Economics Letters 55 (1997) 293–299

economics letters

Measuring efficiency in the National Basketball Association[1]

Richard A. Hofler[a], James E. Payne[b,*]

[a]Department of Economics, University of Central Florida, Orlando, FL 32816, USA
[b]Department of Economics and Finance, Eastern Kentucky University, Richmond, KY 40475, USA

Received 25 November 1996; received in revised form 3 March 1997; accepted 5 March 1997

Abstract

This paper investigates how closely teams in the National Basketball Association play up to their potential. Using the stochastic production frontier model, we provide efficiency measures for each of the 27 NBA teams for the 1992–1993 season. © 1997 Elsevier Science S.A.

Keywords: Stochastic production frontier; Efficiency; National Basketball Association

JEL classification: Z00

1. Introduction

Sports fans frequently debate about which team is the best, usually by quoting won-lost records. Conversely, many coaches talk less about won-lost records and more about their teams 'playing up to their potential'. This paper wishes to answer the question, 'How well do teams in the National Basketball Association (NBA) play up to their potential?' Estimating a stochastic production frontier model will provide the answers. Since the authors are NBA fans, we chose those teams for our study. This analysis could easily be extended to other teams, both amateur and professional, however.[2]

Several researchers have estimated production functions to measure the relationship between team success (output) and performance inputs (see Scully, 1974; Zak et al., 1979; Zech, 1981; Porter and Scully, 1982; Jones and Walsh, 1984; Scott et al., 1985; Schofield, 1988; Hofler and Payne, 1996). Specifically, estimation of production efficiency in professional basketball by Zak et al. (1979) provides the basis of our inquiry. Our study extends their work in several areas. First, rather than use the production frontier methodology of Timmer (1971); Afriat (1972); Richmond (1974) we utilize

*Corresponding author. Tel.: (1-606) 622-1769.

[1] An earlier draft of this paper was presented at the meeting of the Kentucky Economic Association, Frankfort, KY, October 6–7, 1994. The authors appreciate the comments by Mark Berger, Frank Scott and an anonymous referee. The usual disclaimer applies.

[2] See Cairns et al. (1986) for a review of the literature. Goff and Tollison (1990) have produced an interesting edited volume on economic applications of different facets of the sports industry. Demmert (1973); Quirk and Mohamed (1971) provide detail analyses of the economics associated with professional sports teams.

PII S0165-1765(97)00083-9

the production frontier methodology advanced by Aigner et al. (1977). Second, while Zak et al. (1979) examine the production efficiency of five professional basketball teams, we examine efficiency in a cross-sectional analysis of all 27 basketball teams in the NBA for the 1992–1993 season.

Section 2 will present the production frontier model. Section 3 provides the empirical results while concluding remarks are discussed in Section 4.

2. Production frontier and efficiency

Efficiency will be measured by the application of the stochastic frontier methodology developed by Aigner et al. (1977). The stochastic frontier model which will be employed in this study was developed, and continues to be largely used, in a production context. It is natural to address efficiency in that case and easy to understand what it means. In the case of a sport, however, terms like production, frontier and efficiency may not have clear meanings. For this discussion, 'production' means a team's wins. The 'frontier' and 'frontier production' refer to the maximum attainable (potential) wins that a team can achieve, given its players, coaching and other circumstances. Finally, 'efficiency' describes how closely to its potential a team approaches.

In general terms, let

$$Y_i = \boldsymbol{X_i}\boldsymbol{\beta} + v_i \quad i = 1, \ldots, n \tag{1}$$

represent a non-frontier production function for each team where Y represents actual production (wins). $\boldsymbol{X_i}$ is a row vector of team-specific production determining characteristics and $\boldsymbol{\beta}$ is a column vector of regression coefficients. The error term v_i is assumed to be normally distributed with $E(v_i)=0$, and $\mathrm{VAR}(v_i)=\sigma_v^2$.

In addition, let u be a one-sided random error term that takes on only nonpositive values and represents the distance by which actual production falls short of the maximum attainable (frontier or potential) production. The model to be estimated then becomes

$$Y_i = \boldsymbol{X_i}\boldsymbol{\beta} + v_i + u_i \quad i = 1, \ldots, n \tag{2}$$

where u_i is also an error term specific to team i, and u_i is ≤ 0, $E(u_i)=\mu<0$, and $\mathrm{VAR}(u_i)=\sigma_u^2$. This equation models the actual production of each team. In the context of the model in Eq. (2), the function in Eq. (1) can be reinterpreted as each team's stochastic frontier, which it aspires to produce on so as to maximize its output

$$Y_{f,i} = \boldsymbol{X_i}\boldsymbol{\beta} + v_i \quad i = 1, \ldots, n \tag{3}$$

In other words,

$$[Y_{f,i} = \boldsymbol{X_i}\boldsymbol{\beta} + v_i] > [Y_i = \boldsymbol{X_i}\boldsymbol{\beta} + v_i + u_i]. \quad i = 1, \ldots, n \tag{4}$$

The relationship in Eq. (4) specifies that the team's potential production exceeds or equals its actual production. The error term u_i is a measure of the reduction in output experienced by the ith team due to inefficiency (failure to attain its potential number of wins). That is, the closer this value is to zero, the smaller is the degree of inefficiency, with $u_i=0$ implying no inefficiency (the team is playing at its potential). A related measure of inefficiency, the percentage of frontier output attained, is intuitively clearer and will be used in this study. It is a value that ranges between 0 and 100%, with 100% meaning complete efficiency (the team is playing at its potential).

The frontier model's parameters can be estimated by maximum likelihood based on the error term from Eq. (2)

$$\epsilon_i = v_i + u_i \quad i = 1, \ldots, n \tag{5}$$

where v_i iid $N(0,\sigma_v^2)$, $u_i \leq 0$, and v_i and u_i are assumed to be independent.

There are currently three choices for the distribution of the one-sided error component u:

(a) u is iid $N(0,\sigma_u^2)$ and truncated at zero (a 'half-normal' distribution),
(b) u is exponential, and
(c) u is iid $N(\mu,\sigma_u^2)$ where μ does not equal zero and truncated at zero (a 'truncated normal' distribution).

Researchers now understand that the first two choices are poorer than (c) because their modal values are nearly zero. These tend to underestimate the extent of inefficiency. Consequently, we assume u_i to be iid $N(\mu,\sigma_u^2)$ and truncated at zero from above. The model's parameters will be estimated by ML with the start values coming from initial OLS estimates.[3]

We wish to compare our results with those of Zak et al. (1979). Therefore, we estimate the same homogeneous Cobb-Douglas that they use.[4] Specifically, the production model is

$$\ln Y_i = A + \sum_{i=1}^{7} \alpha_i \ln X_i + \alpha_8 X_8 + v_i + u_i \quad i = 1, \ldots, n \tag{6}$$

and

Y = WINS = actual number of wins,
X_1 = FG = ratio of field goal percentage (+),
X_2 = FT = ratio of free throw percentage (+),
X_3 = OR = ratio of offensive rebounds (+),
X_4 = DR = ratio of defensive rebounds (+),
X_5 = A = ratio of assists (+),
X_6 = S = ratio of steals (+),

[3] See Aigner et al. (1977) for details.

[4] Zak et al. (1979) use the ratio of final scores as their dependent variable. This dependent variable is the season-long mean ratio of final scores in all of its games: the team's average final score relative to the average final score of all of its opponents. Using the ratio of final scores measures the relative closeness of the games. We performed all of the analyses in this paper using that variable (the results are available). However, we decided that using actual wins as the dependent variable would seem more sensible to most readers and make the output easier to follow. These results are presented in the paper. The conclusions drawn from both sets of analyses are nearly the same. This suggests that using wins instead of final scores has not distorted the information that is revealed here.

X_7 = TO = ratio of turnovers (-),
X_8 = BS = difference in blocked shots (+).
(Expected signs are in parentheses.)

The data was collected from the 1994 Sports Almanac published by Sports Illustrated (1994) for the 1992–1993 NBA season. The dependent variable measuring output is each team's actual number of wins during the regular season (the playoffs are excluded). The ratios of field goal and free throw shooting percentages are hypothesized to have positive impacts on team wins. These two variables measure the quality of a team's shooting. The ratios associated with offensive and defensive rebounds are also hypothesized to have positive impacts on winning. The team with better rebounding enhances its chances at scoring and winning while limiting the opposition's chances to score. The ratio of assists proxies ball handling skills and teamwork not measured by shooting percentages. Assists are assumed to increase wins. The ratio of steals and the difference in blocked shots each reflect defensive intensity which should have a positive effect on winning. Finally, the ratio of turnovers is hypothesized to adversely affect a team's wins.[5]

3. Empirical results

Table 1 contains the ML estimation results of the frontier. The F-statistic from the OLS step shows that there is a significant relationship between Y and the set of regressors. Therefore, the adjusted R^2 from the OLS step is nonzero and reveals that this model explains over 88% of the variation in winning. These results seem to show that only more defensive rebounds increase winning. However, this model suffers from severe collinearity, which can cause important variables to be insignificant. First of all, the condition number (a value of over 361) for the regressor matrix indicates severe multicollinearity. Secondly, the variance inflation factors and correlation coefficients indicate that five variables are involved in collinear relationships: field goal percentage, offensive rebounds, defensive rebounds, steals and turnovers.

Table 2 shows that the teams were winning at a high level of efficiency. The league-wide average was 89% during the 1992–1993 season. This is similar to the rating that Zak et al. (1979) found, 99.8%, for an earlier period for only five teams. Table 3 ranks teams by frontier wins: the teams with the most potential. No team was 100% efficient in realizing its winning potential. Denver and Minnesota were the best teams at using their abilities to the fullest and Washington and Dallas were the worst.

Notice how Table 3 clarifies the distinction between winning potential and efficiency in using that potential. Phoenix had the highest winning potential (it could have won 78 games) but it was 25th in efficiency. Its poor efficiency lowered its actual performance (62 wins) below where it could have been had it performed at even average efficiency (69 wins). Denver, conversely, had a below average winning potential (37 wins) but was the most efficient team. This efficiency still translated into poor actual performance because of its low potential. As one might expect, Phoenix, the New York Knicks,

[5] Zak et al. (1979) also introduce a binary variable for home versus away games. Since our study is a cross-section of each team's season averages we could not introduce a binary variable.

Table 1
Maximum likelihood estimates of frontier model (dependent variable is number of wins)

Variable[a]	Coefficient	t-ratio
Constant	36.586	17.964
Field goal %	15.789	0.205
Free throw %	−12.443	−0.248
Offensive rebounds	5.793	0.297
Defensive rebounds	115.860	2.606
Assists	−4.215	−0.340
Steals	17.723	0.860
Turnovers	−39.853	−1.181
Blocked Shots	0.011	0.538
σ_u/σ_v	8.077	0.304
$\sqrt{\sigma_v^2+\sigma_u^2}$	6.148	1.296
Variance components:		
Variance of 1-sided error		0.571
Variance of 2-sided error		37.223
Value of log-likelihood		71.288
R^2		0.918[b]
Adj R^2		0.881[b]
$F_{8,18}$		25.10[b]

[a] All independent variables are ratios and in logarithms except blocked shots. It is a difference (not a ratio) and in levels (not a logarithm).
[b] This is taken from the OLS step used to generate start values for the maximum likelihood step.

Houston (NBA champions the next season), Chicago and Cleveland, were at the top while Milwaukee, Philadelphia, Sacramento, Minnesota and Dallas were at the bottom.

4. Concluding remarks

This paper has attempted to investigate how closely teams in the National Basketball Association play up to their potential. Using the stochastic production frontier model, we provide efficiency measures for each of the 27 NBA teams for the 1992–1993 season. This line of research could easily be extended to other sports, both amateur and professional. Another avenue of inquiry is the extent to which coaching affects team efficiency.

Table 2
Descriptive statistics on selected variables 1992–1993 season

Variable	Mean	Minimum	Maximum	Cases
Actual wins	41.00	11.00	62.00	27
Frontier wins	46.00	18.00	78.00	27
Efficiency (%)	88.96	61.67	98.49	27

Table 3
Each NBA team ranked by frontier wins 1992–1993 season

Team	Wins		
	Frontier[a]	Actual	Efficiency
Phoenix	78	62	79.3%
NY Knicks	64	60	93.5%
Houston	63	55	87.5%
Cleveland	62	54	86.8%
Chicago	60	57	94.2%
San Antonio	58	49	85.0%
Seattle	57	55	96.8%
Portland	57	51	90.9%
Boston	54	48	89.2%
Atlanta	50	43	84.8%
Utah	49	47	96.9%
Detroit	49	40	81.0%
Charlotte	48	44	91.0%
Orlando	45	41	91.7%
Indiana	45	41	91.7%
LA Clippers	45	41	91.0%
NJ Nets	44	43	97.5%
LA Lakers	43	39	90.4%
Miami	41	36	88.3%
Denver	37	36	98.5%
Golden State	35	34	96.4%
Washington	32	22	68.7%
Milwaukee	30	28	92.7%
Philadelphia	30	26	86.7%
Sacramento	27	25	92.6%
Minnesota	19	19	97.8%
Dallas	18	11	61.7%
Averages	46	41	89.0%

[a] Frontier wins rounded to nearest whole number.

References

Afriat, S.N., 1972. Efficiency estimation of production frontiers. International Economic Review 13, 568–598.

Aigner, D., Lovell, C.A.K., Schmidt, P., 1977. Foundation and estimation of statistical production function models. Journal of Econometrics 5, 21–37.

Cairns, J., Jennett, N., Sloane, P.J., 1986. The economics of professional team sports: A survey of theory and evidence. Journal of Economic Studies 13, 1–80.

Demmert, H.G., 1973. The economics of professional team sports (D.C. Heath, Lexington, MA).

Goff, B.L. and R.D. Tollison, 1990. Sportometrics (Texas A&M University Press, College Station).

Hofler, R.A., Payne, J.E., 1996. How close to their offensive potential do National Football League teams play?. Applied Economics Letters 3(11), 743–747.

Jones, J.C.H. and W.D. Walsh, 1984. Mantia for a monopsonist: The World Hockey Association and player exploitation in the National Hockey League, unpublished manuscript (University of Victoria, Australia).

Porter, P.K., Scully, G.W., 1982. Measuring managerial efficiency: The case of baseball. Southern Economics Journal 48, 642–650.

Quirk, J., Mohamed, E.H., 1971. An economic model of a professional sports league. Journal of Political Economy 70, 1302–1319.
Richmond, J., 1974. Estimating the efficiency of production. International Economic Review 15, 515–521.
Schofield, J.A., 1988. Production functions in the sports industry: An empirical analysis of professional cricket. Applied Economics 20(2), 177–193.
Scott, F.A., Long, J.E., Somppi, K., 1985. Salary vs. marginal revenue product under monopsony and competition: The case of professional basketball. Atlantic Economic Journal 13(3), 50–59.
Scully, G.W., 1974. Pay and performance in Major League Baseball. American Economic Review 64(6), 915–930.
Sports Illustrated 1994 Sports Almanac (Little Brown, New York).
Timmer, C.D., 1971. Using a probabilistic frontier production function to measure technical efficiency. Journal of Political Economy 19, 776–794.
Zak, T.A., Huang, C.J., Siegfried, J.J., 1979. Production efficiency: The case of professional basketball. Journal of Business 52(3), 379–392.
Zech, C.F., 1981. An empirical estimation of a production function: The case of Major League Baseball. American Economist 25(2), 19–30.