University of Victoria
Department of Economics

Econometrics Working Paper EWP0513

ISSN 1485-6441

ARE SPORTS TEAMS MULTI-PRODUCT FIRMS?

Kenneth G. Stewart
&
J. C. H. Jones

Department of Economics, University of Victoria
Victoria, B.C., Canada V8W 2Y2

June, 2005

Abstract

The appropriate conception of team outputs is investigated by estimating a two-output factor demand system for baseball teams, relative to which single-output models are rejected. There is, however, some empirical support for output separability, suggesting that team outputs may sometimes be adequately treated as a production aggregate.

Keywords: Sports economics, multi-output production

JEL Classifications: D24, L83

Author Contact:
K. G. Stewart, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada V8W 2Y2; e-mail: kstewart@uvic.ca; Phone: (250) 721-8534; FAX: (250) 721-6214

1 Introduction

The economics of sport serves as a perhaps stereotypical example of the gradual emergence into respectability of a new field within a discipline. Beginning in scattered journal articles and spoken of in hushed and mocking tones as the bastard child of other areas like labor economics and industrial organization, the field gradually became recognizable in occasional books and edited volumes. Its maturity is now evidenced by the establishment of the *Journal of Sports Economics,* survey articles in the *Journal of Economic Literature* (Fort and Quirk, 1995; Szymanski, 2003), a symposium in the *Economic Journal* (2001), and the frequent inclusion of sessions related to sports economics in conference programs. Courses on sports economics have usurped a place in undergraduate calendars as a popular elective, a market now served by several textbooks such as Downward and Dawson (2000), Fort (2003), Leeds and von Allmen (2002), and Sandy, Sloane, and Rosentraub (2004). The field is now well-placed to claim its noble lineage to such pioneering contributions as Rottenberg (1956), Jones (1969), Scully (1974), and Rosen (1981).

This emergence is in large measure attributable to the quality and quantity of data available for this industry. Whereas data on conventional markets and industries is often difficult to obtain—particularly at the level of individual agents such as firms and workers—a vast array of data are readily available on teams and players. Thus sports economics has provided a natural testing ground for ideas that in principle might be of interest in other industries or markets, but where data limitations constrain empirical inquiry. Sports economics has, therefore, emerged as a heavily empirical discipline.

The essential premise of the field is that sports teams may be modeled as profit maximizing firms serving a demand for their product in their output market. This in turn generates a derived demand for their factors of production—most importantly players. The nature of the industry provides almost endless scope for collusive and strategic behavior in these product and factor markets, and most research in the field concerns itself in one way or another with aspects of this product or factor market activity.

However the analogy of players and teams with conventional workers and firms only extends so far, and there is much that is specific to sports as an industry that must be treated in empirical analysis. For example, the dominant component of team variable costs is player salaries, and one view is that the factors of production are properly viewed as the skill characteristics embodied in players. Although these skill characteristics are directly measurable, their prices are not: instead they are hedonic prices that must be estimated. This interpretation of the factor market for players as an "implicit market" for their skills is exposited in detail in Stewart and Jones (1998) and applied in Ferguson, Jones, and Stewart (2000).

Similarly, the nature of output in the industry is less clear-cut than in conventional industries. If teams serve a demand for their output in their product market, what is that output? Consumer demand for the product has traditionally been viewed as a demand for attendance at sporting events: see Cairns (1990), Downward and Dawson (2000, Chaps. 5,6), and Fort (2003, Chap. 2). However, clearly teams do not literally produce attendance; instead attendance is being used as a proxy for the direct outputs of teams that are less readily measurable. But what are these direct outputs?

The conjecture of this paper is that teams are best thought of as multi-product firms with two outputs, *performance* and *entertainment.* In our view it is these outputs that in turn determine attendance and gate receipts (and, peripherally, concession and parking income), media and merchandizing revenue, and ultimately profits. If our notion is correct, and if performance and entertainment can be separated and measured satisfactorily, estimations of behavioral outcomes will, presumably, be more accurate than those based on a single omnibus attendance variable. This, in turn, calls into question the adequacy in sports economics of demand analyses based on the assumption that teams produce a single output.

We explore our conjecture by applying the methodology that is used to study multi-product firms in other industries: a multi-output cost function. In doing this, we adopt the view of Stewart and Jones (1998) and Ferguson, Jones, and Stewart (2000) of the factor market as an implicit market for player skills in which factor prices are hedonic. As in those studies, we focus our analysis on Major League Baseball, the only sport in which the detailed salary and player skill data are available to estimate the necessary hedonic prices. In principle, however, the procedure should be applicable to any team sport.

We find that single output measures are rejected, suggesting that sports teams should be viewed as multi-product firms. There is, however, evidence that these multiple outputs may sometimes be adequately treated as a production aggregate.

2 Modeling Framework

Adopting the premise of sports economics generally, we draw on the standard microeconomic theory of the firm to describe sports teams. It is assumed that teams maximize profit by employing inputs $\boldsymbol{x} = [x_1, \ldots, x_n]$ to produce outputs $\boldsymbol{y} = [y_1, \ldots, y_m]$ using a technology

$$t(\boldsymbol{y}, \boldsymbol{x}) = 0. \tag{1}$$

The dual of this profit maximization problem is the cost function

$$C(\boldsymbol{y}, \boldsymbol{p}), \tag{2}$$

where $\boldsymbol{p} = [p_1, \ldots, p_n]$ is the vector of factor prices. The n factor demand equations may be obtained using Shephard's lemma:

$$x_i = \frac{\partial C(\boldsymbol{y}, \boldsymbol{p})}{\partial p_i} = x_i(\boldsymbol{y}, \boldsymbol{p}) \qquad (i = 1, \ldots, n). \tag{3}$$

Potentially this cost function may be parameterized in a number of ways. By far the most widely used cost function in empirical work, in both single- and multiple-output contexts, has been the translog. However it has several important limitations that have increasingly led researchers to turn to alternative flexible functional forms, notably the symmetric generalized McFadden (SGM) cost function of Diewert and Wales (1987). The most common motivation for using the SGM model is that it permits the imposition of concavity, something that the translog does not.[1] However for our purposes there are even more important reasons for using the SGM model. First, the translog model applies logarithmic transformations to all variables and so is not applicable in contexts in which some observations have zero prices or output levels. Although some modifications of the translog model address this through the use of Box-Cox terms, the cleanest approach is to abandon the model entirely in favor of one that is formulated in terms of the untransformed variables, as is the SGM model. Second, in applying Shephard's lemma to derive factor demands from the cost function, key parameters of the translog model relating to the dependence of cost on output are lost. Consequently some important hypotheses and elasticities—such as those relating to returns to scale—require the estimation of the cost function joint with the system of factor demands, something that requires data on the level of costs. By contrast, the SGM model involves no loss of parameters in going from the cost function to the factor demands, and so the estimation of the factor demand system alone enables the recovery of all the parameters of the cost function. Both these considerations turn out to be of critical importance in studying baseball teams as multi-product firms.

2.1 Separability and Output Aggregation

A special case that is of particular interest in the present context is that in certain circumstances the production technology (1) may be separable of the form $\psi(\boldsymbol{y}) = f(\boldsymbol{x})$. The function $\psi(\cdot)$ may be interpreted as an aggregator so that $\psi(\boldsymbol{y})$ is a single aggregate output.

In this case it is well known from Hall (1973) that the cost function (2) can be written as $C(\boldsymbol{p}, \psi(\boldsymbol{y}))$, and the factor demands take the form

$$x_i = \frac{\partial C(\boldsymbol{p}, \psi(\boldsymbol{y}))}{\partial p_i} = x_i(\boldsymbol{p}, \psi(\boldsymbol{y})) \qquad (i = 1, \ldots, n). \tag{4}$$

The traditional situation of a single output y is the special case in which the production technology is $y = f(\boldsymbol{x})$, the cost function has the form $C(\boldsymbol{p}, y)$, and the factor demands are[2]

$$x_i = \frac{\partial C(\boldsymbol{p}, y)}{\partial p_i} = x_i(\boldsymbol{p}, y) \qquad (i = 1, \ldots, n). \tag{5}$$

In this framework there are two ways of interpreting the question posed by the title of this paper. One is as the hypothesis that modeling teams as multi-output producers is not a significant improvement over viewing them as single-output producers, so that the multi-output factor demand system (3) is no better a description of team behavior than the system (5) based on a single output. In this paper we use model selection criteria and nested and nonnested testing to examine this hypothesis.

The second way of interpreting the question of this paper is as the hypothesis that teams' production technologies are separable, so that even if teams produce more than one output these outputs are treated in their production decisions as a single aggregate $\psi(\boldsymbol{y})$. This is the restriction of the general demand system (3) to the separable system (4). In the context of the SGM model it turns out that as long as we consider aggregator functions of a linear form,

$$\psi(\boldsymbol{y}) = \sum_k \beta_k y_k, \tag{6}$$

separability imposes parametric restrictions on the general demand system (3) that can be tested. Because the aggregator $\psi(\boldsymbol{y})$ is an argument of $x_i(\boldsymbol{p}, \psi(\boldsymbol{y}))$ it is identified only up to a multiplicative constant and so requires an additional normalizing restriction in estimation, something we comment further on below. Essentially, then, $\psi(\boldsymbol{y})$ has the interpretation as a production index with arbitrary base determined by the normalization.

3 Data

Our SGM factor demand system for baseball player skills is based on $m = 2$ outputs $\boldsymbol{y}_t = [y_{1t}, y_{2t}]$ and $n = 4$ inputs $\boldsymbol{x}_t = [x_{1t}, x_{2t}, x_{3t}, x_{4t}]$. These are observed over the 1986–91 seasons for the 26 major league baseball teams that existed during those years, for a total of $T = 156$ observations indexed by t.

As described in the introduction, our conjecture is that teams' outputs are best thought of as performance y_1 and entertainment y_2. The obvious measure of a team's performance is the proportion of the games it wins over the season ("wins" for short). An appropriate measure

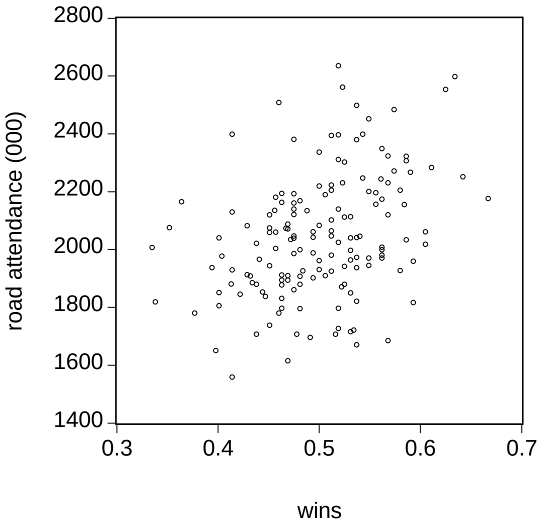

Figure 1: Wins and road attendance, 1986–91: correlation = 0.364

of a team's entertainment value is, on the other hand, more elusive: home attendance as such a measure, for example, is inextricably confounded with team performance. In an attempt to abstract from performance in measuring the entertainment value of a team, we propose the total season attendance at its games away from home ("road attendance" for short). That is, all other things equal, entertaining teams should draw large crowds when on the road. Their performance, on the other hand, works in two offsetting directions that may roughly balance in affecting road attendance. On one hand, it might be thought that high-performance teams should draw large crowds at their away games. But on the other hand, home town fans tend to stay away from games that the home team is likely to lose. Because performance works in offsetting directions in its affect on road attendance, road attendance figures may be regarded as primarily reflecting a team's entertainment value.

In summary, our two outputs are

y_{1t} = team performance, measured by wins;
y_{2t} = entertainment value of the team, measured by road attendance.

It is to be emphasized that neither of these corresponds to the measure of output most often used in the sports literature, which is, as we noted in the introduction,

y_{3t} = attendance at home games ("home attendance" for short).

To be sure, y_{1t} and y_{2t} are positively related to some degree: winning teams tend to draw large crowds at their away games. But as the scatter plot of our complete sample of (y_{1t}, y_{2t}) values in Figure 1 indicates, the relationship is not particularly strong: the sample correlation is 0.364. This suggests that the idea that these outputs play distinct roles in teams' production decisions is not to be dismissed out of hand.

Teams produce these outputs by employing player skills. Following Stewart and Jones (1998), our four skill inputs are defined as follows:

x_{1t} = *experience,* measured by the total years of experience in the major leagues of all players on the team;
x_{2t} = *hitting skills,* measured by the published team slugging average for the current season, scaled by the number of hitters;
x_{3t} = *pitching skills,* measured by the ratio of total strikeouts to walks for all pitchers over the current season, scaled by the number of pitchers;
x_{4t} = *stars,* measured by the number of players who were elected to the *Sporting News* All-Star or Gold Glove teams, or who played in the all-star game, in the current season.

Hitting and pitching skills x_{2t} and x_{3t} are each scaled by the number of players involved in order to yield factor inputs interpretable as total rather than average quantities, consistent with the other two inputs x_{1t} and x_{4t}. The justification for the choice of these skill characteristics and the details of the estimation of the associated hedonic price vectors $\boldsymbol{p}_t = [p_{1t}, p_{2t}, p_{3t}, p_{4t}]$ are discussed in Stewart and Jones.

Is there any reason to accept our premise that estimated hedonic prices for player skills can be treated like observed prices for the purpose of applied demand analysis? As a simple descriptive check on the data, Stewart and Jones computed the cost identity

$$\text{costs} = \sum_{i=1}^{n} p_{it}x_{it} = p_{1t}x_{1t} + p_{2t}x_{2t} + p_{3t}x_{3t} + p_{4t}x_{4t}. \qquad (7)$$

The resulting cost series is obtained from estimated hedonic prices and so we will call it *hedonic costs* to distinguish it from observed cost data. Because it is, in this sense, a synthetically constructed series, it would be inappropriate to use it in place of observed cost data in the estimation of a cost function. Even so, it may be used to obtain the implied cost shares

$$v_{it} = \frac{p_{it}x_{it}}{\text{costs}} \qquad (i = 1, \ldots, 4).$$

The average values for these shares obtained by Stewart and Jones are reproduced in Table 1, and imply an entirely plausible breakdown of team variable costs. The estimated hedonic prices suggest that almost half of the typical team's payroll is a payment to hitting skills, a third is for experience, and 12.4% is for pitching skills. Only about 6% is for stars; that is, controlling for the other skills that stars bring to a team and taking into account the small number of stars on a team, the employment of stars accounts for a fairly small fraction of variable costs.

In the present context where the SGM cost function can be reconstructed from the demand estimates, another simple check on the hedonic approach is available. The estimated demand system can be used to obtain a fitted cost series that can be compared with hedonic costs. Consistency between the series serves as an internal check on the legitimacy of our approach.

Table 1: Mean Hedonic Cost Shares $\bar{v}_i$

experience:	$\bar{v}_1 = 0.338$
hitting:	$\bar{v}_2 = 0.476$
pitching:	$\bar{v}_3 = 0.124$
stars:	$\bar{v}_4 = 0.061$

4 The Multi-Output SGM Demand System

Let us turn to the parameterization of the multi-output cost function (2) that is provided by the SGM model.

The single-output SGM cost function of Diewert and Wales (1987) has been extended to the multi-output context by Kumbhakar (1994):

$$C(\boldsymbol{y}, \boldsymbol{p}) = g(\boldsymbol{p}) \sum_k \beta_k y_k + \sum_i b_i p_i + \sum_i b_{ii} p_i \left(\sum_k \beta_k y_k \right) + \left(\sum_i \lambda_i p_i \right) \sum_j \sum_k d_{jk} y_j y_k$$

where

$$g(\boldsymbol{p}) = \frac{\boldsymbol{p}' \boldsymbol{S} \boldsymbol{p}}{2 \boldsymbol{\theta}' \boldsymbol{p}}.$$

Because the matrix $\boldsymbol{S}$ appears in a quadratic form, without loss of generality it is specified to be symmetric (so $s_{ij} = s_{ji}$) and satisfy $\sum_i s_{ij} = 0$ for all j. For similar reasons the d_{jk} are specified as symmetric and a further identifying restriction is imposed on them that we comment on momentarily. As in Diewert and Wales's (1987) original single-output version of the model, for estimation purposes each component of $\boldsymbol{\theta}$ is set equal to the sample mean of the corresponding factor demand: $\theta_i = \bar{x}_i$ for all i. By the cost identity (7), $\boldsymbol{\theta}' \boldsymbol{p}$ has the interpretation as the value of costs at the mean level of factor inputs.

Applying Shephard's lemma, the factor demand system (3) is

$$x_i = g_i(\boldsymbol{p}) \sum_k \beta_k y_k + b_i + b_{ii} \sum_k \beta_k y_k + \lambda_i \sum_j \sum_k d_{jk} y_j y_k \qquad (i = 1, \ldots, n) \qquad (8)$$

where, letting $\boldsymbol{S}^{(i)}$ denote the ith row of the matrix $\boldsymbol{S}$,

$$g_i(\boldsymbol{p}) \equiv \frac{\partial g(\boldsymbol{p})}{\partial p_i} = \frac{\boldsymbol{S}^{(i)} \boldsymbol{p}}{\boldsymbol{\theta}' \boldsymbol{p}} - \frac{\theta_i \boldsymbol{p}' \boldsymbol{S} \boldsymbol{p}}{2 (\boldsymbol{\theta}' \boldsymbol{p})^2}. \qquad (9)$$

Special Case of Two Outputs

In our case of $m = 2$ outputs this demand system is, using the symmetry restriction $d_{12} = d_{21}$,

$$x_i = g_i(\boldsymbol{p})(\beta_1 y_1 + \beta_2 y_2) + b_i + b_{ii}(\beta_1 y_1 + \beta_2 y_2) + \lambda_i (d_{11} y_1^2 + 2 d_{12} y_1 y_2 + d_{22} y_2^2) \qquad (i = 1, \ldots, n).$$

Inspection reveals that two further normalizing restrictions are needed for the parameters of this demand system to be identified. The first is with respect to the coefficients β_1 and β_2, which are not separately identified jointly with the s_{ij} or b_{ii} in the first and third terms. Although any linear restriction on the β_k would serve to identify them (for example, that they be required to sum to one), estimation is aided by simply setting $\beta_1 = 1$.

The second normalizing restriction is with respect to the d_{jk}, which are not separately identified jointly with λ_i in the fourth term. Again, although in principle any linear restriction would identify the d_{jk}, estimation is aided by setting $d_{11} = 1$.

With these normalizing restrictions, and simplifying notation by relabelling β_2 as β, the SGM demand system with two outputs becomes

$$M_0: \quad x_i = g_i(\boldsymbol{p})(y_1+\beta y_2)+b_i+b_{ii}(y_1+\beta y_2)+\lambda_i(y_1^2+2d_{12}y_1y_2+d_{22}y_2^2) \qquad (i=1,\ldots,n).$$

5 Single- Versus Multiple-Outputs

Is the multi-output SGM model a better description of baseball team production decisions than its single-output counterparts?

Consider the original Diewert and Wales (1987) single-output version of the SGM model, assigning either wins y_1 or road attendance y_2 the role of the single output. The factor demand system (5) is, for each output respectively,[3]

$$M_1: \quad x_i = g_i(\boldsymbol{p})y_1 + b_i + b_{ii}y_1 + \lambda_i y_1^2 \qquad (i=1,\ldots,n) \tag{10a}$$

$$M_2: \quad x_i = g_i(\boldsymbol{p})y_2 + b_i + b_{ii}y_2 + \lambda_i y_2^2 \qquad (i=1,\ldots,n). \tag{10b}$$

Comparing with the multi-output model M_0, these are parametric special cases of that more general model: setting $\beta = d_{12} = d_{22} = 0$ in M_0 yields M_1 and, by the symmetry of the labeling of the outputs y_1 and y_2, M_2 is similarly a special case of M_0. Thus either of these single output models can be tested as a parametric restriction of the multi-output model (although of course M_1 and M_2 are nonnested in relation to one another); in the analysis below we use likelihood ratio tests to do this.

However y_1 and y_2 are not the only candidates for a single output measure. The obvious alternative is the conventional one, home attendance y_3, and so in addition to the above models we define

$$M_3: \quad x_i = g_i(\boldsymbol{p})y_3 + b_i + b_{ii}y_3 + \lambda_i y_3^2 \qquad (i=1,\ldots,n).$$

This is nonnested in relation to all three of M_0, M_1, and M_2.

There are two approaches to comparing nonnested models: a model selection approach based on information criteria, and nonnested testing. We consider each in the sections that follow.

Table 2: Model Selection

	M_0	M_1	M_2	M_3
$\mathscr{L}$	-1532.33	-1547.63	-1564.82	-1548.52
number of coefficients K	21	18	18	18
BIC$= -\mathscr{L} + \frac{K}{2}\ln(nT)$	1599.91	1605.56	1622.74	1606.45

5.1 Model Selection Criteria

Let us begin with the model selection approach to comparing our four models, the two-output model M_0 and the single-output models M_1, M_2, and M_3. Their loglikelihood function values $\mathscr{L}$ are reported in Table 2. The loglikelihoods for the single-output models are directly comparable because they have the same number of parameters, and on this basis the preference ranking is M_1, M_3, M_2. This is intuitively plausible: if it were necessary to choose a single output measure a priori, one would almost certainly choose wins or home attendance rather than road attendance. Note as well that the loglikelihoods of M_1 and M_3 are almost the same, so the preference for wins over home attendance as the output measure is a marginal one.

Turning to the two-output model M_0, it has a higher loglikelihood than any of the single-output models, but it is also more generously parameterized. To penalize these additional parameters we use Schwartz's Bayesian information criterion (BIC) which, compared to other popular model selection criteria such as the Akaike information criterion, imposes a relatively heavy penalty for additional parameters, and in this sense is biased against the multi-output model. The BIC values are reported in the final line of Table 2. The multi-output model has a substantially lower BIC than any of the single-output models, indicating that it is preferred. (A smaller BIC means a preferred model.) Note that, because the single-output models have the same number of parameters, the ranking of them given by the BIC is the same as that yielded by $\mathscr{L}$.

5.2 Nested Testing

Because M_1 and M_2 are each restricted versions of M_0, a likelihood ratio (LR) test can be used to test them. (We use LR tests because our demand systems are nonlinear, a situation in which Wald statisics—the natural alternative test—have poor invariance properties.) Using the loglikelihood function values from Table 2, the LR statistic for M_1 against M_0 is $2[-1532.33 - (-1547.63)] = 30.60$, while that for M_2 against M_0 is $2[-1532.33-(-1564.82)] = 64.98$; both reject the three restrictions in question very strongly.

(For example, $\chi^2_{0.01}(3) = 11.34$.)

Since the data clearly reject both of M_1 and M_2 against the multi-output model, these results are fully consistent with the findings from the model selection criteria.

5.3 Model Selection Tests

The model selection test methodology of Vuong (1989) provides another way of comparing our four models. The likelihood ratio tests just discussed use the traditional Neyman-Pearson framework of treating null and alternative hypotheses asymmetrically, asking whether the data provide compelling evidence against the restrictions of the null. We have found that, in the case of M_1 and M_2, it does.

By contrast Vuong's test, although still likelihood-based, treats alternative models symmetrically, asking whether the data provide evidence that one is closer to a hypothetical true model. Our primary interest is in using this to compare models that are nonnested, namely M_3 versus M_0 and M_1, M_2, and M_3 against one another. Nevertheless it is also of interest to use it to compare the nested models—M_0 with either of M_1 or M_2—comparing the findings with those yielded by the likelihood ratio tests.

The interpretation of Vuong's test differs somewhat from most nonnested tests and so requires an understanding of the mechanics of his procedure. It is therefore useful to begin with a sketch of those mechanics as they apply to our context.

Outline of Vuong's Model Selection Test

Consider any two of the demand systems M_0, M_1, M_2, and M_3. Following Vuong's notation, denote the estimable versions of these two models as the following n-equation multivariate regressions:

$$Y_t = F(Z_t; \theta) + u_t$$
$$Y_t = G(Z_t; \gamma) + v_t.$$

Here $Y_t' = [x_{1t} \cdots x_{nt}]$ is the vector of n factors at observation t; Z_t is the vector of right hand side variables. (In the multi-output model M_0 this would comprise the full set of prices and outputs, $Z_t = [\boldsymbol{p}_t; \boldsymbol{y}_t]$, while in the single-output models this would specialize accordingly.) The disturbances u_t and v_t are $n \times 1$ multivariate normal disturbance vectors, statistically independent across observations indexed by t.

Consider the first of these demand systems, $F(\cdot)$, and let the covariance matrix of u_t be denoted by $\Omega : n \times n$. Let the density of Y_t conditional on Z_t be denoted $f(Y_t|Z_t, \theta)$. The logarithm of this conditional density is

$$\ln f(Y_t|Z_t, \theta) = -\frac{1}{2}T\ln(2\pi) - \frac{1}{2}T\ln|\Omega| - \frac{1}{2}u_t'\Omega^{-1}u_t.$$

The loglikelihood function is the sum of these log-densities:

$$\mathscr{L}_F(\theta) = \sum_t \ln f(Y_t|Z_t, \theta).$$

The loglikelihood function for the second demand system $G(\cdot)$ is defined analogously:

$$\mathscr{L}_G(\gamma) = \sum_t \ln g(Y_t|Z_t, \gamma).$$

Let $\hat{\theta}$ and $\hat{\gamma}$ denote the respective maximum likelihood estimators, and $\mathscr{L}_F(\hat{\theta})$ and $\mathscr{L}_G(\hat{\gamma})$ the associated maximized loglikelihood function values. In the case of model F, for example, in which the estimated covariance matrix is denoted $\hat{\Omega}$ and the $n \times 1$ vector of residuals at observation t is $\hat{u}_t = Y_t - F(Z_t; \hat{\theta})$, the loglikelihood function value is computed as

$$\mathscr{L}_F(\hat{\theta}) = \sum_t \ln f(Y_t|Z_t, \hat{\theta}) = \sum_t \left[-\frac{1}{2} T \ln(2\pi) - \frac{1}{2} T \ln|\hat{\Omega}| - \frac{1}{2} \hat{u}_t' \hat{\Omega}^{-1} \hat{u}_t \right],$$

and similarly for $\mathscr{L}_G(\hat{\gamma})$.

The observation-by-observation components $\ln f(Y_t|Z_t, \hat{\theta})$ and $\ln g(Y_t|Z_t, \hat{\gamma})$ of the loglikelihood functions are used in the calculation of Vuong's test statistic. The null hypothesis that the competing models F and G are equally close to the (unobservable) true model is stated as

$$H_0 : E^0 \left[\ln \frac{f(Y_t|Z_t, \theta)}{g(Y_t|Z_t, \gamma)} \right] = 0, \tag{11}$$

where E^0 denotes the mathematical expectation with respect to the true conditional density of Y_t. Vuong's methodology allows this null to be tested against the alternative that F is closer to the true model than is G,

$$H_F : E^0 \left[\ln \frac{f(Y_t|Z_t, \theta)}{g(Y_t|Z_t, \gamma)} \right] > 0,$$

and against the alternative that G is closer to the true model than is F,

$$H_G : E^0 \left[\ln \frac{f(Y_t|Z_t, \theta)}{g(Y_t|Z_t, \gamma)} \right] < 0.$$

Thus the alternative hypothesis is two-sided.

The test proceeds by using the sample values of the argument of these mathematical expectations,

$$\ell_t \equiv \ln f(Y_t|Z_t, \hat{\theta}) - \ln g(Y_t|Z_t, \hat{\gamma}) \qquad (t = 1, \ldots, T).$$

The expected value of these deviations is consistently estimated by their sample mean

$$\bar{\ell} = \frac{1}{T} \sum_t \ell_t = \frac{1}{T} [\mathscr{L}_F(\hat{\theta}) - \mathscr{L}_G(\hat{\gamma})],$$

while their variance is consistently estimated by

$$\hat{\omega}^2 = \frac{1}{T} \sum_t (\ell_t - \bar{\ell})^2.$$

In analogy with an elementary Student's t test of a population mean, Vuong shows that the statistic

$$\frac{\bar{\ell}}{\hat{\omega}/\sqrt{T}} = \frac{\mathscr{L}_F(\hat{\theta}) - \mathscr{L}_G(\hat{\gamma})}{\sqrt{T}\hat{\omega}}$$

is asymptotically standard normal under the null hypothesis H_0 given by (11).

Notice that, as we have remarked, Vuong's procedure treats the competing models symmetrically in the formulation of the hypotheses. This is in contrast to most nonnested tests, where the models are given the asymmetric roles of null and alternative hypotheses, creating the possibility of inconclusive test outcomes in which both models are rejected or neither is rejected. An appealing feature of Vuong's test is that it does not suffer from such ambiguities of interpretation.

Model Comparisons

Considering the models M_0, M_1, M_2, and M_3, Vuong's test may be applied to any pair taken from the four. The various possibilities are considered in Table 3. (We conduct these tests without imposing concavity or separability in order that they not be contingent on potentially false maintained hypotheses.) It turns out that Vuong's test yields much the same conclusions about these models that the information criteria and LR tests did, although with some new insights. We begin by comparing the two-output model M_0 with the single-output models, and then turn to comparing the single-output models M_1, M_2, and M_3 among themselves.

Comparing the Multiple- and Single-Output Models Consider first the tests of the multi-output model M_0 against each of the single-output models. When the single-output model is M_1 or M_2 the null hypothesis that the multi- and single-output models are equally close to the true model is strongly rejected in favor of the alternative that the multi-output model is closer. The rejection is especially strong for the single-output model M_2 in which the output measure is road attendance y_2. Not surprisingly, these conclusions are entirely consistent with those yielded by the LR tests in Section 5.2: if the data strongly reject restrictions in a Neyman-Pearson framework, they are bound to favor a model in which the restrictions are not imposed when models are treated symmetrically.

Turning to the more interesting case of the nonnested models M_0 and M_3, here the rejection of the single-output model is less strong: the p-value is 0.0500. Somewhat in contrast to the conclusion yielded by information criteria, therefore, M_3 is the least-strongly rejected of the single-output models—a finding perhaps heartening to researchers who have followed the traditional practice of using home attendance y_3 as a single output measure. In any case, however, it seems fair to say that the evidence strongly favors the multi-output

Table 3: Vuong Selection Tests of Multiple- and Single-Output Models

Model Assignment		Test		
F	G	Statistic Value	p-value (one-tailed)	Conclusion implied by a rejection (in the direction given by the sign of the test statistic):
M_0	M_1	2.682	0.0037	Two-output model better than single-output model based on y_1 = wins
M_0	M_2	3.546	0.0002	Two-output model better than single-output model based on y_2 = road attendance
M_0	M_3	1.645	0.0500	Two-output model better than single-output model based on y_3 = home attendance
M_1	M_2	1.579	0.0552	Given choice of a single output, model based on y_1 = wins is better than model based on y_2 = road attendance
M_1	M_3	0.092	0.4633	Given choice of a single output, model based on y_1 = wins is better than model based on y_3 = home attendance
M_2	M_3	-1.727	0.0421	Given choice of a single output, model based on y_2 = road attendance is inferior to model based on y_3 = home attendance

model over any of the single-output ones.

Comparing the Single-Output Models Suppose that we nevertheless limit ourselves a priori to the single-output models. We have seen that information criteria yield the ranking M_1, M_3, M_2, but with M_1 and M_3 being very close. The last three tests in Table 3 are consistent with this. The first and third offer fairly strong evidence rejecting M_2 relative to either M_1 or M_3, with p-values of 0.0552 and 0.0421 respectively. Again, it is intuitively plausible that road attendance alone should be rejected as a satisfactory output measure.

Comparing M_1 and M_3, the one-sided p-value is 0.4633; the null that the two models are equally close to the true model is therefore not rejected. Thus nonnested testing does not clearly favor one or the other of wins or home attendance as a single output measure.

5.4 Conclusions

In summary, we have found that model selection criteria, LR tests of the nested models, and Vuong's model selection tests yield a largely consistent picture of the validity of the competing models. The two-output model is clearly preferred, with the data providing fairly strong evidence against any of the single-output models. But if one limits oneself a priori to a single output measure, on purely statistical grounds there is no strong basis for choosing between wins y_1 and home attendance y_3.

6 Separability

Although these results suggest that performance and entertainment value are both team outputs, it is conceivable that teams may treat them as a single aggregate output. This is the hypothesis of separability.

As result (4) indicates, under separability the SGM factor demands take the form of the single-output demands (10) in which the single output is the aggregate $\psi(\boldsymbol{y})$:

$$x_i = g_i(\boldsymbol{p})\psi(\boldsymbol{y}) + b_i + b_{ii}\psi(\boldsymbol{y}) + \lambda_i\psi^2(\boldsymbol{y}) \qquad (i = 1, \ldots, n). \tag{12}$$

Here $g_i(\boldsymbol{p})$ is the same as in the multi-output case, and so is given by equation (9); since it does not involve output, it is not affected by the restrictions associated with separability.

6.1 Testing for Separability

Suppose that separability holds and $\psi(\boldsymbol{y})$ is an aggregator of the linear form (6). Substituting this into the demand system (12) yields the special case of the multi-output demands (8) in which the following restrictions are imposed:

$$d_{jk} = \beta_j\beta_k \qquad (j, k = 1, \ldots, m).$$

Thus the hypothesis that outputs are linearly aggregable into a single output measure is testable as a restricted version of the multi-output system.

In our case of $m = 2$ outputs, the separable factor demands (12) with a linear aggregator are

$$M_4: \quad x_i = g_i(\boldsymbol{p})(y_1 + \beta y_2) + b_i + b_{ii}(y_1 + \beta y_2) + \lambda_i(y_1^2 + 2\beta y_1 y_2 + \beta^2 y_2^2) \qquad (i = 1, \ldots, n).$$

Notice that, just as in the general model M_0, identification requires that we impose the normalizing restriction $\beta_1 = 1$, which has the effect of eliminating the coefficient on y_1 in all terms. As before, notation is simplified by relabeling β_2 as β.

Comparing M_4 with M_0, it is evident that the hypothesis of separability comprises the two restrictions $d_{12} = \beta$ and $d_{22} = \beta^2$. Likelihood ratio tests of these joint restrictions are presented in Table 4. The test outcomes are similar regardless of whether concavity is imposed on the model. Separability is not rejected at conventional significance levels, although it comes close to being rejected at 10%. This suggests that there may sometimes be a case for treating multiple outputs as a single aggregate.

6.2 Isocost Curves

How does the imposition of separability affect the tradeoff between outputs that teams face in making their production decisions?

Table 4: Tests of Separability

	loglikelihood function values		LR statistic	
	unrestricted $\mathscr{L}_U$	restricted $\mathscr{L}_R$	$2(\mathscr{L}_U - \mathscr{L}_R)$	p-value
without concavity imposed	−1532.330	−1534.471	4.282	0.118
with concavity imposed	−1533.654	−1535.767	4.226	0.121

Reconstructing Costs

As our introductory remarks noted, the SGM model has the remarkable feature that there is no loss of parameters in going from the cost function $C(\boldsymbol{y}, \boldsymbol{p})$ to the factor demands. Consequently estimation of the factor demands permits the reconstruction of the companion cost function, even in the absence of cost data. The reconstructed cost function may be used to obtain the fitted values $\hat{C}_t = C(\boldsymbol{y}_t, \boldsymbol{p}_t)$.[4]

How do these reconstructed costs and their implied factor shares compare with the hedonic costs we obtained earlier in connection with Table 1? Table 5 offers this comparison. It shows that the reconstructed cost function—with or without concavity and/or separability imposed—essentially reproduces the cost shares of Table 1. As well, fitted costs are highly correlated with hedonic costs; in this sense the "goodness-of-fit" of the reconstructed cost function is strong, and this goodness-of-fit is robust to the imposition of concavity and/or separability. (We place "goodness-of-fit" in quotations because it is a goodness-of-fit to hedonic rather than observed costs.) These are important numerical checks on the internal consistency of our analysis.

Table 5: Cost Shares and Correlations of Hedonic and Fitted Costs

		fitted costs			
	hedonic	no separability		separability	
	costs	no concavity	concavity	no concavity	concavity
$\bar{v}_1$ (experience)	0.338	0.343	0.345	0.343	0.345
$\bar{v}_2$ (hitting)	0.476	0.471	0.472	0.471	0.471
$\bar{v}_3$ (pitching)	0.124	0.124	0.123	0.124	0.123
$\bar{v}_4$ (stars)	0.061	0.062	0.060	0.062	0.060
correlation of fitted with hedonic costs		0.983	0.983	0.982	0.982

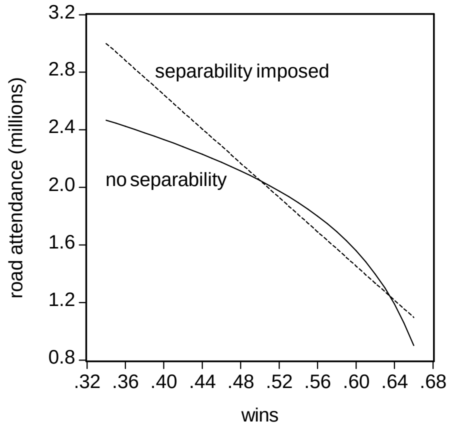

Figure 2: Isocost Curves

Isocost Curves

In addition to evaluating the reconstructed cost function over all points of the observed sample, it may be evaluated at the mean of its arguments: $\hat{C} = C(\bar{\boldsymbol{y}}, \bar{\boldsymbol{p}})$. It is then possible to consider the isocost tradeoff between outputs $\boldsymbol{y} = [y_1; y_2]$ defined by $\hat{C} = C(\boldsymbol{y}, \bar{\boldsymbol{p}})$.

Figure 2 graphs this tradeoff, both with and without separability imposed. (The graph happens to be constructed from the cost function without concavity imposed, but it is unaltered by the imposition of concavity.) The unrestricted model implies a conventional curved transformation frontier, reflecting a marginal rate of transformation that varies with the relative levels of the two outputs. In contrast, the separable model generates a linear transformation frontier, as it must given the a priori specification of a linear aggregator. Thus, although separability is not rejected statistically at conventional significance levels, the implied transformation surfaces do differ qualitatively.

7 Elasticities

Our inferences about multiple outputs and separability are conditional on the legitimacy of our model. What evidence supports the SGM cost function as a credible description of team behavior? As always with flexible functional forms, coefficient estimates have no direct economic interpretation. Instead the economic interpretation of the model comes from the implied elasticities. The following elasticity formulas may be usefully compared with those for the single-output case given by Kumbhakar (1990) and Rask (1995). In all cases we evaluate them at the point of variable means.

Table 6: Allen Elasticities Yielded by Multi-Output SGM Model

	separability not imposed				separability imposed			
	experience	hitting	pitching	stars	experience	hitting	pitching	stars
concavity not imposed	−0.1016	0.0378	0.0427	0.0969	−0.1016	0.0378	0.0417	0.0990
		−0.0046	−0.0375	−0.0924		−0.0041	−0.0379	−0.0974
			0.0602	0.0094			0.0505	0.0469
				0.4533				0.4038
concavity imposed	−0.1077	0.0389	0.0430	0.1224	−0.1080	0.0390	0.0433	0.1216
		−0.0141	−0.0155	−0.0442		−0.0141	−0.0156	−0.0439
			−0.0172	−0.0488			−0.0173	−0.0488
				−0.1390				−0.1370

7.1 Demand Elasticities

The price elasticities of demand are defined as

$$\epsilon_{ij} \equiv \frac{p_j}{x_i}\frac{\partial x_i}{\partial p_j} \tag{13}$$

The relevant derivative is

$$\frac{\partial x_i}{\partial p_j} = \frac{\partial g_i(\boldsymbol{p})}{\partial p_j}\sum_k \beta_k y_k = \left[\frac{s_{ij}}{\boldsymbol{\theta}'\boldsymbol{p}} - \frac{(\boldsymbol{S}^{(i)}\theta_j + \boldsymbol{S}^{(j)}\theta_i)\boldsymbol{p}}{(\boldsymbol{\theta}'\boldsymbol{p})^2} + \theta_i\theta_j\frac{\boldsymbol{p}'\boldsymbol{S}\boldsymbol{p}}{(\boldsymbol{\theta}'\boldsymbol{p})^3}\right]\sum_k \beta_k y_k.$$

Defining the factor shares

$$v_j \equiv \frac{p_j x_j}{C},$$

the Allen elasticities are

$$\sigma_{ij} \equiv \frac{\epsilon_{ij}}{v_j} = \frac{p_j}{x_i}\frac{\partial x_i}{\partial p_j}\frac{C}{p_j x_j}$$

Table 6 gives these Allen elasticities, evaluated for each of the four versions of the SGM model in which concavity and/or separability are imposed.[5] The consideration that motivated Diewert and Wales's development of the SGM model is that concavity is often not satisfied in demand data, and indeed here it is a binding constraint: it is not satisfied unless imposed. Thus the own-price effects are all negative only in the concavity-constrained versions of the model. Cross-substitution effects are sometimes negative, suggesting that there may be some degree of complementarity in the employment of player skills. This is consistent with the finding of Stewart and Jones (1998) based on a single-output Generalized Leontief model. These elasticities are not sensitive to the imposition of separability.

Table 7: Factor Input Elasticities With Respect to Output

	skill factor	separability not imposed: wins	separability not imposed: road attendance	separability imposed: wins	separability imposed: road attendance
concavity not imposed	experience	0.608	0.686	0.633	0.436
	hitting	0.213	0.281	0.231	0.159
	pitching	0.550	0.549	0.621	0.427
	stars	1.792	0.707	1.553	1.068
concavity imposed	experience	0.617	0.688	0.640	0.438
	hitting	0.226	0.296	0.247	0.169
	pitching	0.542	0.529	0.612	0.419
	stars	1.745	0.680	1.512	1.035

7.2 Factor Input Elasticities With Respect to Output

How much additional input x_i is needed to achieve a one percent increase in output y_k? For example, how much additional star status is needed to increase entertainment value? This is the elasticity

$$\frac{y_k}{x_i}\frac{\partial x_i}{\partial y_k}.$$

The relevant derivative is

$$\frac{\partial x_i}{\partial y_k} = g_i(\boldsymbol{p})\beta_k + b_{ii}\beta_k + 2\lambda_i \sum_j d_{jk} y_j.$$

These elasticities are presented in Table 7. All are positive, so that all player skills are normal goods in team production. The values are fairly robust across the alternative versions of the model, in particular being little affected by the imposition of concavity. According to the non-separability-constrained model, for example, a 1% increase in road attendance requires around a 0.7% increase in the number of stars on the team. The largest elasticities are for the stars-wins effect, so that it takes a relatively large increase in the number of stars to improve team performance. The smallest are for the relationships between hitting and the two outputs; it takes a relatively small increase in hitting skills to improve either the entertainment value or performance of teams.

7.3 Cost Elasticities

How costly is it to increase each of the outputs? This is given by the cost elasticities

$$\epsilon_k = \frac{y_k}{C}\frac{\partial C}{\partial y_k}.$$

Table 8: Cost Elasticities and Elasticity of Scale

	separability not imposed			separability imposed		
	cost elasticities			cost elasticities		
	wins	road attendance	elasticity of scale	wins	road attendance	elasticity of scale
concavity not imposed	0.441	0.452	1.120	0.455	0.313	1.302
concavity imposed	0.448	0.457	1.106	0.463	0.317	1.282

The relevant derivative is

$$\frac{\partial C}{\partial y_k} = g(\boldsymbol{p})\beta_k + \beta_k \sum_i b_{ii} p_i + 2\left(\sum_i \lambda_i p_i\right) \sum_j d_{jk} y_j.$$

The elasticity of scale—the percentage increase in the outputs generated by a 1% increase in all inputs—is the reciprocal of the sum of the cost elasticities:

$$\frac{1}{\epsilon_1 + \epsilon_2}.$$

These elasticities are reported in Table 8. All are positive, as they should be, and their values are robust across the various versions of the model. A 1% increase in either output individually implies a slightly less-than-1/2% increase in total costs. This translates into an elasticity of scale slightly greater than 1, so that that a 1% increase in all player skills results in a slightly-more-than-1% increase in team performance and entertainment value.

Although this point estimate is not greatly different from constant returns, the difference is significant. Constant returns to scale may be tested: the restrictions $b_i = 0$, $d_{jk} = 0$ make the SGM cost function linearly homogeneous in its outputs. Introducing these restrictions into our two-output non-separable demand system M_0 reduces it from 21 to 11 coefficients. A likelihood ratio test yields a test statistic of 86.268 which, for 10 degrees of freedom, strongly rejects the CRTS restrictions. (For example, $\chi^2_{0.01} = 23.21$.)

8 Conclusions

We have used data from Major League Baseball to estimate a system of factor demands derived from a multi-output symmetric generalized McFadden cost function. Underlying this approach is a conception of the factor market for players in which teams compete in an implicit market for player skills, paying hedonic prices for those skills. The estimated model is supported by elasticities that are plausible in sign and magnitude.

The adoption of the SGM functional form offers several advantages. First, it permits the global imposition of concavity. Second, it allows the use of hedonic prices even when they take on zero values. Third, it permits the recovery of all the parameters of the cost function even when data on costs are unavailable, as is the case here. Thus we have been able to use the recovered cost function to obtain fitted costs, which have in turn been used for several purposes: the calculation of Allen elasticities, the comparison of predicted and hedonic costs as a check on the internal consistency of the analysis, and to obtain isocost curves.

The question posed in the title of this paper has been investigated with information-theoretic and nested and nonnested testing methods. The information-theoretic approach indicates that the two-output model is preferred over any of the single-output versions, and nonnested testing rejects the single-output models (the least strongly rejected model being M_3 which is based on home attendance y_3, where the p-value is 0.0500). This calls into question the adequacy of analyses in sports economics that are predicated on the assumption that teams produce a single output—wins, attendance, or otherwise.

For researchers who nevertheless find themselves obliged to adopt a single output measure, the empirical evidence offers no strong basis for preferring either wins y_1 or home attendance y_3. In our results, both these single-output models have similar likelihood values and Vuong's test fails to reject the null that they are equally close to the true model. In the Introduction we argued that home attendance is the derived result of team outputs rather than itself being a direct output of the team production function—gate receipts being just one component of revenue, no different formally from other components such as merchandizing, parking, or broadcast revenue. But despite this a priori argument against home attendance as the definition of team output it is not rejected on purely statistical grounds relative to other single-output measures.

Although single-output models are rejected, the hypothesis of separability is not—at least at conventional significance levels. This suggests that it may sometimes be reasonable to view teams as producers of an output aggregate measured by a production index. Our analysis has been based on an index based on performance and entertainment value, measured respectively by wins and road attendance. However this has been taken as a maintained hypothesis, and other researchers will no doubt have their own views about the appropriate conception of team outputs and how best to measure them. Our purpose here is not to argue that our conjectures about team outputs are the only possible ones; on the contrary, hopefully this analysis will spur a deeper consideration of the proper specification of team production functions. Instead our goal has been to advance the view that team production processes should be regarded explicitly as multi-output ones—a view that may have implications for how researchers in the expanding field of sports economics investigate the questions of interest to them.

Notes

[1]The SGM cost function is not unique in permitting the imposition of concavity. An alternative is the Asymptotically Ideal Model (AIM) of Barnett, Geweke, and Wolfe (1991). The importance of the SGM model is that the imposition of concavity does not destroy the flexibility of the functional form. By contrast, Terrell (1995) demonstrates that the flexibility of the AIM is dramatically impaired by the imposition of concavity.

[2]It was these single-output factor demands that were estimated by Stewart and Jones (1998), using Generalized Leontief and translog functional forms. Since the latter expresses factor prices in log form it was estimated over just the subsample of hedonic price vectors consisting entirely of non-zero prices. The sensitivity of the results to alternative measures of the single output was investigated by considering both $y =$ wins and $y =$ home attendance.

[3]In the Diewert-Wales formulation the model is presented with the dependent variables expressed as input-output ratios x_i/y_i. Barring other considerations, this is the preferred form for estimation because it mitigates heteroskedasticity. However Vuong's test procedure used in Section 5.3 requires the alternative models to have common dependent variables, as is typical of most nonnested tests. Thus for the purpose of our implementation of his test the single-output models are estimated in the form shown.

[4]This is of course equivalent to evaluating the cost identity (7) using the fitted factor inputs $\hat{x}_{it}$ yielded by the estimated demand functions. We have checked that this equivalence is satisfied in our calculations.

[5]We do not bother to devote space to reporting the price elasticities ϵ_{ij} as they are not of significant independent interest. Approximate values may be obtained from the Allen elasticities using the cost shares given in Table 5. (The values would be approximate because using mean cost shares is not identical to evaluating the elasticity expression (13) at the point of variable means.)

References

Barnett, W.A., Geweke, J., and M. Wolfe (1991) "Seminonparametric Bayesian estimation of the asymptotically ideal production model," *Journal of Econometrics* 49, 5–50.

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

Diewert, W.E., and T.J. Wales (1987) "Flexible functional forms and global curvature conditions," *Econometrica* 55, 43–68.

Downward, P., and A. Dawson (2000) *The Economics of Professional Team Sports* (London: Routledge).

Economic Journal (2001) "Symposium on the Economics of Sport," ed. by S. Szymanski, 111 (February), F1–F84.

Ferguson, D.G., Jones, J.C.H., and K.G. Stewart (2000) "Competition within a cartel: League conduct and team conduct in the market for baseball player services," *Review of Economics and Statistics* 82, 422–430.

Fort, R.D. (2003) *Sports Economics* (Upper Saddle River, NJ: Prentice-Hall).

Fort, R.D., and J. Quirk (1995) "Cross-subsidization, incentives, and outcomes in professional team sports leagues," *Journal of Economic Literature* 33, 1265–1299.

Hall, R.E. (1973) "The specification of technology with several kinds of outputs," *Journal of Political Economy* 81, 878–892.

Jones, J.C.H. (1969) "The economics of the National Hockey League," *Canadian Journal of Economics* 2, 1–20.

Kumbhakar, S.C. (1990) "A reexamination of returns to scale, density and technical progress in U.S. airlines," *Southern Economic Journal* 57, 428–442.

Kumbhakar, S.C. (1994) "A multiproduct symmetric generalized McFadden cost function," *Journal of Productivity Analysis* 5, 349–357.

Leeds, M., and P. von Allmen (2002) *The Economics of Sports* (Boston, MA: Addison-Wesley).

Rask, K. (1995) "The structure of technology in Brazilian sugarcane production, 1975–87," *Journal of Applied Econometrics* 10, 221–232.

Rosen, S. (1981) "The economics of superstars," *American Economic Review* 71, 845–858.

Rottenberg, S. (1956) "The baseball player's labor-market," *Journal of Political Economy* 64, 242–258.

Sandy, R., Sloane, P., and M. Rosentraub (2004) *The Economics of Sport: An International Perspective* (London: Palgrave Macmillan).

Scully, G.W. (1974) "Pay and performance in major league baseball," *American Economic Review* 64, 915–930.

Stewart, K.G., and J.C.H. Jones (1998) "Hedonics and demand analysis: The implicit demand for player attributes," *Economic Inquiry* 36, 192–202.

Szymanski, S. (2003) "The economic design of sporting contests," *Journal of Economic Literature* 41, 1137-1187.

Terrell, D. (1995) "Flexibility and regularity properties of the asymptotically ideal production model," *Econometric Reviews* 14, 1–17.

Vuong, Q.H. (1989) "Likelihood ratio tests for model selection and non-nested hypotheses," *Econometrica* 57, 307–333.