alpha}$ in (1).

The Hausman–Taylor estimator is based upon a four-way classification of observed variables:

$\boldsymbol{Z}_{1i}$ is a vector of K_1 exogenous, time-invariant variables that are not correlated with θ_i,
$\boldsymbol{Z}_{2i}$ is a vector of K_2 endogenous, time-invariant variables that are correlated with θ_i,
$\boldsymbol{X}_{1it}$ is a vector of L_1 exogenous, time-varying covariates that are not correlated with θ_i,
$\boldsymbol{X}_{2it}$ is a vector of L_2 endogenous, time-varying covariates that are correlated with θ_i.

The error terms in the model are assumed to be normally distributed with zero mean and unknown constant variances:

$$\varepsilon_{it} \sim \text{i.i.d } (0, \sigma_\varepsilon^2)$$

$$E[\theta_i] = E[\theta_i|\boldsymbol{X}_{1it}, \boldsymbol{Z}_{1i}] = 0$$

$$E[\theta_i] = E[\theta_i|\boldsymbol{X}_{2it}, \boldsymbol{Z}_{2i}] \neq 0$$

$$\mathrm{Var}[\theta_i|\boldsymbol{X}_{1it}, \boldsymbol{Z}_{1i}, \boldsymbol{X}_{2it}, \boldsymbol{Z}_{2i}] = \sigma_\theta^2$$

$$\mathrm{Cov}[\varepsilon_{it}, \theta_i|\boldsymbol{X}_{1it}, \boldsymbol{Z}_{1i}, \boldsymbol{X}_{2it}, \boldsymbol{Z}_{2i}] = 0$$

$$\mathrm{Var}[\varepsilon_{it} + \theta_i|\boldsymbol{X}_{1it}, \boldsymbol{Z}_{1i}, \boldsymbol{X}_{2it}, \boldsymbol{Z}_{2i}] = \sigma^2 = \sigma_\theta^2 + \sigma_\varepsilon^2$$

The estimator is reliant upon the investigator's ability to distinguish the sets of variables $\boldsymbol{X}_1$ and $\boldsymbol{Z}_1$ that are not correlated with θ_i from those variables making up $\boldsymbol{X}_2$ and $\boldsymbol{Z}_2$ that are. In the case of football match attendances, we shall demonstrate that relevant variables of the type $\boldsymbol{X}_2$ and $\boldsymbol{Z}_2$ do exist. The resulting correlation between variables and random effects renders any OLS or GLS random effects estimation of (1) inconsistent.

Fortunately, a more efficient estimator is available in the form of the Hausman–Taylor instrumental variables estimator. This uses just information already specified within the model. The procedure is to first estimate (2) using the standard fixed effects (within) estimator. The transformation to deviations from team means will remove that part of the error term that is correlated with $\boldsymbol{X}_{2it}$. Hausman and Taylor (1981) show that group deviations from means can be used as $L_1 + L_2$ instrumental variables for the estimation of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ parameters. $\boldsymbol{Z}_1$ is uncorrelated with the error term, by assumption above, and so it can be used as a set of K_1 instrumental variables. Estimation requires a further K_2 instrumental variables. Hausman and Taylor show that the group means for $\boldsymbol{X}_1$ can be used as these additional instruments. The model is identified if $L_1 \geqslant K_2$. That is, the order condition for identification requires that the number of variables in $\boldsymbol{X}_{1it}$ is at least as great as the number of variables in $\boldsymbol{Z}_{2i}$. Feasible generalised least squares (FGLS) of the model will then be an improvement.

In the first stage the within fixed-effects estimator consistently estimates $\boldsymbol{\beta}$ and generates residuals ($\ln y_{it}$ minus predicted values of $\boldsymbol{X}_{it}\boldsymbol{\beta}$). These residuals are regressed on $\boldsymbol{Z}_i$ using a set of time-varying exogenous variables and time-invariant exogenous variables as instruments. This yields intermediate (consistent) estimates of $\boldsymbol{\alpha}$. Both overall and within residuals are obtained. Together, these residuals are used to estimate the components of variance of the dependent variable. The estimated variance components are used to form weights for FGLS estimation in the second stage. More formally, the group means of residuals from fixed effects regression are in turn regressed on $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ with $\boldsymbol{Z}_1$ and $\boldsymbol{X}_1$ deployed as instruments. The residual variance from this IV regression is a consistent estimator of $\sigma^2 = \sigma_\theta^2 + \sigma_\varepsilon^2/T$ where T is number of time periods in the model. The initial fixed effects regression yields an estimator of $\sigma_\theta^2 = \sigma^2 - \sigma_\varepsilon^2/T$. The estimate of $\xi = \sqrt{\sigma_\varepsilon^2/(\sigma_\varepsilon^2 + T\sigma_\theta^2)}$ forms the weight for feasible GLS. The final stage is then a weighted instrumental variable estimator which is efficient, compared to the unweighted IV estimator, although both IV estimators will be consistent.

Clearly, a major advantage of the Hausman–Taylor estimator is that it permits estimation of the impacts of time-invariant covariates in a panel data setting. Beyond this, the estimator economises on use of instruments. All instruments are derived from within the model. These are: $\boldsymbol{X}_{1it}$ and associated means, $\boldsymbol{Z}_{1i}$ and the deviations of $\boldsymbol{X}_{2it}$ from associated means. A search for external instruments, as would be required in fixed-effects models where covariates are potentially endogenous, is not required.

An alternative two-step estimator would first estimate β using the fixed effects estimator and then regress residuals, inclusive of fixed effects, on Z_i. This estimator of the impacts of Z_i is consistent if $E(\theta_i Z_i) = 0$. In contrast, the Hausman–Taylor estimator is robust to $E(\theta_i Z_i) \neq 0$ and is therefore more general (and more informative) than the two-step fixed effects estimator.

The Hausman–Taylor estimator has been applied in several settings. Among the questions addressed have been the impact of schooling on wages (Baltagi and Khanti-Akom, 1990), the impact of health on wages (Contoyannis and Rice, 2001) and the effects of distance on exports and foreign direct investment (Egger and Pfaffermayr, 2004). Here, we apply the Hausman–Taylor estimator to investigate determinants of attendance at football matches.

3. Context and data

The context for our analysis is English professional soccer where 92 clubs compete in four hierarchical divisions linked by a system of promotion and relegation. The top tier is known as The Premier League and the current brand name of the tier below it is the Football League Championship. It is this second tier, which comprises 24 teams, each playing 23 home games per season, that we choose for our case study. It is more amenable to analysis than the top tier because the proportion of sell-out games is so small (1.1% over our study period of seven seasons) that censoring of data raises no serious concerns. By contrast, capacity is filled regularly in the Premier League. While in principle, the tobit estimator is appropriate where some observations of the dependent variable are censored, the solution becomes untenable where certain clubs, as in the Premier League, sell all their seats every match. Further, the legitimacy of tobit estimation for examining attendance at other clubs is brought into question by the industry practice of restricting access to popular (sell out) games to those who have also purchased tickets for less attractive fixtures. Thus one does not observe 'true' demand even at games where the crowd is not capacity constrained. Tobit is incapable of estimating customer response to match characteristics if true demand is not observed at any game.

Of course, aside from estimation issues, modelling attendance in the Football League Championship should be more useful to clubs than would be the case in the Premier League where measures to increase attendance would be redundant at many venues since capacity constraints are already binding. And of course, while our application of the Hausman–Taylor estimator is to a particular level of English football, the approach should be applicable to other countries' football leagues and to competitions in sports such as baseball that have been the focus of earlier published, but we believe flawed, studies.

Our data period extends over several more seasons than has been customary in this literature, covering 7 years from 1997/8 to 2003/4. This period centres on the 2001 Census: because we investigate demographic influences on attendances, we required reliable local population data estimates and judged that population densities during a period covering three years either side of the Census would be adequately represented by the local area Census statistics. Another influence on attendances that we investigate is the televising of matches. Including seven seasons allowed an adequate number of televised games (158) to be included in the sample, so that television impacts could be estimated with relatively high precision.

Not all of the 3864 matches played over the 7-year period were used in the estimation (this number refers to 'regular season' fixtures; the small number of play-off games held at the end of each season to determine the final promoted club are not considered here). We deleted the opening round of matches from each season because two of our control variables required information on previous league form in the current season. There were also 21 cases of a club failing to declare its wage bill for a particular season. Since wage bills were one of our foci of interest, we deleted all observations involving those clubs in those seasons. This left a final sample size of 2884 matches. Attendance across the sample ranged from 3436 to 44135, with the mean 14 988 (standard deviation 7237).

4. Model

We have unbalanced panel data. The cross-sectional unit is a club playing home matches in a particular season (there are 147 such groups). The time unit is the match (observations per group varied, between 18 and 22, because some observations had been deleted due to missing information). The dependent variable is the natural logarithm of attendance. We hypothesize that the size of crowd at a given game will be influenced by factors affecting the size of the market of the home club; factors influencing the number of away supporters who will travel to the game; scheduling issues; the quality of the teams and players on show; and television coverage of the match and of other football taking place at the same time.

Here we give details of the covariates included in the model. Table 1 presents a complete list (with summary statistics), grouped according to whether they are classified as exogenous time-invariant, endogenous time-invariant, exogenous time-varying or endogenous time-varying. A variable is considered to be endogenous if it is a function of other variables within the model. A variable is deemed to be exogenous if it is not determined by other variables within the model but is set externally or is somehow pre-determined. In the Hausman–Taylor model, a variable can be time-invariant and exogenous, in the sense of being independent of other variables in the model, but still correlated with the team specific θ_i terms. An example of such a variable is local population.

4.1. Exogenous time-invariant covariates

A majority of attendees at a game will normally be local supporters and a considerable influence on the size of the crowd will therefore be the size of the market from which the home club draws its customers. Clubs in a large metropolitan area, so long as their advantage is not eroded by competition from other clubs nearby, would be expected to attract larger crowds for a typical match than those based in smaller centres. Some measure of local population should therefore be included in the model.

Using an approach new to the literature, we exploit modern GIS software to measure population within certain distances of the stadium, with the distances defined sufficiently tightly that travel costs from each part of a zone within a club's catchment area will be of the same order of magnitude. Following Forrest *et al* (2002), we defined a club's catchment areas by a radial distance of 10 miles from its stadium and divided this area into two zones, 0–5 and 5–10 miles from the ground, to ensure rough homogeneity of travel costs from each zone. We measured population in each zone at each club, employing 2001 Census microdata for approximately 175 000 Output Areas, and manipulating them using stadia Ordnance Survey map references and the MapInfo software package.

Home club population within 5–10 miles of the ground proved statistically insignificant in our initial estimation of the model and so the model whose results are reported here includes just one population variable for the home club (the natural logarithm of), the population within 5 miles distance of the ground.

In football, tradition is important and support may build up over time because interest is passed between generations. Older clubs may therefore have a larger following. We include as an additional covariate the duration in years of the club's membership of the Football League, relative to 2001.

The fundamentals of geography and history, represented by the population and duration of membership variables, are clearly exogenous. The impact of these variables would be subsumed under team-specific effects in a standard fixed-effects model. The Hausman–Taylor approach adopted here facilitates explicit treatment of market size and league membership.

4.2. Endogenous time-invariant covariates

The impact of population density on crowd support will be mitigated to the extent that a club has to share its market with one or more rivals. We constructed, again using MapInfo software, an index, termed *market overlap*, to measure the degree of competition faced by each club in a more precise way and this also features as a variable in our model. *Market overlap* is the proportion of the catchment area population that also lies within the catchment area of another club. Where

Table 1 Variables employed in estimation

	Sample mean	*Standard deviation*
Dependent Variable		
Match attendance[a]	14,987.75	7,237.40
Exogenous time-invariant		
Duration in years of home club's League membership	95.33	18.39
Population within 5 miles of home club's stadium[a]	442,044	342,016
Endogenous time-invariant		
Market overlap for home club	2.02	2.00
Home club's relative wage	0.999	0.568
Exogenous time-varying		
Derby match[b]	0.012	0.110
Distance in miles between the home grounds of the two clubs[c]	127.42	70.10
Midweek match[b]	0.260	0.439
Bank Holiday fixture[b]	0.067	0.250
October[b]	0.122	0.327
November[b]	0.115	0.319
December[b]	0.119	0.324
January[b]	0.075	0.263
February[b]	0.099	0.299
March[b]	0.128	0.334
April/May[b]	0.154	0.361
Terrestrial TV coverage of European match with English club[b]	0.054	0.226
Population within 5 miles of away club's stadium[a]	444,144	344,112
Duration in years of away club's League membership	95.42	18.37
Endogenous time-varying		
Away team relative wage	1.00	0.569
Market overlap for away club	2.03	2.01
Points per game in season to date (home team)	1.38	0.479
Points per game in season to date (away team)	1.40	0.482
Match shown on ITV[b]	0.001	0.037
Match shown on ITV Digital[b]	0.006	0.077
Match shown on Sky Sports[b]	0.082	0.275

[a]Variable expressed as a natural logarithm in estimation.
[b]Categorical variable.
[c]Variable also entered in squared form in estimation.
Sources: Fixture and attendance information collected or derived from the *Rothmans* and *Sky Sports Football Year Books*. Points per game calculated from League tables. Distances obtained from the RAC. Club wage data from editions of the *Deloitte and Touche* (formerly *Deloitte*) *Annual Review of Football Finance*. Population and overlap measures derived from the 2001 Census (see text). Television coverage from various issues of *TV Sports Markets*.

there is more than one neighbouring club, these intersections of population are aggregated and *market overlap* may then exceed one. Indeed, it often does and the highest value among the clubs here is 7.62 (for Fulham). This measure is clearly time-invariant, as it is constructed from the 2001 Census. But *market overlap* is treated as endogenous because extra clubs may be spawned where population densities are high. The number of overlapping clubs formed will respond to exogenous population density and our overlap measure will then be endogenous. The feedback from local population to market overlap will be captured in our model.

We expect more people to buy tickets for matches when the quality of the two teams is higher. Since the Bosman ruling effectively made players in Europe free agents, the labour market in European football has become competitive and wages for players should therefore reflect talent. Hence we include, for both the home and away team, the club wage bill for the season as a proxy for the quality of its playing squad. The strength of the influence of the home wage budget on attendances will be of particular interest to clubs which are considering spending more on player talent: an anticipated positive effect on attendances might be necessary if the policy is to be affordable.

Employing wage bill data to proxy team quality follows the use of 'budget' as a variable in García and Rodríguez (2002). However, we make an adjustment to account for player wage inflation over the long period described by our data. A club's *relative wage* is its wage bill over a particular season divided by the mean wage bill for the division in that same season. By construction, our wage variable has a mean of one but its range was very wide, from 0.27 to 3.06, reflecting the difference in resources available between clubs aspiring to be promoted to the Premier League and those struggling to avoid relegation. Team quality is endogenous because the resources available to build a squad of players will depend in part on market size.

4.3. *Exogenous time-varying covariates*

In Europe, in contrast to America, distances between clubs in a national league are typically small enough that significant numbers of supporters travel to watch their team playing away. Therefore, it is standard in European attendance models (Dobson and Goddard, 1995) to include distance between clubs as a proxy for travel costs that will influence how many away fans will attend a game. This is customarily entered as a quadratic in the expectation that increasing distance will deter away fans' attendance but at a diminishing rate. We too include *distance* and *distance squared* in our specification. But we innovate by including in addition measures of the size of market from which away support will be drawn. Away market size is proxied by the same variables as in the case of home clubs: population within 5 miles of the stadium, market overlap and duration of League membership.

The season extends from mid-August to early May. We include a series of categorical variables to represent each individual month from October on (April and May are combined as only a small number of fixtures took place in May). Time of year could influence attendance because of weather conditions and competition from alternative sports and activities; but a particular factor identified in previous work is that interest peaks late in the season as many games become significant in the settlement of promotion, playoff and relegation issues. We also include categorical variables to allow for the effects of scheduling on bank holidays or on midweek evenings (where *midweek* refers to any day from Monday to Friday that is not a bank holiday).

Derby is a categorical variable included to identify matches between local or regional rivals. Such games are often played with particular passion and the results may have an importance to supporters independent of their implications for positions in the League. Derby matches are exogenous as they are set by the composition of teams in a division in any season. Our panel unit is a team-season and so *derby* is an exogenous, time-varying covariate.

There may be a risk to attendance if television is showing a game from a higher level of football at the same time as a Football League Championship match is taking place. European matches in the Champions League were transmitted live on many midweek evenings and these will have been competitive with live attendance at matches included in our sample. We include a covariate which takes a value of one if a match was held on the same evening as terrestrial television was relaying a European game featuring an English club. Sky Sports also showed different European fixtures but preliminary estimation did not reveal any statistically significant effects from this source.

4.4. *Endogenous time-varying covariates*

Over this period, the Football League entered into a number of contracts for its television rights and live coverage was variously relayed through three channels: the mainstream, terrestrial, free-to-air ITV; and two subscription services, Sky Sports and the now defunct ITV Digital, accessible through cable and satellite. We measure the impact of a game being televised by a series of categorical variables set equal to one if a match was screened live on the particular television channel. This will permit separate estimation of the effect of telecasting according to whether the platform is free-to-air or pay. However, since a large majority of screenings were on Sky Sports, it is the effects of coverage by subscription television that will be estimated most precisely. Note that there were no examples of matches shown on pay-per-view television where viewing of each event is billed in addition to subscription charges.

There is limited prior knowledge of how readily, if at all, fans are ready to substitute television for live viewing. In their survey, Borland and Macdonald (2003) remark that 'there is not strong evidence on how TV broadcasts affect attendance': the number of studies in the area is relatively small and they do not yield consistency in sign of impact. Borland and Macdonald suspect that the mixed results can 'be attributed to the difficulties in undertaking empirical analysis... One problem is potential joint endogeneity'. The point is underlined by the fact that, in our data set as in others, mean attendance is actually higher for televised than non-televised games. This is likely to be because the characteristics of a match that appeal to the television company also appeal to those going to the stadium. To isolate the impact of television on crowd size, it follows that it is important to include a full set of control variables to capture match characteristics. Without such careful specification, there is a danger that the greater attendance indicated for televised matches in the raw data will be reflected in coefficient estimates on television variables that are biased upwards (ie that underestimate any propensity for home viewing to be substituted for going to the stadium). The decision to screen Football League matches is endogenous as it will be influenced by other match characteristics included as covariates of the model. That is, selection into categories such as *match shown on ITV* will be endogenous since broadcasters will prefer to show matches involving teams with large market size or large team payrolls as indicators of team quality.

Players in a squad may work together more or less successfully than the market value of its players' services might suggest. We therefore include, as additional covariates, actual measures of current season team performance in the form of the points per game that had been won by the home and by the away team in the current season prior to the match taking place (three League points are awarded for a win and one for a draw). The cardinal measure is preferred to the ordinal measure of League position adopted by some authors.

5. Results

Table 2 displays results from weighted random effects IV estimation, using the observed information matrix variant of maximum likelihood. We use the xthtaylor command, implemented as an ado file, in Stata 9.0. This applies adaptive Gauss–Hermite quadrature to compute the log-likelihood and its derivatives. Adaptive quadrature is more flexible than non-adaptive quadrature. Standard errors are derived asymptotically from maximum likelihood estimation. An alternative procedure in R would be to code up the likelihood, use the 'optim' command to obtain the maximum and get the standard errors shown in Table 2 directly from this.

Maintained assumptions of the Hausman–Taylor estimator are $E(\varepsilon_{it}Z_i) = 0$ and $E(\varepsilon_{it}X_{it}) = 0$. Inspection of correlation matrices shows that these conditions are upheld by our model. The order condition for identification clearly holds (number of exogenous time-varying covariates exceeds number of endogenous time-invariant covariates) and so the Hausman–Taylor estimator delivers consistent estimates.

All variables attract signs and significance consistent with prior expectations. For example, attendances build steadily over the season from December on. Using the formula for marginal effect of a change in categorical variable X from zero to one, $e^{\beta X} - 1$, where β is the estimated coefficient, we find that a bank holiday is associated with an estimated 10.4% boost in attendance relative to a normal weekend. In contrast, an estimated 7.0% contraction may be expected from scheduling midweek. 'Derby' games attract substantial extra interest with a predicted 13.9% rise in the size of crowd in addition to effects from the distance variable taking a low value in such cases. Potential supporters respond readily where visiting teams can draw on expensive squads or where home and away teams have performed well through the season.

Table 2 Results from Hausman–Taylor estimation

Dependent variable: natural log of attendance	Coefficient	$\lvert z \rvert$
Exogenous time-invariant		
Duration in years of home club's League membership	0.004	2.22
Natural log of population within 5 miles of home club's stadium	0.401	2.14
Endogenous time-invariant		
Market overlap for home club	−0.159	2.13
Home club's relative wage	0.759	4.14
Exogenous time-varying		
Derby match	0.130	5.10
Distance in miles between the home grounds of the two clubs	−0.002	12.64
Distance squared	0.000005	9.39
Midweek match (not on television)	−0.068	9.06
Bank Holiday fixture	0.099	7.96
October	0.016	1.61
November	0.008	0.79
December	0.032	3.21
January	0.033	2.90
February	0.045	4.29
March	0.057	5.93
April/May	0.108	11.84
Terrestrial TV coverage of European match with English club	−0.051	3.87
Natural log of population within 5 miles of away club's stadium	0.040	5.22
Duration in years of away club's League membership	0.0008	5.56
Endogenous time-varying		
Away club's relative wage	0.085	16.80
Market overlap for away club	−0.016	6.36
Points per game in season to date (home team)	0.040	4.57
Points per game in season to date (away team)	0.035	6.02
Match shown on ITV	−0.212	2.93
Match shown on ITV Digital	−0.083	2.32
Match shown on Sky Sports	−0.047	4.58
Constant	2.96	1.30
Number of observations	2884	
Wald chi-squared (27)	1338.65	

A strong advantage from employing the Hausman–Taylor estimator, rather than a fixed effects estimator, is that it permits evaluation of the contribution of home club market size to attendance. Coefficient estimates on all three indicators of market size are strongly statistically significant and of a magnitude consistent with the centrality accorded the issue by theoreticians who have analysed the way sports leagues are likely to work. With variables initially set equal to their means, a one standard deviation increase in the size of local population is predicted to increase attendance by 3842 and a one standard deviation increase in our index of market overlap is expected to diminish attendance by 3625. These are substantial impacts relative to a mean attendance in the sample of just below 15 000. It would be important for clubs setting targets for ticket sales, to take into account the precise size of their local market as defined here.

The absence of away club market size in previous match level attendance studies is confirmed by the results to constitute an important omission. All three indicators of market size are again significant. The implication is that visiting fans contribute to crowd size and it is insufficient to recognise this merely by the inclusion of a proxy for travel costs.

In conventional fixed effects models, the impact of a club's wage bill, a measure of the quality of its players, on home attendances is subsumed into the fixed effects. The Hausman–Taylor estimator delivers a consistent estimate of the impact since it allows for endogeneity of wages in the attendance equation. The impact is shown to be high in terms of both estimated magnitude and statistical significance. If a club committed 10% more to the wage budget, this would be predicted to raise home attendance by 7.5%. The estimate would again be useful to clubs wishing to benchmark their levels of attendance because from the results of the model they could assess whether crowds were as large as one would expect given the wage budget committed (and market size). They could also use it to forecast the increase in revenue they might hope to gain if they committed to a higher budget. For example, we simulated, for a club spending the division average on wages in season 2002/3, the direct impact of increasing its wage bill to 1.2 times the division average. This would have cost the club £1.76 million. With all other relevant covariates set at their sample means, attendance at a typical home game would be forecast to have increased by 2179, suggesting an increase in admissions just in excess of 50 000 over the season. It would be optimistic to expect the increase in revenue to be sufficient to pay for the extra team quality in these circumstances.

Televising games appears to shrink the sizes of the crowds they attract to the stadium. All three coefficient estimates on the television coverage variables are strongly significant. The point estimate indicates a negative impact of 23.6% in the case of matches shown on free-to-air television, which exceeds the size of any effect reported in previous studies; but the confidence interval is wide because only a small number of games were transmitted on this platform. Of more interest therefore is the impact from subscription television coverage. However, while the effect of these telecasts is well determined in our model, the extent of cannibalization of the live product is limited in magnitude. Sky Sports was the dominant broadcaster of Football League games over our sample period and is currently the sole provider of live telecasts. In the case of Sky Sports, the point estimate implies an impact of −4.8%. This substantially lower adverse impact of live telecasts by the satellite provider compared to the free-to-air terrestrial provider is a strong feature of our results that is new to the literature.

Some supporters do therefore appear to substitute home for stadium consumption of a match when the choice is available. There is also evidence in our results that some potential customers switch to home viewing of a higher status match when it is televised in competition with a Football League Championship fixture at the stadium. The estimate of the impact of a European game (with English involvement) being shown on terrestial television is −5.2%. In experimentation, no significant influence was felt from European games relayed on subscription television. This is further evidence that, at current levels of penetration by paid-for sports channels, mainstream television coverage is potentially much more damaging to football attendance. We note also that the willingness to stay at home to watch high-level football in preference to going out to watch a more routine game might be more pronounced amongst lower divisions tiers than the Football League Championship whose attendances we analyse here. It is in fact quite common for clubs to play on Tuesday or Wednesday evenings when Champions League football is being broadcast. The loss of ticket sales for clubs often already financially pressured is, our findings suggest, large enough that rescheduling might appropriately be considered.

6. Summary and conclusions

We have illustrated an approach that permits evaluation of time-invariant but observable club characteristics when analysing pooled cross-sectional time series data on attendances in a sports league. From it, we were able to present estimates of the substantial importance of local population density and competition from other clubs. For these indicators of market size, we innovated by using GIS technology to derive more precise measures than those previously attempted in the sports literature. We were also able to estimate effects on crowd size where the home club allocates a higher budget to player wages.

The estimator employed also permits assessment of the effects of television coverage. We found strong evidence that broadcast of games on paid-for television channels diminishes attendance at the games shown but only to a limited extent. There was evidence, albeit less precise, of more substantial inroads into the crowd at the stadium if the television medium was free-to-air. We also identified the potential of screenings of games from higher levels of competition in the same sport to detract from attendance.

The results are of direct relevance to individual clubs. The model is capable of being used for setting benchmarks against which current attendances can be evaluated and permits quantitative guidance when clubs take decisions on issues such as when to schedule midweek games and whether to increase spending on player talent.

References

Baltagi BH (2005). *Econometric Analysis of Panel Data*, 3rd edn. John Wiley: Chichester.

Baltagi BH and Khanti-Akom S (1990). On efficient estimation with panel data: An empirical comparison of instrumental variables estimators. *J Appl Econ* **5**: 401–406.

Borland J and Macdonald R (2003). Demand for sport. *Oxford Rev Econ Pol* **19**: 478–502.

Contoyannis P and Rice N (2001). The impact of health on wages: Evidence from the British Household Panel Survey. *Empiric Econ* **26**: 599–622.

Dobson SM and Goddard JA (1995). The demand for professional league football in England and Wales, 1925–92. *Statist* **44**: 259–277.

Egger P and Pfaffermayr M (2004). Distance, trade and FDI: A Hausman–Taylor SUR approach. *J Appl Econ* **19**: 227–246.

Forrest D, Simmons R and Feehan P (2002). A spatial cross-sectional analysis of the elasticity of demand for soccer. *Scottish J Pol Econ* **49**: 336–355.

Forrest D, Simmons R and Szymanski S (2004). Broadcasting, attendance and the inefficiency of cartels. *Rev Ind Org* **24**: 243–265.

García J and Rodríguez P (2002). The determinants of football match attendance revisited: empirical evidence from the Spanish Football League. *J Sports Econ* **3**: 18–38.

Greene W (2003). *Econometric Analysis*, 5th edn. Prentice-Hall: Upper Saddle River, NJ.

Hausman JA and Taylor WE (1981). Panel data and unobservable individual effects. *Econometrica* **49**: 1377–1398.

Hirotsu N and Wright M (2002). Using a Markov process model of an association football match to determine the optimal timing of substitution and tactical decisions. *J Opl Res Soc* **53**: 88–96.

Hope C (2003). When should you sack a football manager? Results from a simple model applied to the English Premiership. *J Opl Res Soc* **54**: 1167–1176.

Received April 2006;
accepted October 2007 after two revisions