Dilger, Alexander

Working Paper

Never change a winning team: An analysis of hazard rates in the NBA

Wirtschaftswissenschaftliche Diskussionspapiere // Ernst-Moritz-Arndt-Universität Greifswald, Rechts- und Staatswissenschaftliche Fakultät, No. 03/2002

Provided in cooperation with:
Ernst-Moritz-Arndt-Universität Greifswald

Suggested citation: Dilger, Alexander (2002) : Never change a winning team: An analysis of hazard rates in the NBA, Wirtschaftswissenschaftliche Diskussionspapiere // Ernst-Moritz-Arndt-Universität Greifswald, Rechts- und Staatswissenschaftliche Fakultät, No. 03/2002, http://hdl.handle.net/10419/48885

Ernst-Moritz-Arndt-Universität Greifswald

Rechts- und Staatswissenschaftliche Fakultät

Wirtschaftswissenschaftliche Diskussionspapiere

Never Change a Winning Team

An Analysis of Hazard Rates in the NBA

Alexander Dilger

Diskussionspapier 3/02

Mai 2002

ISSN 1437-6989

PD Dr. Alexander Dilger
Ernst-Moritz-Arndt-Universität Greifswald
Rechts- und Staatswissenschaftliche Fakultät
- Wirtschaftswissenschaften -
Lehrstuhl für Betriebswirtschaftslehre, insbesondere Personal- und Organisationsökonomie
Postfach
D-17487 Greifswald

Tel.: +49-03834/86-2474
Fax: +49-3834/86-2470
e-mail: dilger@uni-greifswald.de
Homepage: http://www.rsf.uni-greifswald.de/~dilger

Never Change a Winning Team

An Analysis of Hazard Rates in the NBA

Abstract

We estimate Cox models to determine proportional hazard rates in professional basketball, concerning leaving the league or changing the team by using a database covering all players of the NBA in the 90's. We predict and confirm that league-hazards depend on a player's performance. A team-switch, however, cannot depend on low performance itself because there has to be a team willing to accept the (new) player. Accordingly we find that a good scoring performance and an intense use of a player reduces the probability of a team-switch, whereas high salaries or non-scoring performance do not.

Acknowledgement

Many thanks to Bernd Frick and Joachim Prinz. The paper also benefited from presentations at my habilitation colloquium in Greifswald and at the 16th Annual Congress of the European Economic Association in Lausanne (EEA 2001). I wish to thank all participants and remain responsible for all mistakes alone.

1. Introduction

There is a rising interest in the economics of sports. See as examples the foundation of the Journal of Sports Economics last year, Rosen and Sanderson's (2000) recent article about the professional sports labor Market or Kahn's paper in the Journal of Economic Perspectives reviewing the sports business as a labor market laboratory (Kahn 2000). Professional team sports proposes a unique opportunity to test hypotheses about managers, workers and business companies in general, which are not directly accessible since a lack of empirically data. As a result the economist has to ask where and how to procure "functional equivalent" data that makes an inspection of testable hypotheses feasible. A second reason for the interest in the economics of sports is the economic importance of sports in itself. Today sport has become a big business, with players salaries, franchises values and merchandising gaining in magnitude from season to season. Sport is only one, but increasing part of the entertainment industry among others like the music or movie business with revenues of billions of dollars.

There are quite a lot of studies concerning the salaries in professional team sports. Most of them use the standard Mincer-type wage equation that tries to identify player salary determinants where performance, seniority and race or ethnicy are the major independent variables in their estimates, see Scully (1974a), Jones and Walsh (1988), Harder (1992), Idson and Kahane (1999). The aim of all these studies was to explain the maximum of variance with theoretically meaningful regressors. However, for most players the exact salary in one single year is of minor importance. They are more interested in playing on in further seasons. Whether a player earns $9 millions or $10 millions per year is less important than the question whether one can earn such an amount next season again or not. To analyse the reasons for staying in or dropping out of the league it is convenient to estimate hazard rates implicating the probability for athletes' leaving the league. There are only a few studies investigating hazard rates using data from a professional sport i.e. Chapman and Southwick (1991), Ohtake and Ohkusa (1994), Ohkusa and Ohtake (1996), Borland and Lye (1996). Chapman and Southwick do this for Major League Baseball in the US, Ohkusa and Ohtake study Baseball data from Japan, whereas Borland and Lye analyse hazard rates for football in Australia. For literature on hazard rates using company data see i.e. Campell (1993) and Lazear (1999), while Booth, Francesconi and Garcia-Serrano (1999) use work-history data. Regarding hazard rates we are interested in a different Major-League Sport, the National Basketball Association (NBA), using a database covering all players in the 90's. For the prior decade there are two analyses of NBA-league-hazard by Staw and Hoang (1995) and Hoang and Rascher (1999).

Besides analysing the risk (hazard) of leaving the NBA, it is our particular concern to estimate a model which considers the movement of players between teams in the league through trades and free agency as suggested by Ohkusa (2001). To the best of our knowledge, we are the first to examine team-hazard in addition to league-hazard. Intuitively one might expect player performance figures being the decisive argument for survival in the league, things are more complex at the team level. No team desires a poor performing player, hence he has to quit the league. Every team however wants (needs) a high quality player, consequently he should remain in the league. But should this particular player also stay with his original team? This question depends whether another team has a greater interest for the player's services, maybe just for pecuniary reasons i.e. in the case that his talent is too expensive for his old team, or if he performs a better job in the new than in the old club.

The rest of the article is organized as follows: Specific hypotheses about league- and team-hazard are generated in section 2 along with some theoretical considerations. Following this, in section 3 we describe our dataset and the methodology we have used, especially the proportional Cox model as estimation technique. In section 4 our estimation results are reported. The paper ends with some concluding remarks.

2. Theoretical Considerations

We are interested in the hazard rates for players in the NBA. Hazard can be defined as leaving the NBA. This kind of hazard we denote as league-hazard, whereas hazards for quitting the team we call team-hazards. Firstly, we consider league-hazards and secondly we investigate team-hazards.

From a theoretical point of view we would expect that there is one and only one decisive determinant of the league-hazard. This is the performance of a player. We would anticipate that players with high performance have a low hazard rate, those with bad performance should indicate a high hazard rate. High performing players are valuable in the NBA, therefore they should stay in the league. Low performing players are less valuable, especially if their performance is lower than the expected performance of newcomers or other league entrants from outside the NBA.

There are three reasons why this supposed correlation between (high) performance and (low) hazard may not hold in other professions. First, high performing workers often have better outside opportunities than low performing employees, but this is of no relevance for NBA-players. The compensation in the NBA is much higher than that of any reasonable alternative job including other basketball leagues (with the exception of coaching a team after the player's career). Even the actual minimum wage of about $300,000 annually for 1999/2000 rookies is much higher than what any of the players could expect offside the basketball court. Therefore, it is irrelevant if better basketball players are also better in any other jobs. Every player wants to stay as long as possible in the NBA. This is true for monetary reasons alone. Moreover, most of the players can be assumed to have a preference for playing basketball over doing other things.

Ohkusa (2001, p. 83) argues for Japanese professional baseball players that they "can continue as professional baseball players if they sharply reduce their wages and trade to another team. However, they can choose not to play and so quit. Quitting is unambiguously voluntary." We think this is not true for professional basketball players in the NBA. The relative high minimum wage induces players to receive it instead of quitting, but profit-maximizing teams are not willing to pay this price for bad performance. This is not the case for baseball players in Japan because a minimum wage does not exist and their salaries are generally much lower.

The second reason for a high negative correlation between performance and hazard present in basketball and baseball as well is that the production functions in all team sports are of a limited form. It is ineffective having hundreds of idle players if only five players can form the squad that is allowed to be on the court in basketball and nine per team on the baseball field. The marginal yield even for well performing players is zero if they cannot be used. Moreover, the allowed maximum number of player's in a NBA-team at the same time is 12.

The third explanation for observing a stronger relationship between performance and hazard in the NBA as well as in all professional sports compared to other professions is that problems of asymmetric information in general and of moral hazard in particular are much lower in the former than in the latter. In contrast to most other workers is the professional basketball player's performance highly observable and transparent. All games are open to the public, all moves can be watched on TV, recapped by video again and again and player statistics are almost costlessly available through the Internet. The coach can monitor what the players do, owners can investigate it as well as the competitors and external analysts can. There are a lot of numbers to measure performance and we use some of them for our own analysis (see section 3). Even if one can argue about the actually used performance numbers, everyone is free to construct better measures by the available evidence.

We do not say that the public behaviour on the field is everything that matters. Some areas like the professional training behaviour, the use of drugs, doping or player's leisure-time activities are simply not perfectly visible for the monitoring institution. In this sense agency theory assumes that the individual player is better informed about his level of effort motivation or real physical condition (shirking through simulations etc.) than his team and trainer are. Rising tenure however reduces some of the team's information deficit about the player, whereas other clubs do not profit from that learning process. In addition, players jeopardize their reputation capital through postcontractual opportunism: Shirking, bad performance, deficient co-operation and unfair behaviour on or off the court will definitely not improve future contract terms or impress other potential employers. Hence, incentives to shirk and malfeasance are mitigated. In general we assume that most information is symmetric, especially actual performance is essential and relatively easy to monitor.

A further consideration is that due to the extreme high remuneration in the NBA (nearly) no player wants to go voluntarily. Thus we expect most of the quits to be coerced which results from minor performance statistics. Instead, too high salaries cannot be problematic for high quality players, because their salaries can be decreased until they equal their marginal productivities.

In order to examine the theoretical considerations just made, we translate them to the following hypothesis of empirical testing:

Hypothesis 1: The league-hazard depends mostly if not inevitably on the player's performance, that is the better the player's productivity the lower his risk of departing the league.

Investigating team-hazards is of different nature. Will players swing to another team for just the same reason they quit the league, which is for poor performance? Two kinds of team exits have to be distinguished: A player can drop out of the league while leaving his team or he can change teams inside the NBA. If the team-hazard is also a league-hazard, we would expect both things happen due to poor performance. Things, however, are more complicated in the case of a team-change. Teams are willing to give away low quality players, but why should any club pick up bad players? It takes two to tango! This could be the case if performance is team specific. How well a player works depends on his interaction with his fellow teammates but also on his relationship with the coach. Regarding this aspect Jovanovic's (1979) matching theory is highly relevant.

Undoubtedly, the team's most important capital is it's pool of potential player talent. This player performance potential varies however from team to team. On the one hand a player may warm the bench during the whole season in team X, whereas his better co-workers form the starting-five, while on the other hand he would be a regular player in team Y. Another reason for trading a player may be his salary. A player might just be too expensive for one team but not for another, even if performance stays constant.

In this respect, some institutional NBA characteristics have to be mentioned. Usually, most of he players sign contracts of at least one year which guarantees them a certain wage. Even if their performance becomes lower than expected, they have to be paid for the duration of their negotiated contract. Therefore, teams can be better off by trading low performing but overrewarded athletes, as long as their value is higher in a new than in the incumbent team. Next, player trades can only occur if they do not end up more than $100,000 of the acquired team's salary cap. The salary cap is the maximum dollar amount teams can spend on player contracts. For NBA terms, however, $100,000 are next to nothing. The reason why there are trades nevertheless is that the NBA employs in contrast to other Major-League Sports a soft cap, which is easily circumvented. Both, the salary cap and the trading rules are full of exceptions, which are sharply used by the teams. Player trades can have the form of a simple exchange of players among at least two teams, but trade packages are more regular. For example, a club acquires one top player from another team in exchange for two medium quality players and a future first round draft pick. Additionally, players can be traded for cash and cash can be included in trade packages as well. However, the amount of cash is limited to $3 millions. Regarding these facts teams might be forced to trade good but expensive players, even if they do not receive identical quality in return. This happens in case the team is money constrained and cannot pay the salaries for it's pool of talent or if a team has more high quality players than it needs due to the limited production function. Accordingly, we expect the following hypothesis to be true:

Hypothesis 2: The team-hazard that is a team-change depends on the player's team-specific performance and also on his salary. Therefore, poor team-specific productivity or a high salary rise a player's risk to change the team.

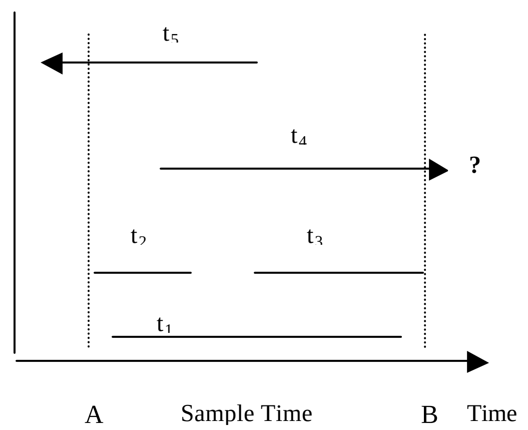
3. Data and Methodology

We use longitudinal data to test empirically the above mentioned hypotheses. The database we accumulated is hand-collected and is primarily drawn from various issues of the *"Sporting News Official NBA Register"* (1993/94-1999/2000) and the *"Sporting News Official NBA Guide"* (1996/97-1999/2000). Furthermore, data for the first three seasons, 1990/90-1992/93, were gathered from *"The Official NBA Basketball Encyclopedia"* edited by Villard Books 1994. Moreover, player salaries and contract information is mostly obtained from Patricia Bender's website at http://www.nationwide.net/~patricia/.

The raw dataset consists of all players that appeared in at least one of the 82 regular season games in any of the NBA-seasons 1990/1991-1999/2000. The total number of observations is 4,522 with some players being active in all 10 seasons and others in only a few or in just one of them. Sometimes there are two or more observations of the same player in one season if he changed the team during the season (see below). The current 29 teams employed 996 different basketball players over the past 10 years. While single performance figures like games played, minutes, field goals, free throws, three points, rebounds, assists, blocks, turnovers, steals, fouls and individual characteristics such as league experience, tenure, a player's age and all star games participation are available for all athletes, this is unfortunately not the case for player salaries and player's race. The former information is missing for approximately four percent and the latter is absent for one percent of the players. Complete player information is readily accessible for 4326 observations.

One feature of our data set is that it contains censored observations, for a more in depth treatment see Kiefer (1988). Censored data arise when a player occurs at one edge of our 10 years sample. On the one hand, we have left-censored observations, that is some individuals have experienced the event of interest (first year in the NBA) prior to the start of our study (before the 1990/91 season). Fortunately, we do know when these players started their career in the NBA. On the other hand, there is also the possibility of right-censored observations indicating that all we know is that the players are still employed by the league in the 2000/2001 season, that is after the ten analysed years. Certainly there are also athletes who begin and quit their job during the time of analysis, see Figure 1.

Figure 1: Spells and Censoring



A: Start of NBA sample (season 1990/1991)

B: End of sample (season 1999/2000)

t_1, t_2, t_3: Completed spells

t_4: Right-censored spell, t_5: Left-censored spell

Special attention must be paid to players that left the NBA and returned after a break. We observed some of such cases, but do not know whether and which players might have played once before in the 80's, left the league, came back or whether players who departed the league in the 90's will re-join the league in the new millennium. To solve this problem we decided to count each uninterrupted spell of seasons as a distinct subject. Therefore, it is possible that the same person is counted as two different subjects if he left the NBA for a while. Subsequently, our league-hazard estimation includes 1,097 subjects although there are only 996 different players in our data set. Tests of the team-change model report 2,039 subjects. Accordingly, these are uninterrupted season spells for one player within a particular club.

Methodologically, we are using Cox's (1972) semi-parametric proportional hazards model, where the problems of censoring can be solved. The Cox model does not limit the pattern of the hazard rate like parametric models with distributions such as the Weibull, exponential or log-logistic do. In the Cox model, the conditional hazard function, given the vector z of covariate values at time t or the corresponding time interval, is assumed to be of the form

$$\lambda(t \mid z) = \lambda_0(t)\exp(\beta z),$$

where β is the vector of regression coefficients, and $\lambda_0(t)$ denotes the baseline hazard function, which is the hazard function when all independent variable values are equal to zero. No particular shape is assumed for the baseline hazard, that is it has to be estimated non-parametrically. The parameter estimates $\hat{\beta}$ are obtained by maximizing the partial log-likelihood function

$$\ln L = \sum_{j=1}^{D} \left\{ \sum_{k \in D_j} z_k \beta - d_j \ln\left[\sum_{i \in R_j} \exp(z_i \beta) \right] \right\},$$

where j indexes the ordered failure times t_j, D_j is the set of d_j observations that fail at t_j, and R_j is the set of observations that are at risk at time t_j. Only event times (dropping from the league or the team) contribute positively in the first expression to the partial likelihood. However, both censored and uncensored observations appear in the subtracted expression, where the sum over the risk set includes all players who are still at risk immediately prior to t_j. Tied failure times when two or more hazards occur at the same time are handled by the Breslow (1974) method, see also Peto (1972). This approximation uses the largest risk pool for each tied failure event that means all failure events are included in the risk pool because the order of failures is not known. For calculating the variance-covariance matrix the robust method of Lin and Wei (1989) has been chosen. For further explanation of the Cox model see Kalbfleisch and Prentice (1980), Kiefer (1988), Kahn and Sempos (1989), Altman (1991), Selvin (1996) or Klein and Moeschberger (1997).

One problem that arises is that some players played for more than one team in a single year. Indeed, there are several players who changed the team once, twice or even played for three clubs within one season. Because our data do not tell at which time these changes occurred or whether a player left the league within the season or at its end, it is necessary to drop multiple cases of one player in a single season. For estimating the league-hazard, always the last case of one season remained in the data set. Only this case is at risk for leaving the league, at least ex post. Other team separations of a particular player during one season were eliminated.

For estimating the team-hazard, things are more complicated. The last case in one season could be one without team-change, if the player changed within the season but then remains in his new team. Therefore, we selected from multiple cases of one player in one season the last one with a team-change at its end. This can be the last case of the season or the last but one. The last but one must end with a team-switch because otherwise it would not be followed by an additional case in the same season.

Table 1 exhibits the descriptive statistics for the first of these data sets (those for the second one are nearly identical). We divided players' performance statistics in two groups. First we created the variable scoring performance. Scoring performance is the accumulation of all points (single points

Table 1: Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Max.	No. of Cases
Scoring performance	0.3801	0.1468	0	3	4,117
Non-scoring performance	0.2537	0.0980	-1	2	4,117
Personal fouls	0.1130	0.0778	0	2	4,117
Team years	2.4367	2.2278	1	16	4,117
Team years squared	10.6936	24.2466	1	256	4,117
League years	5.0675	3.8838	1	22	4,117
League years squared	40.7600	56.2703	1	484	4,117
Games	54.4460	26.4115	1	82	4,117
Win-percentage	0.4948	0.1673	0.13	0.88	4,117
Draft number	29.4341	19.8682	1	59	4,117
Salary (ln)	13.6880	1.3914	7.78	17.11	3,982
Roster	15.0189	1.4501	11	19	4,117
Colour	0.7780	0.4156	0	1	4,072
Height in cm	200.8762	9.7893	160.02	231.14	4,117
All star games	0.1567	0.3635	0	1	4,117

through free-throws, double points from field goals and the three points shots which provide a bonus for scoring from a long distance beyond 23 feet 9 inches) a player scored in one season divided by the minutes he played in the whole season. In addition, we follow Harder (1992) and apply a component measure which controls for rather team-contributing player behaviour, the non-scoring performance. This measure is computed by adding an athlete's total amount of rebounds, blocks, steals and assists, subtracting the number of turnovers and dividing this result by his minutes played on the court. The addition of these "hazzle statistics" is justified by the consideration that they support and prevent scoring. In particular, rebounding and blocking are thought of as ungrateful jobs, thus the term "hazzle statistic". We also examined different measures of the non-scoring performance but without great differences in the results. One may wonder the negative minimum on this variable, but Gheorghe Muresan, center with the New Jersey Nets had indeed more turnovers (negative performance) than rebounds, blocks, steals and assists (positive performance). We did also standardize, on the minute basis, players' violations against the rules of the game, that is players who compete unfair against their opponents are sanctioned with a personal foul.

The database indicates that on average a player stayed somewhat more than 2.4 years with his current team at the time of observation. This is not identical with the average length of a career in the NBA because players can play there again next year, whereas players already gone are not observed any longer (at least in the same team). This and the following variables' squared terms are entered to capture possible non-linearity. The variable "league years" counts a player's number of uninterrupted years employed in the NBA at the observation time.

From the Table 1 can be seen that on average players perform 54 games in a 82 regular season period. We limited our investigation to the regular seasons since only 16 of the actual 29 NBA teams will qualify for postseason playoff games. Due to the league's lockout in the 1998/1999 season only 50 regular games were played in that year. We weighted the games of that season accordingly with the factor 1.64.

We did also gather data on the team basis. At the end of a season, each team is ranked by it's win-percentage that it traced during the season. Under the time of our investigation the Chicago Bulls were in 1995/1996 the team with the best win-percentage (they won 88 % of the 82 games) whereas the Dallas Mavericks in 1992/1993 were the worst performing team in our analysis (they lost 87 % of the games).

One particular aspect in all major-league sports is the draft mechanism, which is held yearly at the end of the season. The draft is the major tool for the franchises to secure new talent or improve a team's future performance after a losing season. The rules of the draft dictate the order in which teams select amateur college team-sport players, see Staw and Hoang (1995). The team with the worst win-percentage of the past season has the first chance of picking the most talented college player in the "draft-lottery". The remaining teams then elect players in the inverse order of their prior regular season records, with the best team picking last in each round. Therefore, a low draft number indicates good, a high number not so good talent. In 1985 the NBA executed 10 draft rounds but reduced it to two rounds since the 1989/1990 season. Since our database contains players who were drafted before the 1989/1990 season it is straightforward that some players were picked with quite high numbers. Since most of our data comes from the 90's – the time the NBA switched to the trimmed (two) rounds – we coded draft picks that exceeded the highest possible draft number after the recent team expansion to Canada (the Toronto Raptors and the Vancouver Grizzlies) in 1995 as 59. In other words since there are two rounds today with 29 teams participating in the draft mechanism, 58 players are recruited in this fashion every year. Hence, higher draft picks from earlier rounds do all receive the draft number 59. The same holds for all players that were not drafted at all, especially foreigners. We did yet another change of the original draft number in case that more than ten years have past since the drafting. Then we added in the eleventh and every following year eight numbers to the draft number, until the maximum of 59 was reached. This may sound arbitrary, but this depreciation is necessary to fulfil the conditions of the proportional hazard model and it has also a meaningful interpretation, see below in section 4.

One of the big advantages of this study is the salary information used for the majority of NBA players. We do only miss salary data for 135 observations in the league database. In the following regression analysis we use the natural logarithm to flatten the function out. The average *absolute* salary for a player who was active in the NBA over our study period is about $1,550,000. Was this figure somewhat more than $900,000 in 1990/1991, it is almost $2,8 million today. At present there are 80 NBA players earning more than $5 million per season. In fact, Michael Jordan made about $33 million in his last NBA season 1997/1998. But with so much money to put toward "Superstars", marginal players are often compensated less than the minimum wage, which was $120,000 for rookies in 1990/1991 and $301,875 in 1999/2000. Players making far less than the respective year minimum salary are signed to a so called 10-day-contract. This is a contract that lasts for 10 days. If a player's contract is for less than the full season, the salary he gets is in proportion to the amount of seasons he has already played in the league. So if a 10-day-contract covers three games for the team, then the player's contract (assuming he is a rookie) will be 3/82 of the minimum wage of $301,875 in 1999/2000, or $11,044.

As mentioned above, the league's lockout in 1998/1999 took almost half of the regular season's time, reducing players' income by 39 %. We weighted the salaries of this season accordingly with the factor 1.64 as we did for the number of games played, because we are not interested in the absolute amount of money but the influence of salaries on the hazard rate. For the same reason, we weighted all salaries with the average salary of all years divided by the average salary of the particular year. This is in response to correct for the salary explosion in the last ten years by the factor three. Because salaries in the NBA rose much faster than normal income or inflation, we did not use consumer price or wage indexes as applied in other studies to correct for nominal differences between years.

We did also count the number of players forming a team during the season. NBA rules say that each team is allowed to have 12 players on it's roster but must suit up eight players for each game, while only five – two guards, two forwards and one center – are actively playing on the court. Our data set however has on average 15 athletes in one team, because the clubs are allowed to change their active roster between games through trades. This is not only true for tactical reasons but these additional players do also act as an insurance against too many player injuries within the season. Moreover, these "passive" players do function as a motivation instrument for the whole squad. They produce competition among the starting-five and especially among benchwarmers through the fear of replacement. Furthermore, they can simply be regarded as a reserve for trading players.

A dummy variable signals the colour of NBA professionals. We identified the race of each player by visual examination of his picture in the NBA Official Register. There are considerable more black players (77.8 %) in the league than white ones (22.2 %).

Instead of using position dummies we use a players size in the following regression since there might be a correlation between a players height and exiting the league/team. Staw and Hoang (1995, p. 479) argued that big players who typically occupy the forward and center positions are more likely to have a high number of rebounds and blocked shots. Therefore one might expect that rebounds and blocked shots would emerge as a performance factor for forwards and centers, rather than the smaller guards. We did also regressions with position dummies instead of the players' height. Their effects have been insignificant, however.

With the help of the all star game variable we are able to distinguish "superstars" from others. Although normal players are definitely important for every team, stars are an essential factor for the team's market value as they attract more fans besides their contribution to winning. The status as a (super)star is reflected in the all star games variable, used here as a dichotomous one. It was coded 0 if a player did not participate in the yearly held exhibition game where the best players from the

West compete against the best players from the East and coded 1 if he was at least once a member of the all star happening. Certainly there are players with several all star games. However, we only use the dummy variable of at least one participation because the actual number would violate the proportional hazards assumption of the Cox model (see below in section 4).

The length of a player's contract with his team is not included in the regressions because we have these information only for almost half of the total cases. We did also not include complete data about players not playing for injuries, because they would produce counterintuitive results. Injuries seem to lower hazard rates. This can be explained by the fact that only good players remain in the league or a team in case of injury whereas others drop without notice. Finally, we tried and then decided not use available data about the age and duration of player's professional career (inside and outside of the NBA) because of problems of multicollinearity between these variables and the years played in the NBA.

4. Empirical Results

The results of testing the Cox model for the league-hazard are presented in Table 2. Scoring performance has a negative sign and is statistically significant at the 1-percent-level. For example scoring one point more per minute reduces a player's probability of being cut from the league by 29 %. The variable non-scoring performance has also a significantly and even greater negative impact on the hazard rate. Producing one more unit of this index diminishes the propensity of being fired from the NBA by 45 %. Personal fouls are significantly positive sloped. One more of them every minute increases the likelihood of being released from the league by 48 %. If a player remains longer in one team, the hazard-rate decreases, but this effect of team-specificity is concave, because the quadratic term has a positive sign.

Table 2: League-Hazard

Variable	Coefficient	Std. Dev.
Scoring performance	-0.3387817**	0.1299339
Non-scoring performance	-0.5987001**	0.2064618
Personal fouls	0.3917943**	0.1411024
Team years	-0.1158622*	0.0584928
Team years squared	0.0108804*	0.0049878
Games	-0.0275126***	0.0015874
Win-percentage	-0.8783347***	0.1826118
Draft number	0.0083579***	0.0017707
Salary (ln)	-0.0980379***	0.0189498
Roster	-0.0058813	0.0205072
Colour	0.0061149	0.0629343
Height	-0.0017745	0.0028152
All star games	-0.2555442+	0.1419389

Number of subjects = 1,097

Number of hazards = 732

Number of observations = time at risk = 3,941

χ^2 = 1,055.37***

McFadden R^2 = 0.0774349

+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001, two-tailed tests

The variable representing the highest statistic significance (z-value = -17.332) is the number of games that a player was on the field during a season. The more games a player is in action the higher his chance of playing inside the NBA in the succeeding year. A team's win-percentage is also an important determinant in reducing the hazard-rate: Winning teams keep their members in the league. Further, a higher draft number results in a higher hazard-rate, whereas higher salaries improve the probability of remaining in the NBA. Team size has an insignificantly negative effect on the hazard-rate, whereas black colour insignificantly raises this rate. Also insignificant is the negative effect of a player's height.

Finally, the participation at one or more all star games significantly (but only on the 10-percent-level) reduces the hazard rate elucidating that teams try to keep high quality and "entertaining" players. Not shown in the table is that different tests of the proportional hazards assumption have been passed for the overall model and for each variable. We estimated also with different year- and team-dummies without much effect (the adjusted McFadden R^2 is falling).

How do these results match with hypothesis 1? Direct performance measures like scoring performance, non-scoring performance and personal fouls have the expected signs. We understand the positive coefficient on the personal foul variable as the anticipated sign because fouls in general harm a team's output in form that the rival team gains ball possession and/or free throw attempts. Hence, fouls are clearly a sign of negative player performance, which raises the probability of being ejected from the league. This confirms the hypothesis. However, the statistically significance of the games played variable is much higher than that of the direct performance measures. Nevertheless, this does not mean that the amount of games played is more important than performance figures displayed while playing. The games played are not independent of the player's performance. Indeed, the number of games could be a better indicator of a player's performance than the scores per minute, the non-scoring measure and the fouls per minute. The reason is that the coach always tries to set up his best available players – usually the "starting-five" – whereas performance measures per minute only illustrate the average productivity in the minutes a player was on the field. It is an economically reasonable assumption that performance is concave over time in respect that marginal performance is decreasing. Ceteris paribus we might observe "marathon-players" performing poorer than someone only playing a few minutes. If high quality players are on court for the majority of games and do also receive most of the playing time by the coach, then they do not seem to be so much better in their scoring or non-scoring performance per minute as they really are compared to their inferior teammates who are hardly ever, but well-rested on the field. We interpret the number of games played as performance measure and summarize its significant negative effect on the league-hazard as confirmation of hypothesis 1.

We see the time a player spends in a particular team as a good method to determine a player's team specific human capital and consider it as an indirect performance indicator. The win-loss-record measures team performance and a low draft number predicts prospective star potential. If most playing talent is innate or acquired in early years, we expect that playing quality at the beginning of one's career in the NBA as expressed by the draft number, remains an important factor over an athlete's whole career. Regarding this aspect we do not agree with Staw and Hoang (1995) who interpret the long lasting effect of the draft number as a sign of (irrational) regard for sunk costs. This could only be true if their other performance measures included in the regression are perfect, but they are of no better quality than ours. What we have found is that the effect of the draft number is not stable over time. The assumption of the proportional hazard rates over time is violated for the draft number when the number of years is higher than ten. Therefore we included a depreciation of eight numbers per year after the tenth year in our draft number variable (see section 3 above) which then realizes the proportional hazard conjecture of the Cox model. This finding emphasizes the significance of a player's initial productivity (talent) in the first years of his career, but this is lost and replaced by his current performance if the player survives for more than ten years in the NBA. Staw and Hoang (1995) do not test the proportional hazard assumption of the Cox model, but their study did however, only observe players being in the league for the maximum of eleven years. We found that for players with shorter league duration this central assumption was not violated.

Higher salaries significantly reduce the hazard rate. This would be very surprising if the direct performance measures are perfect. Why should a team pay more for a player of equal quality and performance? High priced players are on the verge of being sacked by a rational thinking manager. The negative relationship between the likelihood of quitting the league and high player salaries can be explained by the fact that well compensated athletes' perform better, even when controlling for the productivity measures applied in our model. For instance, a 10 % higher paycheck improves the chance of staying in the NBA by nearly 1 %. Intuitively, it is perfectly rational that high productive workers are paid better and maintain longer in the league. A team's top but overrewarded player is traded due to economic reasons, which requires that the new team has enough room below the salary cap. In case the player cannot find a new club that accomplishes the latter mentioned aspect - or does simply not accept the wage he calls for - the player has two choices: First, both parties agree on a lower salary and second, the player separates from the NBA. Certainly, regarding his opportunity costs he is better off choosing the lower salary. Who has already a low salary cannot do this to prevent quitting the league due to the minimum wage.

There is no theoretical association between performance on the one hand, roster size or a player's race on the other, accordingly we do not find any significant effects. The hazard for coloured players is higher, but in contrast to the Hoang and Rascher (1999) findings, there is no significant discrimination. Perhaps this can be explained by the fact that our data covers the 90's, theirs the 80's.

Finally, the participation of an all star game considerably reduces the hazard rate, namely by 23 %. Definitely, the all star game variable accounts for stardom player productivity. Moreover, fans decide by electing athletes into an all star game, which stresses a player's popularity apart from the mere athletic performance in professional basketball. Similar to the draft number analyses, the proportional hazard assumption is violated when the "real" numbers of all star games participations are counted for. The problem with the all star game variable is that players with more than a handful of all star games tend to be quite old, which indicates that these players near the end of their working lives.

Besides testing the Cox model to identify the coefficients of our independent variables, it is possible to compute the baseline hazard function and the baseline survival function. The Cox model estimation is semi-parametric such that the functional form of the baseline hazards is not given. The baseline hazard rate is calculated for all variables having zero value. As a result we normalized the variables using a linear transformation to obtain zero means. Figure 2 shows the baseline hazard rate for an average player.

Although one player remains 22 years in the NBA, we do not display the hazard rate after the fourteenth year because there are less than ten hazard events per year and results are neither representative nor stable. From Figure 2 it is obvious that the hazard rate does not follow any specific functional form and must be estimated non-parametrically. This justifies the application of the Cox model. One common alternative would have been to estimate a Weibull model, but the Weibull function has to be monotone and that is not the case for the hazard rate in the NBA.

Figure 3 shows the corresponding survival function to Figure 2. It can be seen that 50 % of the players are gone after five years even when achieving average characteristics. Most rookie players perform below average statistics, which reduces their probability of staying in the league.

The results of estimating the corresponding Cox model for team-hazard are summarized in Table 3. Tests of the proportional hazard assumption illustrate the goodness of the overall model and for the specific variables included in the estimation. Scoring performance significantly hinders a job change. The risk of being traded is 37 % lower for one point per minute more. Amazingly, the non-scoring

Figure 2: Baseline League-Hazard Rate

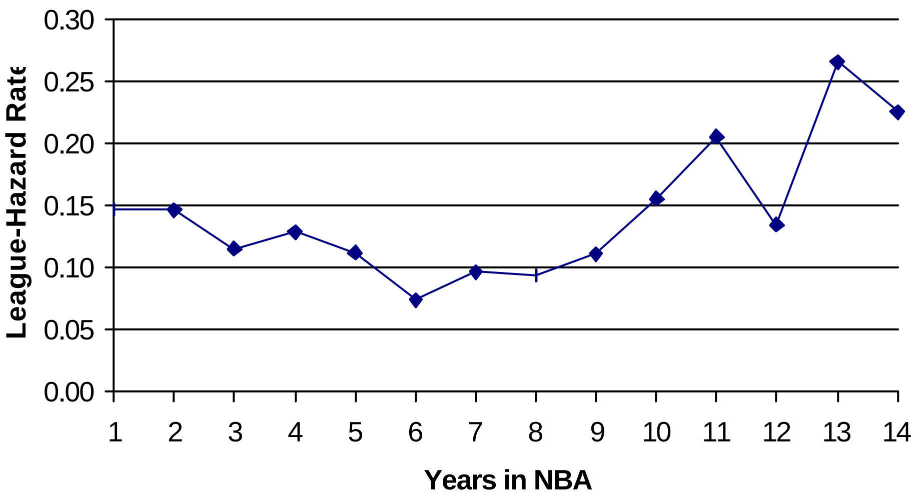
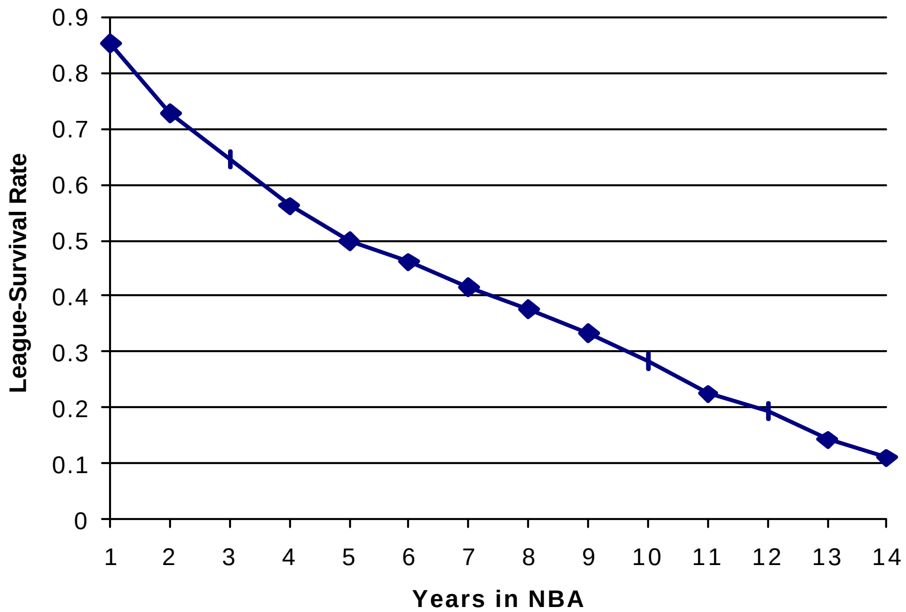


Figure 3: Baseline League-Survival Rate



performance is insignificantly positively sloped. This might be interpreted as follows: There is no need to exchange players of identical skills i.e. replacing one top-scorer for another one. Points are points. However, things are differently regarding the non-scoring performance measures. One team lacks a playmaker who provides the "big boys" with assists in the paint, another club might need to reinforce it's defence and employs a strong rebounder. Teams do not like players producing one foul after the

Table 3: Team-Change as Hazard

Variable	Coefficient	Std. Dev.
Scoring performance	-0.4692518**	0.1686467
Non-scoring performance	0.0065880	0.2140010
Personal fouls	0.2858753*	0.1139573
League years	0.1185835***	0.0233878
League years squared	-0.0047571**	0.0015905
Games	-0.1129830***	0.0010448
Win-percentage	-0.6563446***	0.1527685
Draft number	0.0025397^{+}	0.0014704
Salary (ln)	0.0009563	0.0230358
Roster	-0.0051531	0.0173600
Colour	0.0151138	0.0573018
Height	-0.0036314	0.0026583
All star games	-0.1960048*	0.0786295

Number of subjects = 2,039

Number of hazards = 1,175

Number of observations = time at risk = 3,945

c^2 = 309.67***

McFadden R^2 = 0.0119855

$^{+}$ p<0.1; * p<0.05; ** p<0.01; *** p<0.001, two-tailed tests other; thus they hand them over (with an increased probability of 33 % for an additional foul per minute) as long as they do not have to leave the entire league.

The most statistically significant variable is again the number of games played during the season. This confirms hypothesis 2. A team is not interested in loosing players that it needs most, so it neither lays them off the league, nor trades them to another NBA franchise. The most intensively used players achieve their highest value with the club they are currently working. In every other case the team is well advised to trade the particular player and look out to reach more adequate matches.

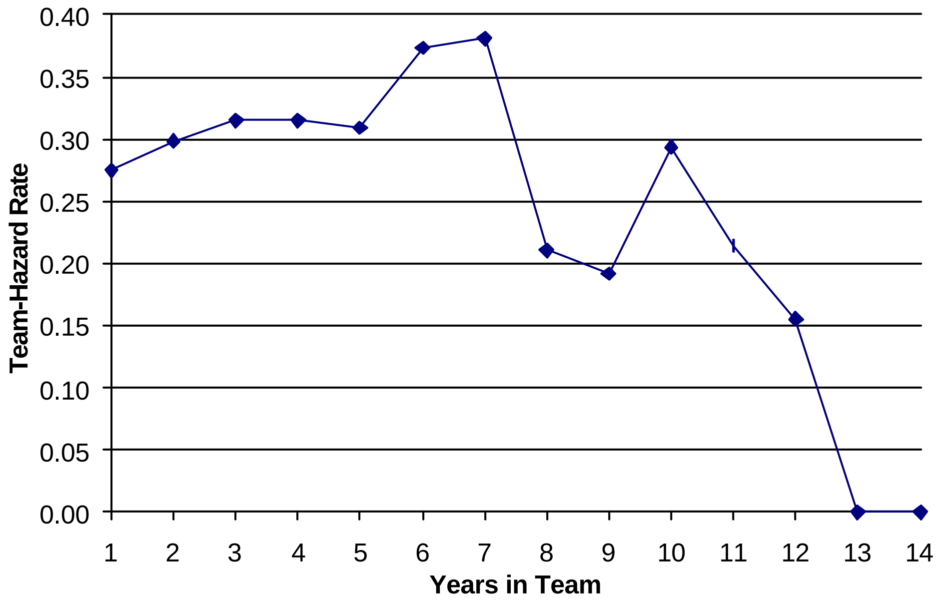
Winning significantly reduces the propensity of a player to separate with his original team: Never change a winning team. It remains open for future research whether a high win-percentage really diminishes the probability of a team-change or whether this causal relationship is rather inverse, such that stable teams win more regular games than teams producing a high degree of labor turnover.

The draft number is positively sloped. A low draft pick hints high quality, although this information is widely available for all teams. The salary coefficient has also a statistically insignificant positive direction. Although one might expect a significantly positive effect this observation does not contradict hypothesis 2, because salaries are not only costs but can be seen as a sign of high performance. Both effects may cancel out each other.

Team roster, colour and size show again no significant effect. Finally, taking part on at least one all star game significantly lowers the probability of a team-change by 18 %. All star games are not only a pure quality signal, they also reveal additional, partly team-specific value to the spectators beyond athletic performance.

Figure 4 depicts the baseline hazard rate for the team change event. Although there are less than ten hazards in the eighth and some later years, we plotted the hazard rate until year fourteen for one team to make a comparison with Figure 2 feasible. As can be seen in Table 1 the maximum achieved years within a particular club is 16, reached by exactly two players in our data set. During the first

Figure 4: Baseline Team-Hazard Rate

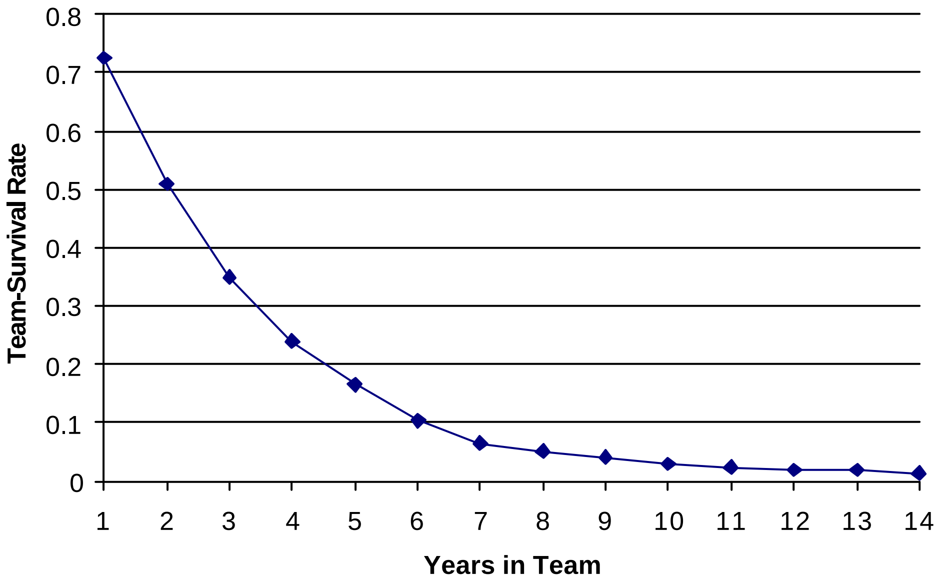


years the probability to switch the team is even higher than the hazard to leave the league. The hazard to change the team finds its maximum in the seventh year at nearly 38 % for a player with average characteristics. Thereafter, it plunges and peaks again in the tenth year. In the thirteenth and fourteenth year the propensity to change the team has fallen to zero. In other words, this means that new team members change with a high probability, whereas old ones have found their perfect match or have lost their will to change teams before retirement. Figure 5 shows the corresponding baseline survival rate.

5. Conclusions

This paper has tested Cox proportional hazard models for leaving the NBA and for changing a team inside the NBA. The results obtained by the regressions are strongly consistent with our main hypotheses. The most important reason for dropping the league is low player performance. Thereby it is not only the direct performance measures that matter. Also the number of games played or the salary do count as indirect measures of productivity.

Figure 5: Baseline Team-Survival Rate



Low performance cannot be the single determinant for a team change, because no club would like to receive such players. This is exactly why these kind of players have to exit the league. A team switch occurs if both teams improve their situation, that is the changed player has a comparative higher value in the new team than in his old franchise. Therefore, the extensive use of a player by his team – measured by the number of games played – inhibits the probability of being traded similar as the high scoring performance does. However, this does not hold for a high non-scoring performance or large salaries. The former has many dimensions so that players with different strengths and weaknesses can be traded, whereas the latter are an indirect performance measure and expenditure at the same time. It remains a question for future research whether the matching hypothesis is true and better matches between teams and players are really reached, or whether the quality of players is general to the league and does not vary between teams before and after trading.

References

Altman, D. G. (1991). *Practical Statistics for Medical Research*. London: Chapman & Hall.

Booth, A., Francesconi, M., & Garcia-Serrano, C. (1999). Job Tenure and Mobility in Britain. *Industrial and Labor Relations Review*, *53*, 43-70.

Borland, J., & Lye, J. (1996). Matching and Mobility in the Market For Australian Rules Football Coaches. *Industrial and Labor Relations Review, 50*, 143-158.

Breslow, N. E. (1974). Covariance Analysis of Censored Survival Data. *Biometrics*, *30*, 89-99.

Campell, C. (1993). Do Firms Pay Efficiency Wages? Evidence with Data at the Firm Level. *Journal of Labor Economics*, *11*, 442-470.

Chapman, K. S., & Southwick, L. (1991). Testing the Matching Hypothesis: The Case of Major League Baseball. *American Economic Review*, *81*, 1352-1360.

Cox, D. R. (1972). Regression Models and Life-Tabels (with Discussion). *Journal of the Royal Statistical Society, Series B*, *34*, 187-220.

Harder, J. (1992). Play for Pay: Effects of Inequity in a Pay-for Performance Context. *Administrative Science Quarterly*, *37*, 321-335.

Hoang, H., & Rascher, D. (1999). The NBA, Exit Discrimination, and Career Earnings. *Industrial Relations*, *38*, 69-91.

Idson, T., & Kahane, L. (1999). Team Effects on Compensation: An Application to Salary Determination in the National Hockey League. *Economic Inquiry*, *37*, 345-357.

Jones, J. C., & Walsh, W. D. (1988). Salary Determination in the National Hockey League: The Effects of Skills, Franchises Characteristics and Discrimination. *Industrial and Labor Relations Review*, *42*, 592-604.

Jovanovic, B. (1979). Job Matching and the Theory of Turnover. *Journal of Political Economy*, *87*, 972-990.

Kahn, L. M. (2000). The Sports Business as a Labor Market Laboratory. *Journal of Economic Perspectives*, *14*, *No. 3*, 75-94.

Kahn, H. A., & Sempos, C. T. (1989). *Statistical Methods in Epidemiology*. New York: Oxford University Press.

Kalbfleisch, J. D., & Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data.* New York: John Wiley & Sons.

Kiefer, N. M. (1988). Economic Duration Data and Hazard Functions. *Journal of Economic Literature*, *26*, 646-679.

Klein, J. P., & Moeschberger, M. L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data.* New York: Springer-Verlag.

Lazear, E. (1999). Personnel Economics: Past Lessons and Future Directions; Presidential Address to the Society of Economists, San Francisco. *Journal of Labor Economics*, *17*, 199-237.

Lin, D. Y., & Wei, L. J. (1989). The robust Inference for the Cox Proportional Hazards Model. *Journal of the American Statistical Association*, *84*, 1074-1078.

Ohkusa, Y. (2001). An Empirical Examination of the Quit Behavior of Professional Baseball Players in Japan. *Journal of Sports Economics*, *2*, 80-88.

Ohkusa, Y., & Ohtake, F. (1996). The Relationship between Supervisors and Workers. The Case of Professional Baseball in Japan. *Japan and the World Economy*, *8*, 475-488.

Ohtake, F., & Ohkusa, Y. (1994). Testing the Matching Hypothesis: The Case of Professional Baseball in Japan with Comparison to the US. *Journal of the Japanese and International Economies*, *8*, 403-411.

Peto, R. (1972). Contribution to the Discussion of Paper by D. R. Cox. *Journal of the Royal Statistical Society*, *Series B*, *34*, 205-207.

Rosen, S., & Sanderson, A. (2000). Labor Markets in Professional Sports. *NBER Working Paper 7573*. Cambridge (MA): National Bureau of Economic Research.

Scully, G. W. (1974). Pay and Performance in Major League Baseball. *American Economic Review*, *64*, 915-930.

Selvin, S. (1996). *Statistical Analysis of Epidemiologic Data* (2nd ed.). New York: Oxford University Press.

Staw, B. M., & Hoang, H. (1995). Sunk Costs in the NBA: Why Draft Order Affects Playing Time and Survival in Professional Basketball. *Administrative Science Quarterly*, *40*, 474-494.

Ernst-Moritz-Arndt-Universität Greifswald

Rechts- und Staatswissenschaftliche Fakultät

Wirtschaftswissenschaftliche Diskussionspapiere

2002 bislang erschienen:

1/02 Stefan Greß; Kieke G. H. Okma; Jürgen Wasem: “Private Health Insurance in Social Health Insurance Countries: Market Outcomes and Policy Implications”

2/02 Ole Janssen; Armin Rohde: Monetäre Ursachen der Arbeitslosigkeit in Currency Board-Systemen?”

Siehe auch unter: http://www.rsf.uni-greifswald.de/bwl/paper.html