Applied Economics

ISSN: 0003-6846 (Print) 1466-4283 (Online) Journal homepage: https://www.tandfonline.com/loi/raec20

Performance and revenue in professional league football: evidence from Granger causality tests

S. M. Dobson & J. A. Goddard

To cite this article: S. M. Dobson & J. A. Goddard (1998) Performance and revenue in professional league football: evidence from Granger causality tests, Applied Economics, 30:12, 1641-1651, DOI: 10.1080/000368498324715

To link to this article: https://doi.org/10.1080/000368498324715

Applied Economics, 1998, **30**, 1641–1651

Performance and revenue in professional league football: evidence from Granger causality tests

S.M. DOBSON and J.A. GODDARD*

Department of Economics, University of Hull, UK and **School of Accounting, Banking and Economics, University of Wales, Bangor, UK*

Using a dataset comprising annual performance (measured by final league position) and gate revenue for 77 Football League clubs which maintained unbroken league membership between 1946 and 1994, the relationship between performance and revenue is investigated using cointegration and causality tests. A cointegrating relationship between performance and revenue is established in only 10 cases out of 77, although it is argued that some caution is required in interpreting these results, due to the low power of the relevant tests in relatively small samples. In Granger causality tests, more evidence is found of causality running from lagged revenue to current performance than of causality in the opposite direction, while the dependence of performance on revenue seems to be greater for the smaller clubs than for the larger. These results lend empirical support to the popular view that, unless checked by mechanisms for revenue redistribution within the league, the natural tendency is for success to become concentrated increasingly among a small group of elite, wealthy clubs.

I. INTRODUCTION

Recent changes in both the financial structure and the organization of professional league football in England and Wales have fuelled popular discussion about the extent to which a relative increase in the financial power of a small group of elite clubs at the top end of the league structure creates a tendency for playing success also to be concentrated increasingly among the same group of clubs.

Such developments include the withdrawal in 1992 of the former division 1 clubs from the Football League to form a break-away Premier League, which since 1992 has been considerably more effective than was the Football League previously in exploiting commercial opportunities in areas such as sponsorship, merchandizing and the sale of television rights. Also significant has been the relaxation of a number of regulations which previously restricted mobility in the players' transfer market, especially at international level. The abolition of the 'three players rule' (which prevented clubs from fielding more than three overseas players at any one time) and the European Court of Justice ruling in the Bosman case (which prevents a club holding the registration of an out-of-contract player from receiving a fee if the player moves to another club in a different EU country) have created greater flexibility in the football labour market, arguably making it easier for the clubs with the most resources to secure the best players simply by paying them more than the rest can afford. However, while the complaint that the odds of success seem to be loaded in favour of the wealthiest clubs may have achieved new impetus as a result of recent developments, there is no doubt that the same complaint has a history as long as that of professional league competition itself.

As pointed out by Davies *et al.* (1995) in a study of rugby league performance and attendance, these considerations are to some extent counter to the approach which has been adopted in most previous econometric studies of the demand for professional team sports. Typically in this literature, a single equation is estimated in which the dependent variable, attendance, is explained by independent variables which, as well as various team performance measures, may include factors such as admission price, income and population. The individual observations are either match attendances (see Peel and Thomas, 1996; Wilson and Sim, 1995;

0003–6846

Baimbridge *et al*., 1996 for recent examples; and Cairns, 1990 for a survey of the earlier literature) or season-by-season average or aggregate attendances (Bird, 1982; Burkitt and Cameron, 1992; Dobson and Goddard, 1995, 1996; Simmons, 1996; Szymanski and Smith, 1997). Implicit in the specification of the empirical model is a presumption that the direction of causality runs from performance to attendance (a successful team will attract more spectators) but not the other way round (a team with large attendances has the resources to attract better players and thereby generate better performance). Davies *et al.* (1995) argue that if reverse causality (from attendance to performance) is present, then at best the dynamics in most previous econometric models have been specified incorrectly, and at worst, misleading policy recommendations may have arisen in some cases. Granger causality tests (Granger, 1969) provide a suitable framework within which such patterns of causality can be investigated, as well as allowing the dynamics of performance and attendance time series to be modelled with considerable flexibility and without needing to impose any prior assumptions (e.g. about lag length and structure) before estimation takes place.

This paper follows the approach developed by Davies *et al.* (1995), using an England and Wales Football League (and Premier League) dataset on season-by-season team performance and gate revenues, which encompasses the period between the 1946–7 and 1993–4 seasons (inclusive). Because the two papers adopt similar methodologies, we now comment in some detail on the two main improvements which we believe this paper offers *vis-a-vis* the Davies *et al.* study.

The first improvement comes by testing for Granger causality between gate revenue and performance (rather than attendance and performance). The use of revenue data offers two advantages over attendance. First, if it is access to financial resources which allows clubs to attract better players so as to improve performance, then revenue is a more appropriate measure of financial strength than attendance. Secondly, if successful performance generates more demand, clubs may respond either by allowing attendances to increase, or by raising admission prices. An attendance series captures the former response but not the latter, whereas a revenue series captures both. Where a club is already operating close to its ground's physical capacity, it is likely to respond to excess ticket demand by raising prices.[1] All of these points suggest that it is more appropriate to test for Granger causality between performance and revenue, rather than attendance. A limitation which remains is that total revenue (from all sources) would be preferable in many respects to gate revenue. However, the former could only be obtained from club company accounts, which are not available over a sufficient time period to permit an analysis of the type implemented here. In any event, it seems likely that revenues from sources such as merchandizing, sponsorship and television are closely correlated with gate revenues for most clubs.

Our second improvement derives from the fact that our dataset is both comprehensive, and more extensive in its duration than that of Davies *et al.* Although we have performance and revenue data for all 103 clubs which competed in the Football League between 1946 and 1994, we have restricted the Granger causality analysis to the 77 clubs which maintained Football (or Premier) League membership continuously during this period. Nevertheless, a cross section of 77 clubs allows us to derive some very much more general conclusions than were possible in the earlier study (which used data on just five major rugby league clubs). Indeed, we do not even attempt to draw conclusions about the dynamic structure of the performance–revenue relationship which are specific to individual clubs in the manner attempted by Davies *et al.* Instead, we search for patterns of Granger causality which appear common across all 77 clubs taken as a whole, or across five groups of clubs with similar characteristics which are used to classify the sample of 77.[2]

The five group definitions are based on three criteria which are distinct from performance or revenue: local population, date of initial entry into the league, and broad geographical location (even though, as shown below, these characteristics are highly correlated with both performance and revenue indicators). Broad group definitions are as follows: Group 1 (G1) comprises major clubs from the largest cities (with populations greater than half million in the 1961 Census) which entered the League before its early-1920s expansion;[3] G2 comprises clubs from other large

[1] Recently, the coincidence of reductions in many ground capacities (resulting from the replacement of standing by seated accommodation) with an increase in spectator demand for attendance has elicited precisely this response from many leading clubs.

[2] We take the view that it is unrealistic to search for strong conclusions about differences between clubs in the dynamic structure of the performance–revenue relationship from the significance or otherwise of individual lagged variable coefficients obtained from estimations using 30 (Davies *et al.*) or even 48 (in the present study) time series observations. This is simply because the relevant tests lack power in such small samples. Like Davies *et al*., we prefer to carry out separate estimations for individual clubs, and therefore do not wish to pool our data cross-sectionally. However, by compiling the results of the individual tests at an aggregate level, we are able to address the problem of low power, at least in an informal manner. So although we do not consider the result of any individual significance test to be conclusive, if a coefficient is significant in say 30 cases out of 77, this certainly indicates a stronger overall effect than if the proportion is 15 out of 77.

[3] In the 1920–21 season, Football League membership increased from 44 clubs to 66 with the incorporation of 22 former Southern League clubs into a new division 3 (south). In 1921–22, a 20 club (subsequently expanded to 22) division 3 (north) was created, increasing total membership to 86. The combined membership of the Premier League and Football League stands currently at 92, after further small adjustments in size in 1923–24 and 1950–51.

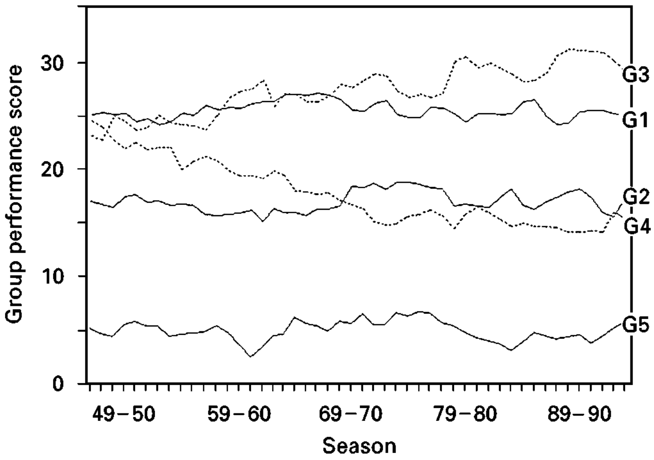

Fig. 1. *Group performance* (% *total score*), *1946–47 to 1993–94*

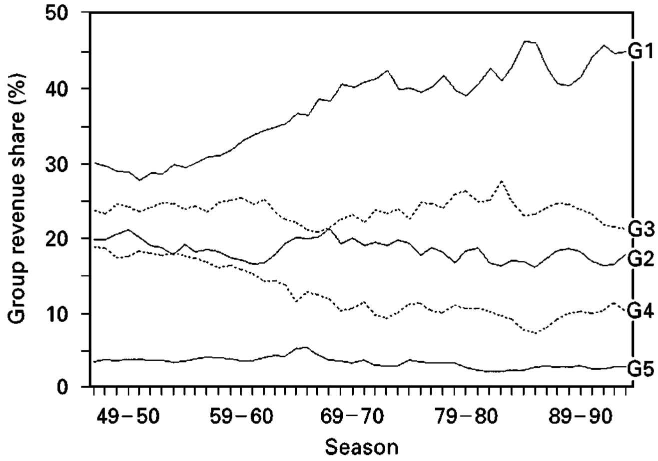

Fig. 2. *Group revenue* (% *share*), *1946–47 to 1993–94*

(around $\frac{1}{4}$ to $\frac{1}{2}$ million) population centres in the Midlands and the North, which entered before the 1920s; G3 comprises all clubs from London, the South of England and South Wales not in G1 (most of which entered during or after the early-1920s expansion); G4 and G5 comprise clubs from the smaller towns in the Midlands and the North which entered before and after the early-1920s expansion respectively. The membership of the five groups is identified in the left hand column of Table 1.

Figures 1 and 2 show performance and gate revenue share measures for the clubs in each of the five groups between the 1946–7 and 1993–4 seasons. For each club, we measure performance in season t, p_t, on a scale of 1 to 92, where $p_t = 92$ indicates the club finished 1st in division 1 (or the Premier League), $p_t = 2$ indicates 2nd, and so on. Each club's revenue measure, r_t, is its percentage share of aggregate league gate revenue. The performance data were obtained from various issues of *Rothmans Football Yearbook*, and the revenue data from the Football League's archives and (for recent seasons) *Digest of Football Statistics*. Figure 1 shows the total performance score for clubs in each of the five groups, expressed as a percentage of the total performance score for all league clubs (including clubs not in the sample). Similarly, Fig. 2 shows the percentage shares of aggregate league gate revenue for clubs in each of the five groups. The levels of the plots in Figs 1 and 2 are influenced by the numbers of clubs in each group (so the more numerous G3 clubs 'outperform' the G1 clubs in Fig. 1). However, our main interest is in the movements of the plots rather than their levels.

The main features of Figs 1 and 2 are as follows. G1 clubs have enjoyed an increase of around 15 percentage points in their revenue share, even though their performance has remained approximately static. G2 clubs have held steady in terms of both indicators. The performance of G3 clubs has improved steadily while their revenue share has been roughly constant. This primarily reflects a catching up process for a number of relatively strong southern clubs, which entered the League relatively late (many in the early-1920s), and which as a group have been consolidating and improving their average position throughout most of the time since. The improved revenue share of G1 and performance of G3 are both mainly at the expense of clubs in G4, which have experienced a marked decline in both indicators. The fortunes of the nine surviving G5 clubs have held reasonably steady, although a pattern of decline would be more evident if G5 were expanded to include other clubs with similar characteristics which have not maintained a continuous League presence.

II. EMPIRICAL TECHNIQUES

In this section, bivarite tests of Granger causality between performance and revenue are carried out for the 77 clubs which enjoyed continuous membership of the Football (or Premier) League between the 1946–7 and 1993–4 seasons inclusive. Multivariate tests (which might include other variables impacting on performance and/or revenue, such as demographic or economic variables at local level) are not considered, since we lack suitable time series data. For each club, the evolution of p_t and r_t (see Section 1) can be represented in its most general form as a vector autoregression (VAR) as follows:

$$\begin{pmatrix} p_t \\ r_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} d_1 t \\ d_2 t \end{pmatrix} + \begin{pmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} \end{pmatrix} \begin{pmatrix} p_{t-1} \\ r_{t-1} \end{pmatrix} + \cdots + \begin{pmatrix} \phi_{11}^{(m)} & \phi_{12}^{(m)} \\ \phi_{21}^{(m)} & \phi_{22}^{(m)} \end{pmatrix} \begin{pmatrix} p_{t-m} \\ r_{t-m} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \quad (1)$$

where u_{it} $(i = 1, 2)$ are random error terms, c_i, d_i and $\phi_{ij}^{(k)}$ $(i, j = 1, 2;\ k = 1 \ldots m)$ are constants and t represents a deterministic time trend. The time trend is included to allow for the long term trends in performance and revenue share identified in Section I among the clubs in some of the groups. However, for some sets of values of the parameters $\phi_{ij}^{(k)}$, the trend becomes redundant and we can set $d_i = 0$ (see below).

Before any tests of the relationship between p_t and r_t can be carried out, two issues concerning Equation 1 need to be addressed. The first is the choice of the number of lagged terms, m, to include on the right hand side of the VAR to capture the dynamics of the process generating p_t and r_t. The second is the question whether the series p_t and r_t are stationary, in which case Equation 1 represents a suitable formulation of the data generating process for purposes of estimation and hypothesis testing, or whether p_t and r_t are nonstationary, in which case it is more appropriate to work with a transformed version of Equation 1, expressed in either difference or error correction form.[4]

The lag length can be determined by comparing estimations of Equation 1 with $m = m_0$ against $m = m_0 - 1$, with $m = m_0 - 1$ against $m = m_0 - 2$, and so on. Additional lagged terms are deleted successively until a point is reached at which further deletions significantly reduce the quality of the fitted model. The likelihood ratio test statistic for the joint significance of the coefficients on the mth lagged terms, $\phi_{ij}^{(m)}$, whose distribution under the null is $\chi^2(4)$, is:

$$\lambda = (T - m)\,\{\ln|\hat{\Omega}_{m-1}| - \ln|\hat{\Omega}_m|\}$$

where T is the total number of time series observations; $\hat{\Omega}_m$ is the (2×2) residual covariance matrix obtained by estimating Equation 1 as a system of two seemingly unrelated regressions with m lags, estimated using the observations for $t = m + 1 \ldots T (\hat{\Omega}_m = \{\hat{\omega}_{ij}\}, \hat{\omega}_{ij} = \Sigma\, \hat{u}_{it}\hat{u}_{jt}/(T - m)$ for $i, j = 1, 2$); and $\hat{\Omega}_{m-1}$ is the covariance matrix from the VAR with $m - 1$ lags also estimated using the observations for $t = m + 1 \ldots T$. Carrying out this procedure for each of the 77 clubs individually, $H_0 : m = 3$ was rejected in favour of H_1: $m = 4$ in 7 cases; $m = 2$ was rejected in favour of $m = 3$ in 4 cases; $m = 1$ was rejected in favour of $m = 2$ in 16 cases; and $m = 0$ was rejected in favour of $m = 1$ in all 77 cases.[5] As it seems desirable to adopt the same specification for all clubs to facilitate comparisons between clubs of subsequent test results, these findings suggest that $m = 2$ is the appropriate lag length for the VAR.[6]

Intuitively, the stationarity or nonstationarity of p_t and r_t hinges on the question whether, for each club, there exist 'equilibrium' levels of performance and revenue share towards which the club has a tendency to revert in the long run, even though it may depart from them (perhaps by a large distance) in the short term; or whether it is possible for a club, as a result of a few seasons of exceptional success or failure, to affect a permanent change to its expected position in the future, without any automatic tendency to gravitate back towards its former status.[7] Technically, the stationarity or nonstationarity of p_t and r_t is important because if these series are nonstationary, likelihood ratio (and other) tests of $H_0 : \phi_{12}^{(k)} = 0$ for $k = 1 \ldots m$ (which if accepted implies lagged revenue does not appear in the performance equation) and $H_0 : \phi_{21}^{(k)} = 0$ (lagged performance does not appear in the revenue equation) do not follow the standard asymptotic distributions, and will therefore give rise to misleading inferences if applied indiscriminately (Hamilton, 1994). This is the case even though the same difficulty does not arise with the tests for lag structure reported above. In order to establish whether adjustment to the model specification is needed before estimation, it is convenient to reparameterize Equation 1 (with $m = 2$) as follows:

$$\begin{pmatrix} \Delta p_t \\ \Delta r_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} d_1 t \\ d_2 t \end{pmatrix} + \begin{pmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{pmatrix} \begin{pmatrix} \Delta p_{t-1} \\ \Delta r_{t-1} \end{pmatrix} + \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} \begin{pmatrix} p_{t-1} \\ r_{t-1} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \quad (2)$$

[4] An issue which we do not pursue in this paper is that the data generating process for p_t is in fact rather more complex than is captured by standard assumptions concerning u_{1t} in Equation 1. For example, p_t is bounded above at 92 and below at 1; this imposes constraints on the distribution of u_{1t} for clubs at or near the top or bottom of the league. Promotion and relegation also imposes constraints on the direction of movement in p_t between certain values in successive seasons. Without attempting any justification, we assume that the impact of these constraints on the tests which follow is sufficiently small to be ignored.

[5] Testing using the bivariate version of Akaike's Information Criterion, which recommends the lag length which minimises $\mathrm{AIC} = \ln|\Omega_m| + (2k^*/T)$, where k^* is the total number of parameters in the VAR with m lags, produces a very similar distribution of results, although with a slight tendency towards the recommendation of less parsimonious specifications than the likelihood ratio test for a few clubs. The equivalent rejection rates using AIC were 14 ($m = 3$ in favour of $m = 4$); 8 ($m = 2$ in favour of $m = 3$); 22 ($m = 1$ in favour of $m = 2$) ; and all 77 ($m = 0$ in favour of $m = 1$).

[6] Seven rejections of H_0: $m = 3$ against H_1: $m = 4$ gives a rejection rate of 0.091, which does not test as significantly different from 0.05 at the 5% level. Therefore, the 7 rejections can plausibly be considered as Type I errors, suggesting that a lag length of $m = 4$ is not necessary. On the same criterion, the 16 rejections of H_0: $m = 1$ against H_1: $m = 2$ are too numerous to be attributable to Type I error, so we conclude $m = 2$ is the appropriate lag length.

[7] Anecdotal evidence can support either view. For example, evidence of stationarity can be found in the case histories of 'big' clubs which have hit hard times and subsequently recovered (e.g. Sheffield Wednesday in the 1970s, Wolverhampton in the 1980s) and 'small' clubs which have temporarily prospered before returning to relative obscurity (e.g. Northampton in the 1960s, Swansea in the late-1970s and early-1980s). On the other hand, the continued success of Wimbledon, which entered the League in 1977 and spent several years in the lower divisions attracting small attendances before a meteoric mid-1980s ascent to division 1 status, argues persuasively in favour of a nonstationarity assumption.

where $\xi_{ii} = -\phi_{ii}^{(2)}$, $\xi_{ij} = -\phi_{ij}^{(2)}$, $\pi_{ii} = \phi_{ii}^{(1)} + \phi_{ii}^{(2)} - 1$, $\pi_{ij} = \phi_{ij}^{(1)} + \phi_{ij}^{(2)}$, and $\Delta p_t = p_t - p_{t-1}$ etc.

Johansen (1988, 1991) has shown that the stationarity or nonstationarity of p_t and r_t, and consequently the choice of specification for purposes of estimation and inference, depends upon whether the parameters $\phi_{ij}^{(k)}$ are such that it is possible to impose restrictions on the rank of the matrix $\Pi(=\{\pi_{ij}\})$ in Equation 2. There are three cases. First, if $\text{rank}(\Pi) = 2$ (so no restrictions are possible), the VAR does not contain a unit root. p_t and r_t are stationary, although they may contain a deterministic time trend. The following (the same as Equation 1 with $m = 2$) then represents the most appropriate specification:

$$\begin{pmatrix} p_t \\ r_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} d_1 t \\ d_2 t \end{pmatrix} + \begin{pmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} \end{pmatrix} \begin{pmatrix} p_{t-1} \\ r_{t-1} \end{pmatrix} + \begin{pmatrix} \phi_{11}^{(2)} & \phi_{12}^{(2)} \\ \phi_{21}^{(2)} & \phi_{22}^{(2)} \end{pmatrix} \begin{pmatrix} p_{t-2} \\ r_{t-2} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \tag{3}$$

The relevant hypotheses for the causality tests are $H_0: \phi_{12}^{(1)} = \phi_{12}^{(2)} = 0$ (lagged revenue does not affect current performance) and H_0: $\phi_{21}^{(1)} = \phi_{21}^{(2)} = 0$ (lagged performance does not affect current revenue). These can be tested using the likelihood ratio statistic $\lambda = (T-2)\{\ln|\hat{\Omega}_R| - \ln|\Omega_U|\}$ where $\hat{\Omega}_R$ and $\hat{\Omega}_U$ are the restricted and unrestricted residual covariance matrices obtained by estimating Equation 3 as a system of two seemingly unrelated regressions. Under each null, λ follows an asymptotic $\chi^2(2)$ distribution.

Secondly, if $\text{rank}(\Pi) = 1$, this implies that the VAR does contain a unit root. p_t and r_t are nonstationary, but it is possible to find a linear combination of the two series, $z_t = p_t - \beta r_t$, which is stationary. In other words, p_t and r_t are cointegrated. The condition $\text{rank}(\Pi) = 1$ implies a (2×1) vector $\alpha = \{\alpha_i\}$ can be found such that

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} (1 - \beta) = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix}$$

The VAR can be written in error correction form as follows:

$$\begin{pmatrix} \Delta p_t \\ \Delta r_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{pmatrix} \begin{pmatrix} \Delta p_{t-1} \\ \Delta r_{t-1} \end{pmatrix} + \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} z_{t-1} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \tag{4}$$

As p_t and r_t are nonstationary, any time trend is captured by the constant terms c_i, so we can set $d_i = 0$. All series in Equation 4 are stationary, so the same likelihood ratio procedure can be used to test H_0: $\xi_{12} = \alpha_1 = 0$ (lagged revenue does not appear in the performance equation) and H_0: $\xi_{21} = \alpha_2 = 0$ (lagged performance does not appear in the revenue equation). This formulation also allows us to distinguish inter-relationships between the series affecting the short run dynamics (by testing $\xi_{12} = 0$ and $\xi_{21} = 0$) from those affecting the long run relationship through the error correction term (by testing $\alpha_1 = 0$ and $\alpha_2 = 0$).

Thirdly, if $\text{rank}(\Pi) = 0$ (i.e. $\pi_{ij} = 0$ for $i, j = 1, 2$), the VAR does contain a unit root, and p_t and r_t are nonstationary. In this case it is not possible to find a stationary linear combination of p_t and r_t, and the data generating process should be expressed in pure difference form as follows:

$$\begin{pmatrix} \Delta p_t \\ \Delta r_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{pmatrix} \begin{pmatrix} \Delta p_{t-1} \\ \Delta r_{t-1} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \tag{5}$$

As before, the presence of a unit root implies the deterministic trend is redundant. Although there is no long run, equilibrium relationship between p_t and r_t, the specification does allow for short run effects. Likelihood ratio tests of H_0: $\xi_{12} = 0$ and H_0: $\xi_{21} = 0$ can be used to establish whether there is causality from revenue to performance, or from performance to revenue, respectively.

Johansen (1988, 1991) also provides a framework to determine which of the three specifications discussed above should be adopted. For a VAR with two equations, this is a two-step procedure as follows:

(1) Test H_0: $\text{rank}(\Pi) = 0$ against H_1: $\text{rank}(\Pi) \geq 1$. If H_0 is accepted, the VAR is diagnosed as nonstationary with no cointegrating relationship, so Equation 5 is adopted. If H_0 is rejected, proceed to step 2. Two alternative tests of H_0, based on the maximal eigenvalue statistic and the trace statistic, are available.
(2) Test H_0: $\text{rank}(\Pi) \leq 1$ against H_1: $\text{rank}(\Pi) = 2$. If H_0 is accepted and provided $\hat{\beta} > 0$, the VAR is diagnosed as nonstationary with a satisfactory cointegrating relationship (in accordance with expectations of a positive relationship between performance and revenue) so Equation 4 is adopted. If H_0 is accepted and $\hat{\beta} < 0$, the VAR is diagnosed as nonstationary with no satisfactory cointegrating relationship, so Equation 5 is adopted. If H_0 is rejected the VAR is diagnosed as stationary, so Equation 3 is adopted.

III. RESULTS

Table 1 shows the diagnoses of the status of the VARs for each of the 77 clubs obtained by following the procedure described in the previous paragraph. Columns 1 and 2 show the results of the tests for H_0: $\text{rank}(\Pi) = 0$ using the maximal eigenvalue statistic and the trace statistic respectively (step 1), and column 3 shows the test for H_0: $\text{rank}(\Pi) \leq 1$ (step 2). Since the maximal eigenvalue statistic and the trace statistic can give contradictory results, two sets of overall diagnoses for the VAR (stationary; nonstationary and cointegrated; nonstationary and not cointegrated) are possible at each significance level. These are reported (at the 5% and 10% levels) in columns 4 to 7. As the results from the two procedures are generally similar, the commentary will refer only to the tests based on the maximal eigenvalue statistic.

Table 1. *Cointegration tests*

	H_0: rank(Π) = 0		H_0: rank(Π) ≤ 1	Diagnosis			
				5% level		10% level	
	ME	Trace		ME	Trace	ME	Trace
GROUP 1							
Arsenal	23.38**	27.62**	4.24**	s	s	s	s
Aston Villa	11.65	18.65**	6.99**	n	s	n	s
Birmingham	9.59	14.60***	5.01**	n	n	n	s
Chelsea	19.56**	24.13**	4.58**	n	n	n	n
Everton	12.58***	19.79**	7.20**	n	s	s	s
Leeds	11.34	13.45***	2.11	n	n	n	c
Liverpool	6.53	7.51	0.98	n	n	n	n
Manchester Utd	14.16**	22.81**	8.14**	s	s	s	s
Manchester City	14.09**	15.21***	1.12	n	n	n	n
Sheffield Utd	12.03	17.18**	5.15**	n	n	n	n
Sheffield Wed	5.91	8.43	2.52	n	n	n	n
Tottenham	17.70**	22.65**	4.96**	s	s	s	s
West Bromwich	15.47**	16.19**	0.72	c	c	c	c
West Ham	5.56	7.57	2.01	n	n	n	n
GROUP 2							
Bradford	8.62	13.66***	5.04**	n	n	n	s
Coventry	6.61	8.52	1.91	n	n	n	n
Derby	14.87**	18.74**	3.87**	s	s	s	s
Hull	10.90	19.40**	8.51**	n	s	n	s
Leicester	9.94	16.41**	6.47**	n	s	n	s
Newcastle	16.82**	29.99**	13.17**	s	s	s	s
Nottm Forest	10.26	15.89**	5.64**	n	s	n	s
Notts County	10.65	14.91***	4.26**	n	n	n	s
Port Vale	10.04	14.17***	4.13**	n	n	n	s
Stoke	13.83***	21.51**	7.68**	n	s	s	s
Sunderland	13.16***	22.13**	8.97**	n	s	s	s
Wolverhampton	24.72**	29.80**	5.08**	s	s	s	s
GROUP 3							
Bournemouth	12.72***	19.01**	6.30**	n	s	s	s
Brentford	6.04	9.78	3.74***	n	n	n	n
Brighton	11.21	14.40***	3.19***	n	n	n	s
Bristol City	10.42	15.51**	5.09**	n	s	n	s
Bristol Rovs	16.20**	18.75**	2.55	c	c	c	c
Cardiff	12.73***	15.20***	2.47	n	n	c	c
Charlton	12.57***	19.64**	7.06**	n	s	s	s
Crystal Palace	15.21**	16.48**	1.27	c	c	c	c
Exeter	9.46	11.23	1.77	n	n	n	n
Fulham	22.33**	24.70**	2.38	c	c	c	c
Ipswich	9.05	13.02	3.97**	n	n	n	n
Leyton Orient	6.74	11.09	4.35**	n	n	n	n
Luton	19.13**	23.82**	4.68**	s	s	s	s
Millwall	9.13	11.73	2.60	n	n	n	n
Northampton	13.28***	17.72**	4.44**	n	s	s	s
Norwich	15.63**	20.10**	4.47**	s	s	s	s
Plymouth	11.86	15.89**	4.03**	n	s	n	s
Portsmouth	9.37	10.78	1.41	n	n	n	n
QPR	11.81	16.31**	4.50**	n	n	n	n
Reading	6.96	9.30	2.34	n	n	n	n
Southampton	10.97	14.02***	3.05***	n	n	n	s
Southend	11.84	17.24**	5.40**	n	s	n	s
Swindon	13.90***	15.74**	1.84	n	c	c	c
Torquay	37.42**	39.78**	2.36	c	c	c	c
Watford	13.42***	16.66**	3.24***	n	c	s	s

Table 1. (*continued*)

	H_0: rank(Π) = 0		H_0: rank(Π) ≤ 1	Diagnosis			
				5% level		10% level	
	ME	Trace		ME	Trace	ME	Trace
GROUP 4							
Barnsley	17.44**	21.77**	4.32**	s	s	s	s
Blackburn	6.81	9.41	2.60	n	n	n	n
Blackpool	10.49	10.58	0.09	n	n	n	n
Bolton	15.78**	18.55**	2.77***	c	c	s	s
Burnley	13.93***	14.46**	0.53	n	n	c	c
Bury	12.16***	13.11	0.95	n	n	c	n
Crewe	9.64	14.91***	5.27**	n	n	n	s
Doncaster	9.62	12.53	2.91***	n	n	n	n
Grimsby	25.69**	28.10**	2.41	c	c	c	c
Huddersfield	10.86	15.41***	4.56**	n	n	n	s
Middlesbrough	23.09**	25.50**	2.41	c	c	c	c
Oldham	17.83**	28.20**	10.38**	s	s	s	s
Preston	8.15	10.19	2.04	n	n	n	n
Rotherham	14.16**	15.69**	1.53	c	c	c	c
Stockport	12.94***	16.60**	3.66***	n	c	s	s
GROUP 5							
Carlisle	11.90	16.70**	4.80**	n	s	n	s
Chester	11.15	14.36***	3.21***	n	n	n	n
Chesterfield	16.13**	19.40**	3.27***	c	c	s	s
Hartlepool	20.26**	23.86**	3.59***	n	n	n	n
Mansfield	14.06***	18.36**	4.30**	n	s	s	s
Rochdale	29.26**	36.29**	7.09**	s	s	s	s
Swansea	21.91**	31.46**	4.55**	s	s	s	s
Tranmere	15.06**	20.89**	5.84**	s	s	s	s
Walsall	8.26	12.54	4.29**	n	n	n	n
Wrexham	13.11***	18.58**	5.47**	n	s	s	s
York	17.68**	27.34**	9.66**	s	s	s	s
Overall summary (number of diagnoses)							
		stationary (3)		14	30	26	42
		cointegration (4)		10	13	12	12
		non-stationary (5)		53	34	39	23

Note: Columns 1 to 3 show the results of tests on rank (Π) in (2). Columns 4 to 7 show the diagnoses produced by the tests, using the maximal eigenvalue and trace statistics, at the 5% and 10% significance levels.
ME = Maximal Eigenvalue Statistic; Trace = Trace Statistic.
** = Significant at 5% level; *** = Significant at 10% level.
s = Stationary diagnosis; c = nonstationary/cointegration diagnosis; *n = nonstationary/no cointegration diagnosis.

At the 5% level, in 53 of the 77 cases, the VAR is diagnosed as nonstationary with no cointegrating relationship. Performance and revenue are diagnosed as nonstationary but cointegrated in 10 cases, and in a further 14 cases the VAR is diagnosed as stationary. Only one of the ten cases for which cointegration is diagnosed occurs among the 26 clubs in G1 and G2, whereas there are nine diagnoses of cointegration among the 51 smaller clubs in G3, G4 and G5. A link between performance and revenue therefore seems to be more evident for the smaller clubs than for the larger.

Although the overall proportions of diagnoses of cointegration and stationarity are relatively small, both require the rejection of H_0: rank (Π) = 0 (see Section II) and cointegration also requires the acceptance of H_0: rank(Π) ≤ 1. In view of the relatively low power of most unit root and cointegration tests in small samples (see e.g. Maddala, 1992 for general discussion and Kremers *et al*., 1992 for a Monte Carlo analysis) the evidence against cointegration (or stationarity) is less persuasive than would be apparent from a superficial reading of these results. Twenty-four rejections of H_0: rank(Π) = 0 out of 77 is considerably more than could plausibly be attributed to Type I error at the 5% level, and there must also be a certain (unknown) proportion of Type II errors among the 53 acceptances.

In view of the results shown in Table 1, it seems difficult to argue on empirical grounds that a common specification for

Table 2. *Granger causality tests*

Equation	3		4						5		Preferred specification	Preferred specification	
Direction of causality	p to r	r to p	p to r			r to p			p to r	r to p		p to r	r to p
Test statistic	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}			
GROUP 1													
Arsenal	2.55	4.39	0.29	2.07	2.29	9.53*	11.02*	16.39*	0.22	5.36**	(3)		
Aston Villa	3.85	6.96**	1.65	5.36**	6.47**	0.73	3.45***	6.67**	1.11	3.22***	(5)		
Birmingham	5.31***	5.32***	2.72***	6.35**	7.92**	0.82	0.07	0.84	1.57	0.77	(5)		
Chelsea	12.15*	4.47	16.48*	14.77*	22.43*	7.78*	9.50*	13.43*	7.65*	3.93**	(5)	yes	yes
Everton	2.01	3.11	0.02	2.91***	3.44	0.10	2.79***	3.97	0.53	1.18	(5)		
Leeds	2.41	7.17**	0.19	4.30**	4.65***	0.03	11.29*	12.62*	0.35	1.33	(5)		
Liverpool	0.92	5.09***	0.29	0.15	0.35	0.17	6.39**	6.54**	0.20	0.15	(5)		
Manchester Utd	0.29	1.59	2.32	5.13**	5.38***	0.31	14.15*	14.73*	0.25	0.58	(3)		
Manchester City	0.98	1.21	2.98***	1.54	3.14	1.20	14.08*	14.10*	1.60	0.02	(5)		
Sheffield Utd	9.99*	1.52	10.58*	10.38*	17.74*	3.29***	4.74**	5.95***	7.36*	1.21	(5)	yes	
Sheffield Wed	1.22	9.85*	2.25	0.06	2.25	4.20**	2.67	11.09*	2.19	8.41*	(5)		yes
Tottenham	0.29	3.69	0.96	2.26	2.36	0.74	17.61*	17.66*	0.10	0.05	(3)		
West Bromwich	2.38	1.64	0.85	14.10*	15.64*	1.88	2.13	2.82	1.54	0.68	(4)	yes	
West Ham	0.10	0.04	0.26	0.70	0.77	0.02	5.38**	3.24	0.07	0.02	(5)		
GROUP 2													
Bradford	2.06	9.30 *	2.03	5.19**	5.26***	3.55***	0.11	2.12	0.07	4.98**	(5)		yes
Coventry	0.30	5.84***	0.36	3.51***	4.15	1.55	0.41	8.54**	0.64	2.18	(5)		
Derby	5.19***	8.30**	4.46**	3.12***	5.33***	0.23	14.66*	14.99*	2.21	1.21	(3)		yes
Hull	2.27	2.06	0.37	6.60**	6.80**	0.00	8.94*	0.93	0.20	0.61	(5)		
Leicester	3.47	7.13**	3.19***	1.10	3.39	8.40*	9.35*	1.08	2.28	3.78***	(5)		
Newcastle	2.47	7.56**	0.05	0.07	0.10	0.46	13.57*	0.26	0.03	5.32**	(3)		yes
Nottm Forest	0.76	0.07	0.56	4.27**	4.31	0.04	0.42	5.39***	0.05	0.01	(5)		
Notts County	1.18	4.96***	0.02	5.18**	5.33***	0.12	0.11	10.65*	0.14	0.16	(5)		
Port Vale	0.42	9.92*	0.06	0.05	0.08	6.18**	8.37*	18.47*	0.03	7.08*	(5)		yes
Stoke	3.38	5.66***	0.60	3.05***	3.77	0.03	3.07***	5.21***	0.72	1.05	(5)		
Sunderland	6.59**	2.02	1.33	8.21*	9.68*	0.33	0.47	2.59	1.47	0.10	(5)		
Wolverhampton	0.78	4.78***	0.42	2.15	2.47	0.04	5.54**	0.41	0.32	1.38	(3)		
GROUP 3													
Bournemouth	3.72	5.48***	2.75***	11.68*	13.19*	6.90*	2.58	14.58*	1.51	4.79**	(5)		yes
Brentford	8.39**	0.41	1.98	5.84**	8.66**	0.01	3.11***	3.69	2.81***	0.13	(5)		
Brighton	4.84***	7.97**	0.03	5.40**	6.76**	0.08	12.32*	5.09***	1.36	3.45***	(5)		
Bristol City	0.39	7.19**	0.23	10.39*	10.85*	9.76*	3.05***	15.87*	0.46	7.60*	(5)		yes
Bristol Rovs	0.82	3.70	0.03	4.14**	4.41	0.63	15.96*	9.55*	0.27	2.51	(4)		yes
Cardiff	0.71	7.50**	0.71	3.16***	3.52	2.04	1.58	13.13*	0.37	3.63***	(5)		
Charlton	4.03	4.51	1.28	2.36	4.67***	1.13	1.48	18.89*	2.31	0.64	(5)		
Crystal Palace	3.43	6.92**	0.19	14.68*	15.20*	6.34**	6.21**	0.43	0.52	2.33	(4)	yes	
Exeter	1.21	5.98***	0.03	4.83**	6.24**	5.93**	8.38*	0.27	1.41	6.25**	(5)		yes
Fulham	2.66	5.74***	2.61	0.24	2.91	2.49	14.66*	15.44*	2.67	0.33	(4)		yes
Ipswich	3.35	2.47	0.32	7.51*	8.04**	0.64	2.05	4.12	0.52	0.11	(5)		
Leyton Orient	4.25	0.41	1.50	1.16	2.56	0.01	0.81	0.57	1.40	0.12	(5)		
Luton	6.61**	1.16	2.08	9.39*	11.07*	0.41	0.10	6.91**	1.68	0.31	(3)	yes	
Millwall	0.09	0.30	0.00	8.68*	8.69**	0.51	0.95	13.05*	0.01	0.13	(5)		
Northampton	1.68	8.56**	1.08	0.31	1.13	0.04	8.17*	3.72	0.82	5.02**	(5)		yes

Norwich	1.62	1.30	1.64	5.44**	6.02**	0.42	0.42	8.58**	0.59	0.97	(3)		
Plymouth	6.91**	4.61***	0.65	14.66*	16.80*	0.55	11.63*	4.93***	2.14	2.95***	(5)		
Portsmouth	0.08	3.32	0.55	1.94	2.25	0.00	2.59	2.44	0.31	0.57	(5)		
QPR	1.65	0.09	0.14	7.25*	8.04**	0.00	0.22	11.39*	0.79	0.04	(5)		
Reading	3.33	1.29	6.52**	7.35*	9.60	1.85	9.66*	5.35***	2.25	1.21	(5)		
Southampton	0.65	2.54	0.08	1.32	1.81	0.00	3.18***	12.26*	0.49	0.51	(5)		
Southend	1.20	5.25***	2.89***	0.09	3.06	3.05***	1.80	23.84*	2.97***	6.18**	(5)		yes
Swindon	1.04	10.04*	0.02	4.25**	4.91***	4.62**	0.36	17.68*	0.66	7.26*	(5)		yes
Torquay	3.69	12.28*	5.41**	6.35**	7.04**	0.03	24.35*	2.46	0.69	2.42	(4)	yes	
Watford	2.57	11.47*	0.39	0.14	0.97	0.94	5.12**	1.65	0.83	6.18**	(5)		yes
GROUP 4													
Barnsley	1.06	15.11*	0.29	5.84**	5.85***	6.10**	3.11***	5.97***	0.01	9.94*	(3)		yes
Blackburn	2.19	5.90***	0.09	0.05	0.21	0.04	3.49***	12.60*	0.16	0.23	(5)		
Blackpool	2.10	1.76	0.30	0.11	0.34	1.03	8.49*	1.38	0.23	0.09	(5)		
Bolton	0.79	6.09**	0.00	1.67	1.70	0.81	2.16	7.87**	0.03	2.77***	(4)		yes
Burnley	2.86	2.46	1.85	8.63*	10.31*	2.33	0.02	6.58**	1.68	2.42	(5)		
Bury	3.65	2.26	1.60	4.60**	4.77***	0.70	11.13*	7.37**	0.18	0.26	(5)		
Crewe	3.64	6.07**	0.31	4.53**	6.53**	2.27	7.64*	15.77*	2.00	4.62**	(5)		yes
Doncaster	3.76	18.83*	0.06	3.20***	3.32	3.27***	8.27*	14.63*	0.12	15.56*	(5)		yes
Grimsby	4.80***	13.22*	1.71	2.21	7.37**	3.33***	5.13**	2.16	5.16**	12.55*	(4)	yes	
Huddersfield	6.95**	2.74	0.09	13.04*	13.35*	2.39	0.38	13.19*	0.31	2.08	(5)		
Middlesbrough	2.62	3.38	0.50	10.32*	11.13*	0.71	0.36	1.39	0.81	1.29	(4)	yes	
Oldham	1.63	3.99	1.83	4.80**	6.35**	5.67**	0.25	3.16	1.55	5.73**	(3)		
Preston	7.57**	3.22	0.01	2.86***	3.35	1.44	12.31*	10.87*	0.49	0.29	(5)		
Rotherham	1.13	0.04	0.09	3.03***	4.60	0.94	0.98	7.98**	1.56	0.40	(4)		yes
Stockport	0.68	8.71**	2.44	0.63	2.45	2.85***	1.61	8.62**	1.82	6.26**	(5)		yes
GROUP 5													
Carlisle	10.71*	0.30	5.92**	7.73*	13.45*	0.01	0.11	7.62**	5.71**	0.00	(5)	yes	
Chester	7.27**	0.49	6.92*	16.79*	18.31*	2.02	6.49**	26.77*	1.52	0.43	(5)		
Chesterfield	0.47	0.99	0.27	0.02	0.40	0.01	4.38**	11.29*	0.38	0.97	(4)		yes
Hartlepool	0.07	2.37	0.58	5.92**	6.12**	0.00	11.82*	0.11	0.20	0.25	(5)		
Mansfield	6.86**	7.92**	7.10*	12.39*	12.91*	0.07	16.62*	6.92**	0.52	3.59***	(5)		
Rochdale	1.62	5.70***	0.21	3.35***	5.61***	0.55	10.26*	12.07*	2.26	5.94**	(3)		
Swansea	3.05	3.00	1.92	0.17	1.97	0.04	7.17*	20.21*	1.80	1.45	(3)		
Tranmere	3.87	5.06***	0.61	2.71***	2.71	0.02	6.73*	16.20*	0.00	0.99	(3)		
Walsall	3.58	0.46	4.27**	11.96*	13.17*	1.11	6.26**	7.72**	1.21	0.32	(5)		
Wrexham	3.24	7.32**	6.40**	2.24	6.93**	2.39	2.47	8.43**	4.69**	5.96**	(5)	yes	yes
York	0.47	7.22**	3.10***	8.05*	8.69**	7.43*	0.55	7.90**	0.64	7.35*	(3)		yes
Overall summary Number of cases significant at:													
1% level	3	9	4	21	17	6	29	31	2	8			
5% level	11	25	10	41	34	13	39	46	5	22			
10% level	15	39	16	51	44	19	47	53	7	29			

Notes: Columns 1 to 10 show the results of Granger causality tests based on Equations 3, 4 and 5, for lagged performance in the revenue equation (p to r) and lagged revenue in the performance equation (r to p).

Column 11 shows the preferred specification (out of Equations 3, 4 and 5), diagnosed by the cointegration tests using the maximal eigenvalue statistic and a 5% significance level, as reported in Table 1.

Columns 12 and 13 identify cases in which Granger causality is diagnosed in the preferred specification at the 5% significance level (using λ_5 and λ_8 in cases where Equation 4 is the preferred specification).

λ_j are likelihood ratio statistics (distributed $\chi^2(r)$ where r = number of restrictions) for the following null hypotheses:

λ_1: $\phi_{12}^{(1)} = \phi_{12}^{(2)} = 0$ in Equation 3 — λ_2: $\phi_{21}^{(1)} = \phi_{21}^{(2)} = 0$ in Equation 3 — λ_3: $\xi_{12} = 0$ in Equation 4 — λ_4: $\alpha_1 = 0$ in Equation 4

λ_5: $\xi_{12} = \alpha_1 = 0$ in Equation 4 — λ_6: $\xi_{21} = 0$ in Equation 4 — λ_7: $\alpha_2 = 0$ in Equation 4 — λ_8: $\xi_{21} = \alpha_2 = 0$ in Equation 4

λ_9: $\xi_{12} = 0$ in Equation 5 — λ_{10}: $\xi_{21} = 0$ in Equation 5

*significant at 1% level **significant at 5% level ***significant at 10% level the VAR (Equations 3, 4 or 5) should be used in the Granger causality tests for all clubs. On the other hand, lacking any strong theoretical argument as to why the VAR should be nonstationary or cointegrated for some clubs and stationary for others, the use of different specifications for different clubs also seems questionable. Basing the Granger causality tests solely on the specification diagnosed individually for each club is likely to imply specification error in a large number of cases, due to Type I and II error as discussed above.

Our approach is to remain agnostic on the issue of model specification. In columns 1 to 10 of Table 2, we report three complete sets of Granger causality tests, using Equations 3, 4 and 5 respectively as a common specification for all 77 clubs. For Equation 4, the tests for Granger causality affecting the short run dynamics, the error correction term, and both together, are shown separately. Column 11 shows the 'preferred' specification (Equations 3, 4 or 5) diagnosed using the maximal eigenvalue statistic at the 5% significance level (as in Table 1). Columns 12 and 13 identify cases in which the tests for Granger causality are positive in the preferred specification (again, testing at the 5% level). If similar patterns are evident in the three original sets of results, and in those compiled using the preferred specifications, this should allow us to draw inferences about the patterns of Granger causality, without needing to commit ourselves either to a common VAR specification for all 77 clubs, or to different specifications for individual clubs or groups of clubs.

Comparing the results of the Granger causality tests shown in Table 2, it is apparent that the null of no Granger causality is rejected more frequently using the error correction specification Equation 4 than with either Equations 3 or 5. This is because with Equation 4, the null implies the deletion from one of the equations of the entire error correction term, which is a more severe restriction than those tested under Equations 3 and 5. Nevertheless, drawing comparisons between the results within each set of tests, a consistent pattern does emerge. The number of cases in which a relationship is detected between lagged revenue and current performance is almost always greater (and usually much greater) than the number in which a relationship is detected between lagged performance and current revenue. The same pattern is evident when the results from the preferred specifications for each club are compared, where there are only ten diagnosies of Granger causality from lagged performance to current revenue, but 24 diagnoses from lagged revenue to current performance.

Comparing the 'preferred specification' results between the five groups of clubs, the number of diagnosies of Granger causality from revenue to performance appears to be proportionately smaller among the 'big city' clubs of G1 than among the other four groups. Using the maximal eigenvalue statistic and a 5% significance level, in G1 there are only two rejections of the no Granger causality null out of 14, a proportion of 14.3%. For G2, G3, G4 and G5 combined, there are 22 rejections out of 63, or 34.9%. In recent years, this might reflect the lesser dependence of the larger clubs on gate receipts, given their greater potential to attract television revenues, sponsorship and other sources of income; however, it is not entirely clear whether this sort of explanation would suffice for the post-war period as a whole. Other than this, there appears to be little evidence of marked differences among the other four groups of clubs in the pattern of acceptances and rejections of the null in the various formulations of the Granger causality tests.

IV. CONCLUSIONS

The following general conclusions can be drawn from the analysis. First, after classifying clubs according to characteristics other than performance and revenue, it is clear that different types of club have had markedly different experiences as regards both performance and revenue since the 1940s. Specifically, the major clubs from the largest cities have enjoyed a large increase in their percentage share of revenue which is not explained by a corresponding improvement in performance. Southern clubs (other than the major London clubs) have enjoyed a steady improvement in performance but no corresponding trend in market share. Correspondingly, the relative position of the older clubs from smaller towns in the Midlands and North has deteriorated in terms of both performance and revenue.

This evidence of heterogeneity among clubs may in part explain the diversity in the results obtained when the 77 club-specific VARs are tested for stationarity, nonstationarity and cointegration, and for Granger causality between performance and revenue. However, a recurrent argument of this paper has been that the lack of power of the relevant tests (even with time series comprising 48 annual observations) could quite easily account for the differences in test results between clubs. In view of this, for purposes of presentation and interpretation of the results, we prefer to look for patterns in results aggregated across clubs, rather than attempt to draw strong inferences from individual results.

Despite these caveats, several conclusions do emerge from the empirical analysis. Most importantly, there is strong evidence that the influence of lagged revenue on current performance is greater than the influence of lagged performance on current revenue. This complements a similar finding by Davies *et al.* (1995) concerning performance and attendance, based on a much smaller rugby league dataset. Somewhat ironically, our results also suggest that the link between lagged revenue and current performance may be weaker for the wealthiest clubs than for the rest; i.e. the clubs which draw the largest attendances and charge the highest admission prices are also those whose future success depends least on their current gate revenues.

Limited evidence is found to support the notion that our performance and/or revenue measures are stationary, or if not, that a long run cointegrating relationship exists between them. Evidence of cointegration appears stronger for the smaller clubs (G3, G4 and G5) than for those from the largest cities or other major provincial centres (G1 and G2). However, although the sample proportions of stationary or cointegration diagnosies are too high to be attributable purely to Type I error, we are unable to reject null hypotheses of non-stationarity without cointegration for a majority of clubs in all groups.

Observers who prefer sporting competition to be meritocratic may be disappointed in (but perhaps unsurprised by) these results, which lend empirical support to the popular notion that the chances of success are loaded in favour of the wealthiest clubs. Although we have examined only the relationship between league gate revenues and performance, it seems plausible to suggest that a continuation of the current tendency for the leading clubs (in particular) to augment receipts from ticket sales with additional revenue from other sources can be expected to lead to a further concentration of playing success among a small group of elite clubs in the future. A possible policy implication for the football authorities is that to preserve some degree of competitive balance within the league, there will be a continued and perhaps enhanced need for revenue redistribution between clubs, through mechanisms such as the transfer system (the future of which looks uncertain, however, in the light of the 1995 Bosman ruling) or the pooling of television and sponsorship proceeds.

REFERENCES

Baimbridge, M., Cameron, S. and Dawson, P. (1996) Satellite television and the demand for football, *Scottish Journal of Political Economy*, **43**(3), 317–33.

Bird, P. (1982) The demand for league football, *Applied Economics*, **14**(6), 637–49.

Burkitt, B. and Cameron, S. (1992) Impact of league reconstruction on team sport attendances: the case of rugby league, *Applied Economics*, **24**(2), 265–71.

Cairns, J. A. (1990) The demand for professional team sports, *British Review of Economic Issues*, **12**(28), 1–20.

Davies, B., Downward, P. and Jackson, I. (1995) The demand for rugby league: evidence from causality tests, *Applied Economics*, **27**(10), 1003–07.

Digest of Football Statistics, various issues, The Football Trust, Leicester.

Dobson, S. and Goddard, J. (1995) The demand for professional league football in England and Wales, 1925–92, *Journal of the Royal Statistical Society Series D, The Statistician*, **44**(2), 259–77.

Dobson, S. and Goddard, J. (1996) The demand for football in the regions of England and Wales, *Regional Studies*, **30**(5), 443–53.

Granger, C. W. J. (1969) Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*, **37**(1), 24–36.

Hamilton, J. (1994) *Time Series Analysis*, Princeton, New Jersey.

Johansen, S. (1988) Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control*, **12**(2/3), 231–54.

Johansen, S. (1991) Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregression models, *Econometrica*, **59**(6), 1551–80.

Kremers, J. J. M., Ericsson, N. R. and Dolado, J. J. (1992) The power of cointegration tests, *Oxford Bulletin of Economics and Statistics*, **54**(3), 325–48.

Maddala, G. S. (1992) *Introduction to Econometrics*, 2nd edn, Macmillan, New York.

Peel, D. and Thomas, D. (1996) Attendance demand: an investigation of repeat fixtures, *Applied Economics Letters*, **3**(6), 391–94.

Rothmans Football Yearbook, various issues, Headline, London.

Simmons, R. (1996) The demand for English league football: a club level analysis, *Applied Economics*, **28**(2), 139–55.

Szymanski, S. and Smith, R. (1997) The English football industry: performance, profit and industrial structure, *International Review of Applied Economics*, **11**(2), 135–54.

Wilson, P. R. D. and Sim, B. (1995) The demand for semi-pro football in Malaysia 1989–91: a panel data approach, *Applied Economics*, **27**(1), 131–38.