Measuring Habit Persistence Effects in Attendance at Professional Team Sports Encounters: a cautionary note

Alistair Dawson and Paul Downward[1]

Abstract

The authors argue that the standard procedure for capturing habit persistence in empirical models of attendance at professional team sports is deficient. A mis-specification traceable to an early (influential) paper appears to affect subsequent work, casting some doubt on its correct interpretation.

1. Introduction

THERE IS A burgeoning literature on the demand for professional team sports. Downward and Dawson (2000) provide the most recent survey of the literature updating a previous paper by Cairns (1990). An important issue identified is the importance of long-run determinants of demand. These had previously been neglected in the literature. Thus, recent long-run studies of attendance at professional association football stress (rightly we believe) the importance of tradition and cultural factors (for example Dobson and Goddard, 1995; and Simmons 1996). In economic language, the influence of these factors can be described as habit persistence. However, while such authors accept the necessity of incorporating habit persistence in their analyses, there is little discussion of the identification problems associated with this endeavour.

While not the sole method of modelling habit persistence, since Borland's (1987) locus classicus of the genre, the incorporation of a lagged dependent variable in regression analysis seems to be universal practice. Thus attendance (A) at professional team sports encounters is modelled by adding lagged attendance to an (attendance) demand function such as

$$A_t = b_1 + b_2P_t + b_3A_{t-1} \qquad [1]$$

ignoring explanatory variables other than price (P) and for ease of exposition sticking at one lag and ignoring the error term. The estimate of b_3 is then interpreted as a measure of habit persistence, although it looks suspiciously like a reduced form parameter.

It is unusual for a writer to justify the procedure, Borland (1987) who seems to have pioneered it, and Simmons (1996) being noteworthy exceptions. Others refer to Borland when justifying the inclusion of lagged attendance data in their equations (Dobson and Goddard, 1995).

Borland was well aware that there was an identification problem insofar as lagged attendance might also pick up (ticket) price expectations formation and/or lagged adjustment by spectators. He dismissed these two possibilities on the grounds (respectively) that ticket prices are set in advance, and that there did not appear to be significant adjustment costs facing spectators. The former argument is more persuasive (although the ticket price is only a portion of the price of attendance) than the latter, and for two good reasons. First, the relative preponderance of season ticket holders nowadays implies significant short-run adjustment costs for spectators. Second, a subset (of unknown magnitude) of those who buy match tickets on the day are regulars[2] who effectively buy a consumer durable ('the season') on a game by game instalment plan. This suggests there are significant adjustment costs even to some non-

season ticket holders. Neale's (1964) 'league standing effect' recognised that the consumption of professional teams sports, as an activity is not limited to the period of attendance or even to that section of the population who attend.

In this note, beginning with a general model embodying both habit persistence and partial adjustment Borland's original analysis is revisited. Restrictions are imposed on the model to derive implied structural forms for two sub-models in each of which only one of the two hypotheses is maintained. The paper concludes that one cannot pick up habit persistence by simply adding lagged attendance to demand equations.

2. *The general model*

To begin with, it is clear that Borland had not intended to capture either partial adjustment or expectations formation. However, as will be demonstrated, his model is one of partial adjustment or of adaptive expectations, contrary to his intentions. To see why, let us return to the habit persistence model due to Brown (1952), which Borland explicitly adopted. We start with a general model embodying both habit persistence and partial adjustment. The 'standard' model of attendance (equation 1) is intended to be the sub-case that excludes partial adjustment.

In the general model attendance at time t depends linearly (*inter alia*) on the 'stock' of habits (H_t) for the team's home games. Habits are derived from past attendance according to the geometrically weighted function

$$H_t = A_{t-1} + hA_{t-2} + h^2A_{t-3} + \ldots\ldots \quad [2]$$

where h $(0 < h < 1)$ is the rate of habit formation. Equation 2 implies that in the long run stationary state $H^e = A^e/(1 - h)$ where superscript e denotes a long-run equilibrium value. Equation 2 implies (equation 3 below) that the rate of change (ΔH_t) of the habit stock between time (t - 1) and time t is a linear function of current attendance and of the previous level of the stock

$$\Delta H_t = A_{t-1} + (h - 1)H_{t-1} \quad (0 < h < 1) \quad [3]$$

Equation 3 is a net investment function for the attendance habit; current gross investment is given by the previous attendance level, while depreciation is simply the habit decay rate $(1 - h)$ multiplied by the lagged habit stock (H_{t-1}).

As an aside it is worth noting that writers have recognised that attendance observations are truncated by the capacities of stadia. Only Kuypers (1996) seems to have taken this factor into account in his estimation procedure. Other writers have treated capacity constraints as a source of heteroscedasticity. Reporting White's robust standard errors has been the typical correction procedure employed, as Downward and Dawson (2000) note. More generally, for English soccer this problem is probably of minor significance outside of the Premier League and prior to the 1990's.

The desired demand for attendance is given by

$$A^*_t = a_1 + a_2P_t + a_3H_t \quad [4]$$

where A^* is the equilibrium desired attendance level. We assume partial adjustment of the form

$$\Delta A_t = g(A^*_t - A_{t-1}) \quad [5]$$

Parameter g $(0 < g < 1)$ defines the proportion of adjustment completed in any single period. Substitute equation 4 into equation 5 to obtain an expression for A_t which does not include the unobserved A^*_t.

$$A_t = ga_1 + ga_2P_t + (1-g)A_{t-1} + ga_3H_t \quad [6]$$

Equation 6 cannot be estimated as it stands since H is never observed. So, substitute equation 2 into equation 6 to remove H and we arrive at

$$A_t = ga_1 + ga_2P_t + (1-g)A_{t-1} + ga_3(A_{t-1} + hA_{t-2} + h^2A_{t-3} + ...) \quad [7]$$

which has an infinity of RHS terms. Apply the Koyck transformation to equation 7 to get to equation 8

$$A_t = (1 - h)ga_1 + ga_2Pt - hga_2P_{t-1} + [(1-g) + ga_3 + h]A_{t-1} - h(1-g)A_{t-2} \quad [8]$$

This is the general specification, embodying both partial adjustment and habit persistence. It seems that two lags in attendance and one in price are required if both Habit Persistence and Partial Adjustment are to feature in the model.

It is clear that if (unlike Borland) the researcher desires to model partial adjustment and habit persistence, equation 1 is a mis-specified form of equation 8. The latter contains five behavioural parameters (the three a_is, g and h) and five reduced form parameters, thus identification may be obtained. Estimates of the reduced form coefficients on A_{t-1} and A_{t-2} have by themselves no significance for the existence, persistence or extent of habit persistence. Thus it may be unwise to base economic advice on such estimates alone. The coefficient of A_{t-1} compounds the rate of habit formation, the rate of partial adjustment and the instantaneous response of attendance to the habit stock. The coefficient of A_{t-2} compounds the rates of habit formation and of partial adjustment.

3. The habit persistence model

A researcher, starting from the general model (equation 8) who dismisses partial adjustment and who aims to model habit persistence alone will set $g = 1$ (complete adjustment within the period) in equation 8 and obtain equation 9.

$$A_t = (1- h)a_1 + a_2P_t - ha_2P_{t-1} + [a_3 + h]A_{t-1} \quad [9]$$

The double lag in attendance drops out, but unlike equation 1 there is one lag in price, so equation 1 is now seen to be a mis-specification of equation 9. The latter equation is what one ought to obtain, given that habit persistence (or at any rate Brown's version) is retained and partial adjustment dropped from the general model.

4. The partial adjustment model

A researcher, operating within this general framework, who rejects habit persistence will impose $h = a_3 = 0$ in equation 8, and retain partial adjustment, arriving at a reduced form (equation 10) which is identical to equation 1.

$$A_t = ga_1 + ga_2P_t + [(1 - g)]A_{t-1} \quad [10]$$

Here the estimated reduced form coefficient on lagged attendance picks up delayed adjustment (or price expectations formation) only; as asserted earlier, Borland's is a model of partial adjustment rather than of habit persistence. At this point one might wonder how Borland arrived at a reduced form categorically excluding the very type of behaviour he had intended to investigate. The answer is very simple - he omitted the vital Δ in his equivalent (Borland, page 229, equation A3) of our equation 2 - with the effect of making a stock the difference between two flows. His subsequent argument for the inclusion of lagged values of attendance proceeded informally. Otherwise it is virtually certain that he would have discovered the missing Δ.

5. Conclusions

This note argues that one cannot pick up habit persistence by simply adding lagged attendance to demand equations; some explicit modelling is necessary. The standard practice, estimating equations like equation 1 and interpreting the estimated coefficients on the lagged attendance variable as habit persistence parameters might under some circumstances be justifiable - but only under a

structure different from that proposed by Borland. We would recommend starting with a more general specification (not necessarily equation 8) and testing down. This would give some clear guidelines for empirical testing, such as jointly testing for the presence of the current and lagged terms on price in equation 9 for example, and thereby highlight the distinction between habit persistence and partial adjustment.

Endnotes

1. Staffordshire University Business School; our thanks are due to two anonymous referees and to the Editor. Their comments enabled us to remove at least some of the errors and infelicities that inevitably attend writing.

2. Simmons (1995) suggests that the price elasticity of demand is lower for season ticket holders than for those sometimes termed 'theatre goers'. This is quite consistent with the existence of a subset of non season-ticket holders who yet are regular supporters of their local team and who face higher time and travel costs should they decide to switch to another club.

References

Borland J (1987) 'The Demand for Australian Rules Football', *The Economic Record*, 63, 220-230.

Brown T M (1952) 'Habit Persistence and Lags in Consumer Behaviour', Econometrica, 20, 355-371.

Cairns J (1990) 'The Demand for Professional Team Sports', *British Review of Economic Issues*, 12, 1-20.

Dobson S M and Goddard J A (1995) 'The demand for professional league football in England and Wales, 1925-92', Discussion Paper, Department of Economics, University of Hull.

Downward P M and Dawson A J (2000) 'The Economics of Professional Team Sports', Routledge (forthcoming).

Hart R A, Hutton J and Sharot T (1975) 'A Statistical Analysis of Association Football Attendance', *Journal of the Royal Statistical Society*; Series C (Applied Statistics), 24, 1, 17--27.

Kuypers T (1996) 'The beautiful game? An econometric study of why people watch English football', University College London Discussion Papers in Economics, 96-01.

Neale W C (1964) 'The Peculiar Economics of Professional Sport', *The Quarterly Journal of Economics*, 78, 1-14.

Simmons R (1996) 'The Demand for English league football; a club-level analysis', *Applied Economics*, 28, 139-155.