PEER EFFECTS IN SPORTS:

EVIDENCE FROM NCAA

RELAY TEAMS

by

LISA ELIN HAGLUND

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF ARTS IN ECONOMICS

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2007

ACKNOWLEDGEMENTS

I would like to thank Dr. Craig A. Depken II for his continuous help and support in writing my thesis. I also want to thank Dr. Courtney LaFountain and Dr. Michael R. Ward for being part of my committee.

April 18, 2007

ABSTRACT

PEER EFFECTS IN SPORTS:

EVIDENCE FROM NCAA

RELAY TEAMS

Publication No. ______

Lisa Elin Haglund, M.A.

The University of Texas at Arlington, 2007

Supervising Professor: Dr. Craig A. Depken II

This paper investigates whether disparity in team member quality impacts team production using NCAA 4x400m relay teams. As a measure of quality I use the team member's individual rankings. The net peer effects are estimated on a team level rather than on an individual level, and are found to have both an absolute and relative negative effect on the team performance. This paper is differentiated from the existing literature by using a direct measure of quality compared to other indirect measures of worker quality such as wages. The evidence provided herein shows that a greater disparity in team member quality increases NCAA relay team times, which suggests that net

negative peer effects exist. These results support the “team cohesiveness hypothesis” for NCAA relay teams.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER 1

INTRODUCTION

In an individual sport such as track and field, athletes are often primarily concerned with their own events and performance. Therefore, it can sometimes be hard to make them compete as a team, since this might put them at a disadvantage in their individual competition. In Europe, where track and field athletes compete for a club rather than a school, athletes compete on their own terms without having to answer to anyone other than their coach. In the United States, on the other hand, most track and field participants compete for a school team, and sometimes sacrifice their personal preference and performance for the "greater good," that is, the team's overall performance. For example, an athlete might be better at the 800m than the 1500m and would prefer to run the 800m if he could decide himself. However, if the team has other athletes who are better than him in the 800m but no one in the 1500m, the coach might decide to put the athlete in the 1500m to give the team an opportunity to score in a wider range of events. On larger teams, an individual team member might see a reduction in his own value since his performance is just one out of a large group of athletes.

One particular instance where this can be seen is on relay teams. Although it might be easier to be part of a relay team as the output is a combination of several team members, it still might not always be what the individual member wants to do or prioritize. Therefore, an individual might not do the best he can, i.e., shirk or free-ride on the other members of the relay team. However, these are only some of the negative peer effects that can arise when members work together as a team. Other effects can include mistrust and jealousy.

This paper deals with the problem of peer effects in 4x400m relay teams in National Collegiate Athletic Association (NCAA) Division I track and field. The empirical model suggests that negative peer effects are a greater problem on teams with more disparity in member quality. To measure team member quality, I use the runner's individual ranking, where a higher ranking indicates a faster runner, that is, a runner of higher quality. Therefore, a team with more disparity in member quality is a team with more uneven runners.

Research on negative peer effects such as shirking is based on the assumption that team members have a harder time monitoring member input when disparity is low, that is, when the team members are equal in abilities, and the coach has a harder time monitoring member input when disparity is low. When it is harder to monitor behavior it is less costly for an individual to shirk. When a team is composed of four runners of equal capacities, it will be harder to notice which of the team members did not perform to their fullest ability and whether the reduced performance was the result of shirking or some other random influence. However, if there is one athlete on the team who is much better than the other members, it will be obvious if he decides to shirk, since the final time will depend much more on how fast he runs. However, negative peer effects can arise from several different sources beyond shirking.

To estimate whether relay teams suffer from negative peer effects, I use a model were a team's relay time is a function of the average ranking of its members, the standard deviation of its members' ranking, the type of season, the type of race, the number of different runners on the team from the previous race, and the year of the race. The general results show that as the average ranking increases, time decreases, that is, the relay team runs faster, but as the standard deviation of the team members' ranking increases, the time increases. The evidence also shows that relay teams perform better outdoors than indoors, and the more important the race is the faster the team will run. Finally, teams with less turnover in team members tend to run faster, ceteris paribus.

Within the NCAA, athletes are considered amateurs and are not allowed to earn money on their sporting performance. Instead, schools recruit athletes based on their academic and athletic abilities and compensate them with various levels of scholarship support. That is, instead of monetary wages, the athletes are paid in-kind. All schools are restricted in what they can offer the athlete, and schools can only offer a certain number of scholarships, divided evenly between male and female sports. The total number of scholarships that a school is allowed to offer a men's track and field team is 12.6 full scholarships. The coach can divide the scholarship money as he or she pleases, but no athlete can receive more than a full scholarship. To ensure that all schools have the same basic chance to recruit an athlete, the Ivy League schools[1] have decided not to give out any athletic scholarships (Kotlyarenko and Ehrenberg 2000, 140). With their academic and financial opportunities, it would not be fair for the other Division I schools to compete with them for an athlete. Schools that violate NCAA rules can be punished by the NCAA Committee of Infractions, as happened to the men's track and field team at Texas Christian University (TCU) in 2005. After academic and financial violations by the head coach and several assistant coaches were discovered, the TCU team received a set of punishments "including scholarship and recruiting limitations and a ban on postseason competition by the track team through 2006-7" (Lederman 2005). The TCU team was also put on two-year probation and all coaches involved were fired.

When an athlete is part of a college team, the coach can choose to increase or decrease the scholarship for the following year if the athlete performs above or below his expectations during a year, but other than the scholarship the coach does not have much of a "hold" on the athlete. Since no wages are involved in NCAA sports, the coach has to provide other motivational factors to ensure that the athlete performs to the best of his ability. Motivational factors can include promises to go to bigger or better meets, to let the athlete run events of his preference, or to expose the athlete more to the campus/community media. Other personal factors that can help an athlete's performance are a strong school or team spirit, and personal pride. On the other hand, the coach can impose threats to get an athlete to work harder. Examples include having the athlete practice more (e.g. more morning runs), not letting them run certain meets, cutting them from the relay team, or cutting them from the team altogether.

This paper is organized in the following matter. First there is a review of the existing literature concerning peer effects and teams. In the next chapter, I will explain

---

[1] The Ivy League Conference is composed of the following teams: Brown, Columbia, Cornell,

the empirical application. The following section provides an explanation of the hypotheses and what the data measure, along with the data description and sources of data. Chapter 5 provides the empirical results. The last chapter provides conclusions.

Dartmouth, Harvard, Pennsylvania, Princeton, and Yale.

CHAPTER 2

LITERATURE REVIEW

The difference between a team and individual production function is that a team's production function is more than just the "summing up" of the individual team members' production functions. Instead, teams produce a combined output, where "there is a source of gain from cooperative activity involving working as a team, wherein individual cooperating inputs do not yield identifiable, separate products which can be summed to measure the total output" (Alchian and Demsetz 1972, 779). Although this in many cases means that the production output from a team will be greater than it would from the same number of individuals, a team also has to deal with internal issues that an individual does not. These issues can affect the performance of workers dramatically, and can have positive or negative effects on the output of the team. Positive peer effects include knowledge transfers, increased esprit de corps, and improved marginal products. Negative peer effects, which exist if the output of worker *i* systematically decreases the output of worker *j* and vice versa, include problems such as jealousy, free riding, and shirking (Falk and Ichino 2006, 41).

Much of the existing literature in this area is based on problems of moral hazard. Whenever one is dealing with a situation where there is more than one worker, there is often a concern about moral hazard. According to Bengt Holmström; moral hazard "refers to the problem of inducing agents to supply proper amounts of productive inputs when their actions cannot be observed and contracted for directly" (Holmström 1982, 324). Any uncooperative behavior "will always yield an inefficient outcome if joint output is fully shared among the agents" (Holmström 1982, 325). Therefore, by separating ownership and labor, he suggests that one can resolve the problem of free-riding, because then you can see what the individual member contributes to the total output.

To accomplish this, much research supports the idea of monitoring the workers. As suggested by Grossman and Hart (1983), the incentive to shirk, for example, will increase as the information available to the principal decreases based on what he can monitor (35). However, monitoring can be very costly or even infeasible. An alternative solution is to have the employees monitor each other. Barron and Gjerde (1997) coin the term "peer pressure environment" as a combination of employee-provided standards, monitoring, and sanctions accepted by the workers before they decide on their individual levels of work effort. However, since a worker does not earn any higher payment by discovering another worker's lazy behavior, he will only exert effort to where his "marginal gain of detection activity" equals "the marginal cost of detection" (Alchian and Demsetz 1972, 780), which is said to imply "a lower rate of productive effort and more shirking than in a costless monitoring, or measuring, world" (780). Also, even if a team member knows that another team member is not doing as much as he could, it might not be in his best interest to tell the manager about this behavior even though it decreases the team's performance. If the person shirking is better than another team member would be even if he did not shirk, then it might be worth that frustration or feeling of unfairness to keep the member on the team despite his behavior.

Team externalities exist when individual performance cannot be completely measured. In situations where team externalities exist a “complete sharing of blame or credit among individuals” arise, and thereby also the ability or incentive to shirk or free-ride (Drago and Turnbull 1988, 101). A lower effort by one individual might decrease the effort of another individual, causing them to share the blame for the decrease in production. As Drago and Turnbull (1988) put it; “under a tournament, incentives to free-ride increase with the externalities as individuals will correctly perceive that shirking is less likely to result in loss of the tournament” (101). Therefore, as long as there are some externalities, one cannot achieve efficiency (Holmström 1982, 325).

Much of the research done in this subject is based on how negative peer effects or moral hazard can be decreased by using the right compensation scheme or incentives for workers. According to Nalebuff and Stiglitz (1983); “workers will supply effort until their marginal disutility from work is just balanced by their increased chance of winning the value of the prize” (26). So if workers do not find their compensation fair, or if they feel that an increase in compensation is too hard to achieve, they will find it more tempting to engage in moral hazard. This is also true for individuals who are to work as a team. Within the context of an NCAA relay team, for example, if a runner knows that he is given a much lower scholarship than someone else on the team who has the same capacity as himself, the runner might decide that it is not worth the effort to push himself to the best of his ability.

Problems can also occur if the objectives of individual workers are not the same as the objectives of the team, that is, there might be a "conflict of interest" (Groves 1973, 618). According to Groves (1973), team decision making is "a multi-person joint decision problem in which the decision makers base their decision choices on different information, yet are motivated by a common goal" (618). In the current context, relay teams can be compared to a conglomerate, which is "an organization consisting of many partially autonomous units linked only through a central administration" (Groves 1973, 622), by looking at the team members as separate subunits and the coach or school as the central administration. In this model, the team members are linked together only through the coach, and might therefore not agree with all the decisions made for them by the coach. For example, their opinion about maximum performance or racing decisions might not be the same as that of the coach. Therefore, Grove suggests that it might be in the team members' interest to lie and instead "send false information" (Grove 1973, 624) to the central administration (or the coach). That is, if the team members do not agree with the coach's decisions, they might choose to give the coach imperfect information, and thereby make it easier for themselves to shirk or free-ride. However, this is likely less of a problem at schools with great team spirit or a strong history in a certain event, such as the 4x400m relay, where team members have a common goal, and the athletes who go there do so because they want to be part of that particular team.

Along with the problem of incentives is the concern of how the workers should be compensated. Lazear and Rosen (1981) suggest that when "inexpensive and reliable monitors of effort are available, then the best compensation scheme is a periodic wage based on input" (842). But, on the other hand, "when monitoring is difficult, so that workers can alter their input with less than perfect detection, input-wage schemes invite shirking" (842). Therefore, when monitoring becomes a problem, they instead suggest that one should pay workers based on rank order, that is, they should get paid depending on their relative position in the organization, and not on the relative production output. This second alternative is usually how track athletes within NCAA are compensated, that is, based on their relative rank or performance within the team as a whole. Anyone with a full scholarship should generally be one of the athletes who scores the most points on the team and can qualify for the championships.

According to Nalebuff and Stiglitz (1983), there are "three critical characteristics of any reward structure," namely risk, incentive levels, and flexibility (22). These characteristics should always be considered, as they are vital to receive the wanted performance from the worker. If a worker is risk averse, shirking when compensation is based solely on output might hurt the worker, since there will be a significant loss in welfare for that person. On the other hand, when compensation is not directly related to output, a worker might not produce the right level of output or make the right decisions (Nalebuff and Stiglitz 1983, 22). This is supported by Grossman and Hart (1983) as they argue that incentive problems only exist when a worker is risk averse and not when the worker is risk neutral (38). What also should be noticed is that what might be an incentive for one worker might not be an incentive for another worker, or even for the same worker in another situation.

Compensation can also be paid in piece rates versus contests. Piece rates are given when output is easily observed, whereas contests are preferable when output is harder to observe (Lazear and Rosen 1981). A different opinion is expressed by McAfee and McMillian (1991) as they propose that “the principal can do no better when he monitors than when he simply bases payments on team output” (563). According to the authors, group-payment schemes are just as effective as piece rate schemes; “the jointness of production and the unobservability of individuals’ contributions need not create a free-rider problem” (McAfee and McMillian 1991, 563). Although, in the end, Lazear and Rosen argue that as long as the worker is risk neutral, independent of whether one pays the workers in piece rates, fixed standard or contest, one can always avoid moral hazard “by shifting all risk onto the agents” (Green and Stockey 1983, 350).

One compensation scheme with great flexibility is the competitive compensation scheme (Nalebuff and Stiglitz 1983, 22) or contest scheme. In this type of compensation, there has to be someone who will be the “loser”, whereas in other compensation schemes everybody might be rewarded (Nalebuff and Stiglitz 1983, 40). However, this works best when all the workers are similar. If everyone knows there are one or more individuals who will be losers, the incentive to work hard disappears. One way to solve this problem is by introducing handicaps, as for example in golf, which will bring back the competitive environment into the competition (Nalebuff and Stiglitz 1983, 40).

Yet another way to organize compensation is to pay team members under a profit-sharing scheme, as suggested by Alchian and Demsetz (1972). With profit sharing, output can increase as the incentives for the workers to shirk would decrease (786). However, they do suggest that this is better for smaller sized teams, as the incentives to shirk are positively related to the size of the team. In the context of an NCAA relay team, compensation might take the form of how many other events in which team members are good enough to participate. Therefore, it is important that team members be compensated in a manner consistent with their performances, so that none of them feel treated unfairly. An alternative is to consider the dollar value of the team member's scholarship as the form of compensation.

Information is another important factor when dealing with peer effects or moral hazard problems and workers. If the worker has full information he will never be caught shirking or free-riding (Baiman and Demski 1980, 204), whereas if the principal (or as in this paper, the coach) has perfect information, the workers will not be able to shirk without being caught. The problem with information is not just a concern between the coach and the team, but also between the team members themselves; "each individual member of a team decides about a different action variable, and each member's decision is based, in general, on different information" (Marschak and Radner 1972, 123). There will always be some information concerning the total team action that is available only to the individual member.

When a worker's ability is known only to himself this is termed "adverse selection" (McAfee and McMilliam 1991, 561). This, however, will never be fully true in the case of a relay team. Although there might not be full information, the coach always has some information about the various team members from watching them during workouts and other races.

In a situation with imperfect information for the principal, workers will want to group themselves with relatively better workers, and create an inefficient situation: groups will not produce at the expected level (Lazear and Rosen 1981). McAfee and McMillian add that although a worker will not overstate his ability since this would create an imbalance with his payment function, he does want to "belong to teams whose members have high productivity" (McAfee and McMillian 1991, 568). However, it is also said that if payment would be "less sensitive to output as ability increased, low-ability agents would have an incentive to overstate their abilities" (McAfee and McMillian 1991, 570). A member of a relay team might be able to overstate his ability to a certain extend by telling the coach and the team members that he has been doing better times that he actually has, but eventually his performances during competition and practice will be observed and the athlete will be put in a group that more closely matches his actual ability.

In addition, various types of organizations encourage different levels of peer effects. Alchian and Demsetz (1972) write in their article "Production, Information Costs, and Economic Organization" that the level of shirking or free-riding should be higher in nonprofit organizations and mutually owned enterprises, as any wealth is not remitted to the stockholders despite any improvements in the management of the firm (789). They compare nonprofit and mutually owned enterprises, among others, to a partnership, which are smaller to avoid shirking or free-riding, and are usually created among people who know each other well and thereby know each other's "work characteristics and tendencies to shirk" (790). As a relay team continues to practice and compete together, the tendency to shirk or free-ride should, therefore, decrease among the members as they get to know each other better. However, an athlete can only compete for NCAA Division I for four years, and therefore it is possible that many teams have more turnover than is optimal because of the NCAA rules.

Another topic discussed in the context of negative peer effects and moral hazard is the allocation of workers. Sanderson and Siegfried (2003) suggest that "where spillovers are significant, high-ability workers are more valuable to other high-quality workers, and workers should be more homogeneously sorted" (266). The only problem in this situation is how to sort these common groups so that output is maximized (266).

The confidence in other team members is based on how much trust there is between the members. Trust can help a team by increasing the information available to the principal and the other team members (Jones and George 1998, 533), which in turn leads to better cooperation and improved performance. If a team can create unconditional trust, the team members will have a "sense of mutual identification" (537), and will avoid problems with conflicts of interest. Teams with strong team spirit have a strong advantage in these circumstances as "shared values result in strong desire to cooperate, even at personal expense" (Jones and George 1998, 539), which the authors suggest will decrease the risks of negative peer effects. On the other side, the principal can also hurt the ability of the team members to cooperate if that type of behavior is not supported or reinforced within the organization. However, it is not clear how long it takes to build this level of trust, and the rules of NCAA eligibility may hurt some relay teams.

Sports teams, as well as many other organizations, try to encourage a strong sense of team spirit and loyalty. As proposed by Alchian and Demsetz, all workers want to be part of a team where no one shirks or free-rides, and by creating stronger team feelings, the team can "enhance a common interest in nonshirking" (790) and become more efficient. The increase in efficiency is a result of the reduction of shirking and not to any other factor built through team spirit or loyalty. They even take it as far as saying that "there will be a tendency to overinvest in training athletes and developing teams" which will create something called "reverse shirking" (Alchian and Demsetz 1972, 791). Many teams and athletes have this problem if they put too much time and effort into their sport at the expense of other activities that have their own benefits. Not all athletes can become superstars; unfortunately, hard training is not enough. Therefore, it might not help if a coach "overinvests" in an athlete, or that the athlete "overinvests" in a sport, because it might be a waste of effort. Hence, both shirking and reverse shirking might be problems for a team.

Even though athletes under NCAA Division I are considered amateur athletes and are not allowed to receive monetary wages for their performance, NCAA sports teams act much like the teams and organizations discussed above. Alchian and Demsetz (1972) explain team production as a production where "several types of resources are used" and "the product is not a sum of separable outputs of each cooperating resource," and lastly that "not all resources used in team production belong to one person" (779). Relay teams have many of these characteristics: the four runners all bring their own resources to the production of the team, several types of resources are used, and the resources belong to different people.

In this paper I am considering the outcome of a team effort rather than the performance of individual track athletes, and therefore cannot fully differentiate the individual efforts. Instead, I will therefore look at the aggregate peer effects across the entire team. As discussed by Carmichael, Thomas and Ward (2000); "[i]n many sports, the isolation and identification of individual contributions is problematic, if not impossible, due to the continuously interactive nature of a match or game" (32). In the case of relay races, it might be possible to use the individual splits of the runners. However, many other factors and tactics play a role in the separate times and the coach might not be able to say exactly who did what as it relates to the final outcome. For example, a runner might be pushed during the baton handoff and therefore have a much slower split than what he is capable of.

Since peer effects can be hard to measure directly, one has to find another common factor that can work as a substitute. Depken (2000), and several authors following his example, have investigated the quality of professional sports teams as a function of average salaries and the disparity of salaries, that is, they use salaries as a proxy for the quality of the player. Depken found that a higher disparity in wages on professional baseball teams led to a reduction in team performance. This result supports the "team cohesiveness hypothesis," developed by Levine (1991), which posits that "greater wage disparity motivates jealousy and mistrust among workers in a firm (team) and possible reduction in overall team performance" (Depken 2000, 5).

The main difference for this research is that there are no wages in the NCAA; athletes are instead given an in-kind payment. Therefore, my approach is to take a direct measure of quality, rather than wages, to test the team cohesiveness hypothesis. That is, I will test if a higher disparity in worker quality leads to a reduction in team performance due to negative peer effects.

CHAPTER 3

A MODEL OF A RELAY TEAM AND PEER EFFECTS

Consider the actual performance of a relay team as a function of the actual performance of the four runners, where actual performance of the runners is a function of the particular quality of that runner, that is, the team has a production function defined as $Time = f(A_1(Q_1), A_2(Q_2), A_3(Q_3), A_4(Q_4)) + \varepsilon$, where A_i is the team member *i* and Q_i represents team member *i*'s expected quality, and ε is a random error term. This team production function would suggest each player's effort is independent of the other players abilities.

However, this would not be an accurate functional form for the type of relay teams considered in this paper, since it does not account for any possible peer effects. A runner's performance in a relay race depends not only on their own quality, but also of the quality of the other team members, which suggests that the basic functional form is naïve. It does not take into account the influences that one runner's quality has on another runner's actual performance or any effects that comes out of competing as a team. Therefore, a more appropriate production function for a team member would instead be a function where $A_i = f(Q_1, Q_2, Q_3, Q_4)$. The partial effect of Q_j on A_i reflects the peer effects of runner *j* on runner *i*. This combination of peer effects on runner *i* will contribute to the actual production for runner *i*. The total peer effects experienced by a team member will be ambiguous; the net effect can be positive or negative. A net positive peer effect will induce the team member to run faster, whereas negative peer effects would induce the runner to run slower. In this study, I will only be able to estimate a reduced form of the net impact of peer effects on a team level. That is, I cannot identify how any particular runner influences the other runners on the team.

A coach can try to improve the performance of a team by changing the combination of the team members, and replacing one runner with a runner of higher quality. In this case, the coach is hoping to get a reduction in the team's time by changing the jth runner, that is, $\frac{\partial A_j}{\partial Q_j} > 0$. However, this will only be the case if, in fact, the total output of the new team member is positive, $\frac{\partial T}{\partial Q_j} = \sum_{i=1}^{4} \frac{\partial f}{\partial A_i} * \frac{\partial A_i}{\partial Q_j}$, where $\frac{\partial A_i}{\partial Q_j}$ reflects positive or negative peer effects of team member *j* on team member *i*. The actual impact of changing a runner can be positive or negative depending on the peer effects caused by the change in quality of the jth team member.

If an increase in quality improves the aggregate performance of the team members, this would mean that the team members responded positively to the new combination of runners, and the time of the relay would decrease. Here, the gains from improved quality would be greater than any negative effects experienced by any of the other runners. On the other hand, if the actual performance of the team decreases after the change of team members, despite any increase in quality, then there are negative peer effects that reduce team output.

To test for positive or negative peer effects, I follow the approach used by Depken (2000), by controlling for average team member quality and the disparity of quality using the standard deviation of team member quality. By including both the average and the standard deviation, the impact of team member quality disparities, which can lead to positive or negative peer effects, can be investigated. If the standard deviation is insignificant, there would be no evidence of net peer effects caused by disparity of team member quality. If the standard deviation is significant and positive, this means that there exist net negative peer effects on the team, and that the greater the disparity of talent on a relay team, the worse the team performs, ceteris paribus.

If negative peer effects exist, why doesn't the coach fix the problem? The coach may not be able to remedy this problem because the coach does not have an unrestricted ability to alter the composition of the team to remove all of the aggregated peer effects. One reason is the limitations imposed by the NCAA rules. Since the runners cannot receive any wages, a coach can not affect a runner's desire to perform better or to disregard the other team members' qualities. Other reasons could be that the coach is unaware of the negative peer effects on the team, that the coach has a different objective function, or that he or she has limited coaching experience or ability.

To test for possible negative peer effects, three different approaches are taken. First, a basic OLS regression with team time as the dependent variable measures the effect of the disparity in team member quality on team performance in an absolute sense. Second, a measure of predicted team quality is generated and compared to the actual team's production. The difference between how the team is predicted to perform and how they actually perform is created and used as the dependent variable. This model tests whether peer effects reduce team effectiveness in a relative rather than an absolute sense, and also helps determine whether a team is under-performing or over-performing given its inputs. Finally, a probit model is estimated where the dependent variable is a dummy variable that takes a value of one if the team is under-performing. In this regression, the parameter estimates reflect the marginal contribution of the right hand side variables to the probability that the team will under-perform given the team members.

If a team performs worse than predicted, on the margin, the parameter estimate will be positive, and if a team performs better than expected, on the margin, they will have a negative parameter estimate. A positive parameter estimate on the standard deviation of team member quality suggests negative net peer effects caused by disparity in team member quality.

3.1 Data and Empirical Specification

My hypothesis states that the greater the disparity of quality on a relay team, the greater the problem of negative peer effects will be on the team. The standard deviation of team member quality is used to test for net aggregate peer effects in the sense that a team with more disparity in team member quality may be faster (slower), ceteris paribus, if the disparity induces positive (negative) peer effects.

In this paper, I am investigating possible peer effects at the team level rather than the individual level. The main question is whether team performance is greater or less than the sum of the individual parts, that is, the expected performance of the individual runners. If the team under-performs, this could occur for one of two reasons: either a random shock or a systematic shock to the relay team. For example, a random shock might be dropping the baton or that one of the team members gets injured. An example of a systematic shock would be that team performance is influenced by some kind of peer effects. Negative peer effects could arise from free-riding or jealousy. Positive peer effects might arise from increased morale and team spirit or reverse-shirking by one or more team members. If it was the case that under-performance was only due to random shocks, then the variables included in the regression analysis will be insignificant.

The team production function that I use is of the form *TIME* = f (inputs) + e, where *TIME* is the dependent variable, the inputs include the average ranking of the team members, the standard deviation of team member ranking, the season (indoors or outdoors) during which the race was run, the type of race (regional, prelims or final), the number of different runners on the team from the previous race, and year dummy variables, and e is a zero-mean random error term. From this, the estimated model is:

$$
\begin{aligned}
TIME &= \beta_0 + \beta_1 AVERANK + \beta_2 AVERANK^2 + \beta_3 SDRANK \\
&+ \beta_4 SDRANK^2 + \beta_5 OUTDOORS + \beta_6 FINALS \\
&+ \beta_7 PRELIMS + \beta_8 NUMDIFF + \gamma \cdot YEAR + \varepsilon
\end{aligned}
$$

The dependent variable *TIME* is the time a relay team runs the 4x400m in seconds. The variable *AVERANK* is the average ranking of a particular team's members, and *SDRANK* is the team member's ranking. Whereas $AVERANK^2$ and $SDRANK^2$ are the team's average ranking squared and the standard deviation of that team's ranking squared respectively. The variable *NUMDIFF* is the number of different runners on a team compared to the previous race. The other variables are dummy variables to control for the season, type of race, and year. A race can be run either indoors or outdoors, and in this dataset, the type of race can be a regional, prelim or final race.

The variable *AVERANK* is expected to have an inverse relationship with time, that is, a one unit increase in the average ranking of the team will result in a faster time. The team is expected to run faster since the average ranking of the team members is better. The parameter *SDRANK*, on the other hand, could be either positive or negative since the range of the ranking on the team can result from any combination of runners. The team could include three average runners and one all-star, or it could be three all-star and one average runner. However, when comparing two teams with the same average ranking, the sign of the standard deviation of the team member rankings provide evidence as to whether the average team has a problem of negative peer effects or not, ceteris paribus.

A negative sign on the standard deviation in this situation would mean that as the range of the runner's ranking increases with one unit, that is, the spread of the runner's ranking is larger, the time of the relay will decrease. The purpose for the parameter $AVERANK^2$ and $SDRANK^2$ are to allow for curvature in the model and to capture any diminishing returns in the average ranking of the team or the standard deviation of the team members' ranking.

The parameter *NUMDIFF* is expected to have a positive relationship with the dependent variable. The more runners that have been part of a particular team over the year, the slower the time will be as the team members will have less confidence and trust in each other. The other variables are dummy variables, and are controlling for the season and the type of race of a particular observation. All these variables are expected to have an inverse relationship with time and compared to their base level.

3.2 Sources of Data

The data come from NCAA Division I track and field meets, specifically NCAA regional and NCAA championship meets, and cover teams running the 4x400m relay. The data were collected from year 2002 and through 2006, yielding a total of 1220 individual runner observations, that is, a total of 305 competing teams. To make sure that all team members would be individually ranked, I only used results from the NCAA Regional meets as well as the NCAA Championship meets, which all have qualifying standards. If results from open meets during the ordinary season were included, two problems would be introduced. First, Division II and III as well as unattached teams would be included in the sample. The problem with this is there are no combined rankings for runners within and outside of Division I track and field, and this would make it impossible to compare performances across teams. Second, a lot of runners would not be ranked at all. As the national ranking list from NCAA Division I only included somewhere between 20 and 110 athletes per event and season, there would be no way that a smaller team would have one or more of their athletes individually ranked. For the regional meet and the indoor championship, only finals take place, whereas in the outdoor championship, teams run both prelims and finals. For the entire data sample, results are taken from 24 different races. However, due to the fact that some of these results are several years old, it was not possible to get all of the races of interest. For example, starting in 2003 the NCAA divided the Division I schools into four regions: the East, the Mideast, the Midwest, and the West region. However, in 2006, I was not able to obtain the results from the Mideast region. Therefore, only three of the four regions are represented in my data sample for that year. For 2003, none of the Regional meets are represented because of a lack of information. The same problem occurred for the prelims in the NCAA Championships in 2003 and 2002, and therefore, only the finals are part of the data sample.

The ranking is based on the national rankings generated by Trackshark.com, where each NCAA Division is separated and ranked in each individual event. The ranking lists are conduced yearly, and are separated into an outdoor and an indoor season. The lists include different numbers of athletes depending on the year and season, ranging from 20 to 110 per event per season. Since the observed race is a 400m race, the individuals are only considered if they are ranked in the 200m, 400m, 800m, or 400m hurdles, since these are the most similar events to the event of interest. Some of the individuals in the data sample were ranked in other events such as the high jump or the long jump, but since these events are much different from the 400m these rankings are not comparable to a 400m time. The reason why I did not use the ranking list for the actual relay was that most of the time, the four team members who ran the relay when the time was posted on the ranking list, were not the four members that ran the race included in the sample.

Next, the ranking is converted to a point system from the International Association of Athletics Federation (IAAF), where each time is worth a certain amount of points. The time considered in this case is the time which the individual ran to get on the ranking list. Since the time is a season best rather than a personal best, the individual's time, and thereby its ranking can change from year to year. The point system is separate for indoor and outdoor races, but helps to create a common variable between athletes with different specialty events. The points given for a certain time on an indoor track will be higher than for the same time on an outdoor track. However, since we are only comparing runners in the same season, this difference will not change any results. Also, all times on the ranking lists are converted to a non-banked 200m track at sea-level even if the runner raced on a banked or over-long track or if they were on high altitude. Even if the athlete was ranked in the 400m hurdles, his time can be converted to where it is comparable to either a 400m time or a 200m time. In the end, all ranked runners will have a set of points, and can be compared directly to each other.

The points for each team member are summed for each team, and the mean and standard deviation for each team is calculated. These two variables are used as a measurement of the team quality. The average ranking for the team should have an inverse relationship with the overall team time. The same could be true for the standard deviation of the team as well, but it could go in either direction. Depending on the members of the team, and what rankings they have, the spread could be caused by one faster runner or one considerably slower runner, or one slower and one faster for that matter.

The model also takes into account what kind of competition the relay is run. The time of the relay will differ depending on if it is run on a regional meet or a championship meet, if it is a final or a prelim, as well as if the race is run indoors or outdoors. Overall, a team will run faster the more important the meet is, both because it is more prestige and important to run good at an important meet, but also because the competition from the other teams will be greater in bigger meets. For example, only the very best teams compete in the championship finals. In the regression model, I use the indoor races and the regional meet as the base level to compare to. What is expected here is that running outdoors will have a faster time of the relay compared to indoors, just as the time will be faster for races run in the prelims compared to a regional meet, and final races are faster than the prelims. The reason for the faster time outdoors relative to indoors is usually explained by the fact that you have to run more laps indoors compared to outdoors and that most people feel more comfortable running outdoors than indoors. All these variables are entered as dummy variables, taking a value of 1 if the race is of that particular type or a value of 0 if it is another type of race.

Another variable in the data set is the total number of runners that has competed for a particular team during the years of interest. The total number of different runners on a team is calculated by looking at each race and determining how many of the members are different from one race to the next. This was done by creating a dummy variable which takes a value of 1 if the runner was not part of the team that ran the previous race, and a value of 0 if the runner was part of the team that ran the previous race. This dummy variable was summed for each team and race. For this variable, it is expected that an increase in the number of different runners will cause the time of the relay to increase. The logic behind this is straight forward; the more different team members over time, the less the runners have a chance to compete and practice with each other, and therefore they will make more mistakes while running and during the exchange of the baton. Also, as the athletes are more familiar wit each other, the risk for negative peer effects such as moral hazard is reduced; “When the same situation repeats itself over time, the effects of uncertainty tend to be reduced and dysfunctional behavior is more accurately reveled, thus alleviating the problem of moral hazard” (Holmström 1979, 90).

The last variable is the year that the race took place. The base year is 2002, which is the first year of observations, and the years are controlled by dummy variables. Overall, the regression output should show that the relay time and year has an inverse relationship, but without a big difference in the values of the coefficients from year to year.

In my regression, I only included teams on which all players were individually ranked. The reason for this was that I was not sure of how to quantify the other runner’s relative ranking or capability compared to the ranked runners. Obviously, since these runners were not on the ranking lists, they are slower that the ranked runners. But the question is by how much? Also, how are their capabilities compared to the other non-ranked runners? One solution could be to give all non-ranked runners a certain amount of (minimum) points, somewhere below the ranked runner’s scores, but high enough to represent a NCAA Division I runner in these types of meets. However, the problem with this solution is that it does assume all non-ranked players are equally good.

By only including teams with all players ranked, 928 observations were deleted, that is, 232 teams, leaving me with 292 observations (73 teams). However, of those 73 teams, three of them did not have an official time from the race. That is, they were either disqualified during the race, or did not start the race at all. Therefore, I also decided to drop these to be able to complete the regression. For the final regression output, I was thereby left with 280 observations for 70 teams.

CHAPTER 4

EMPIRICAL APPLICATION

Peer effects on relay teams are tested by using NCAA Division I track and field results from men's 4x400m relay. The time of the relay, which is run by four individuals each running a leg of 400m, will then be used as a measure of team production output where each of the team members contribute not only with their own quality, but where their individual quality is also affected by each of the other team members as well.

In the indoor season, the runners each run two laps before handing the baton over to the next runner. The first runner in each team have to keep separate lanes for the first three curves, that is, the first 250m before entering in to a common lane. There are no regional races during the indoors season, so teams can only qualify for the NCAA championship based on a qualifying time. The number of teams qualifying to the NCAA championships fluctuates from year to year depending on the rankings for that year, but is usually between 10 and 12 teams. Since the NCAA indoor championship only lasts for two days, there are also no prelims during the meet. Rather, the teams that qualified run the finals directly, with the placing of the teams based solely on time. That is, the fastest team wins independent of the section in which they run.

The outdoor season is structured rather differently. To qualify for the NCAA championship, a team must first qualify for the NCAA regional meet, which they do either by winning their conference in the 4x400m relay, or by qualifying by running a specified time earlier during the season. Once a team has qualified for the regional meet, they must be one of the top five teams from their region to qualify for the NCAA championship. Without competing at the regional meet, there is no way to get to the NCAA championships, even though a team might have the fastest time in the nation before the meet. During the regional meet, qualified teams only run finals, and again the teams compete solely on time. At the championship, which is a four day meet outdoors, teams first run prelims and the eight fastest teams qualify for the final. When the meet has both prelims and finals, the teams advance from being the heat winner or from time. Depending on how many teams are in the prelims, there will be a set number of teams from each heat that will advance to the final purely based on their place in that heat and independent of what time they had. Other than this set number of teams, some additional teams can advance to the finals based on their times. How many additional teams advance depends on how many heats there are and what the set number of teams is, since the total number of teams in the final is eight. The reason for having a total of eight teams advancing to the finals is so that the race can be run in one heat only, and since there only are eight lanes around a track, only eight teams can advance.

For the outdoor relay it is again the first three curves, which here measures 500m, that are run on separate lanes before the runners can cut in to the inner lane. A team will be disqualified from stepping into another team’s lane during the first 500m if this gives the team an advantage compared to the other teams (that is, if they step into a lane on the inside of their own), or if the team members do not exchange the baton within the exchange zone (they can accidentally exchange the baton before the start of the zone, or not make the exchange before the zone ends). The exchange zone measures 20m both indoors and outdoors, 10m before and 10m after the starting line. Other reasons for disqualification include a team member dropping the baton or someone pushes or cuts off another team's runner. If the baton is dropped, the team member can pick it up if and only if they enter the track at the same spot they exited, if the baton falls of the track. The main purpose is so that a team does not gain an advantage over other teams.

The members of the team actually running the championship or the regional meet do not have to be the same members that qualified for the race, and the team members can also be changed between the prelims and the finals. Often a team will save their better runner(s) for the final race, especially if this athlete is competing in many other events during the meet. However, this is clearly only an option for the better schools with more than four runners who are capable of running fast enough.

The dependent variable in this research, the time of the relay, is a summation of the four runner's individual splits. Despite this, one cannot completely compare these times since there are small differences in every leg. The first leg, for example, starts in starting blocks, whereas the other members have a "running start", that is, where the runner can start jogging before receiving the baton and that way is able to more quickly get the speed they need for the race, which will reduce time. The first full leg and parts of the second leg are run on separate lanes, while the other two legs are run on a "common lane", which can be a good or a bad thing depending on what lane you have (it is usually preferable to have one of the lanes in the middle compared to the most inner or outer lane) and how the race is run (e.g. if there is a lot of pushing or if there are a lot of people etc.). Also, since the runners have an exchange zone in which the exchange of the baton must take place, the actual lengths of the different runner's legs are different, and can be either longer of shorter than the actual 400m. The forth leg, for instance, will almost always run the longest leg, since he starts before the starting line, and must run through the finish line.

CHAPTER 5

EMPIRICAL RESULTS

The results of the model are based on 280 observations. The results are reported in three parts. First, a regression using basic OLS model (See table 1). In the reduced form model, all coefficients are statistically significant at the 5% level expect for the standard deviation of the team member's ranking variable and the standard deviation squared. The model's overall fit is also significant, with an R-squared of 0.5932 and an adjusted R-squared of 0.5749. Since the data are semi cross-sectional, an R-squared of this size is good.

The results from the regression output for the basic OLS model are used to look at the absolute effect of the disparity on team performance. The average ranking has an inverse relationship to the dependent variable, and the average ranking squared has a positive relationship with *TIME*. As the average quality of the athletes within a team increases *TIME* will drop, hence the relay team will run faster. But at the same time, for every unit increase in the average ranking squared, the time will increase due to the diminishing returns of the average quality. I also tested the hypothesis that the two variables, *AVERANK* and $AVERANK^2$, are jointly equal to zero. This hypothesis is rejected with an F-test statistic of 64.42 and a p-value of 0.0000.

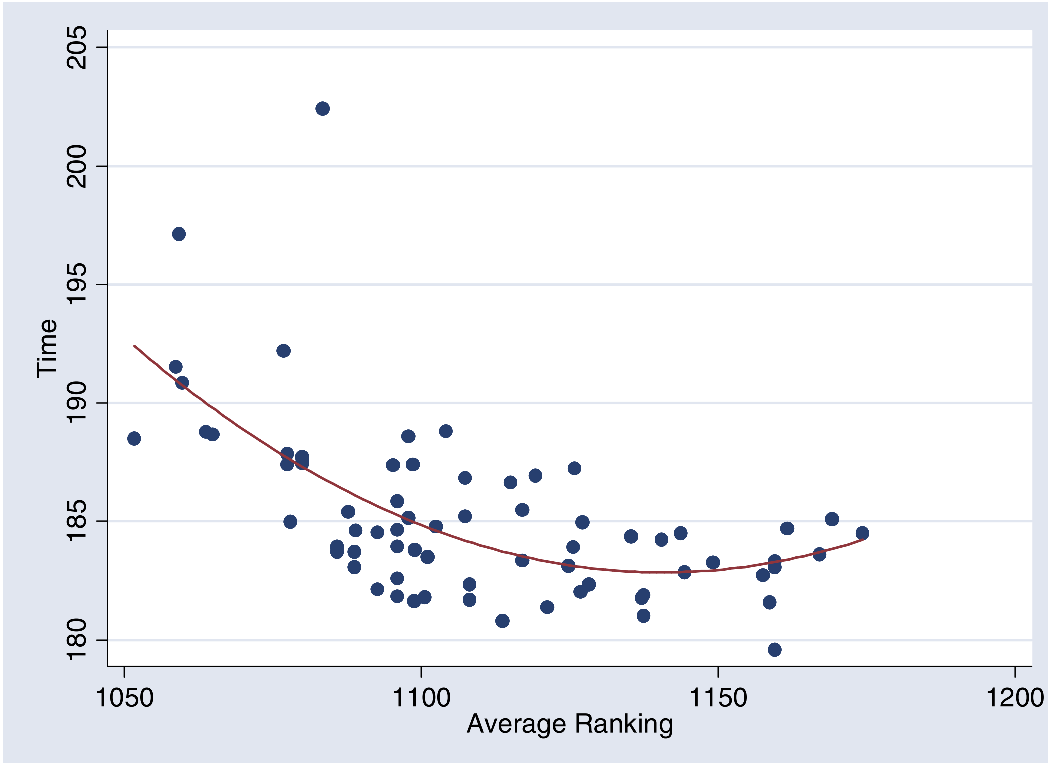

Figure 5.1: Scatter plot of time and the team members' average ranking

The scatter plot *TIME* against the team members' average ranking (*AVERANK*) shows slight curvature and not just a straight line, suggesting there is a point where the average ranking would maximize performance, but that too much quality actually lowers team performance as will too little quality.

The results for the standard deviation parameter suggests that, according to the model, the further away the athletes are to each other in quality, the slower the time of the relay will be. However, the variable came out as insignificant, that is, the result is not distinguishable from zero. The results for the standard deviation of the team members' ranking squared also came out as insignificant, and we can therefore not differentiate the estimate from zero. A hypothesis test that these two variables are jointly equal to zero yields an F-test statistic of 1.17 and a p-value of 0.3134; there is not enough evidence to differentiate these variables from zero.

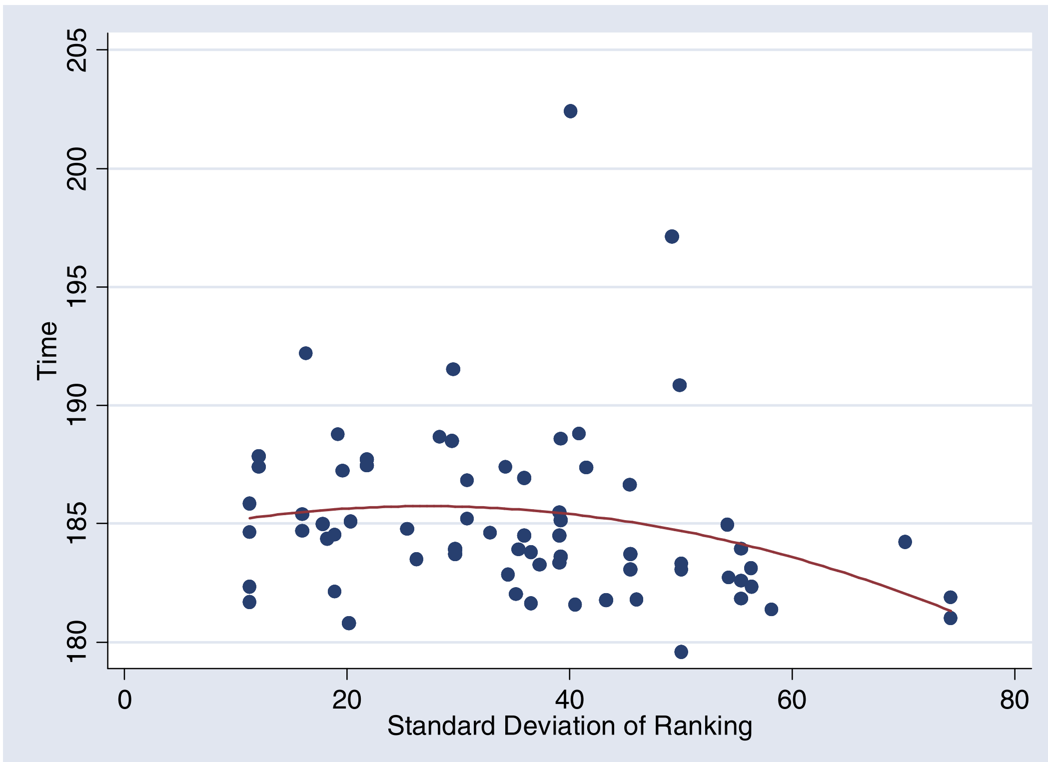

Figure 5.2: Scatter plot of time and the standard deviation of the team members' ranking

The scatter plot *TIME* against the standard deviation of the team members' average ranking (*SDRANK*) shows some curvature just as the previous scatter plot. However, this time the fitted values give a concave curve, suggesting that either a smaller or a larger disparity might be better for the teams' performance than being somewhere in the middle. However, one thing to consider with this graph is some the outliers, as for example, the three observations to the very right of the scatter plot. These suggest that with more disparity on a team, performance improves. Although, what might be one of the reasons in this case is that all of these three observations have two all-stars on their team, which obviously will affect both the range of the quality as well as the performance of the team. However, the slightly negative relationship between time and the standard deviation does confirm the econometric results from the regressions.

For the other variables, *TIME* will be faster for outdoor races compared to indoors, and finals will be faster than prelims, which in turn will be faster than regional races, and for every change in the number of runners from one race to the other, *TIME* will increase.

The basic regression was re-estimated with an additional dummy variable, *ALLSTAR*, to control for the possibility of a team having an all-star. This was done to test whether an all-star on the team changes the parameter of the standard deviation of the team members' ranking since an all-star definitely would increase the range of the runners' quality. The variable is based on the ranking of the individual team member. An all-star is defined as a runner who was more than one and a half standard deviations above the mean of the total points. The mean of the total points over the 1220 player observations is 1112.0 points and the standard deviation is 45.7 points. An all-star would thereby be a runner who had over 1180.5 points. In the total data sample, there were 57 runners who qualified as an all-star. For the smaller sample size, where only the fully ranked teams were considered, only 32 of the runners qualified as all-stars.

However, in the basic regression, the variable *ALLSTAR* was insignificant. Therefore, we can conclude that this variable does not have an effect on the quality of the team based on the standard deviation, and therefore, *ALLSTAR* was not included in any of the other regressions.

To be a bit more specific about the results, the basic model (See table 1) shows that a one unit increase in the team average ranking will reduce the team's time by 1.83 seconds, whereas an increase in the average ranking squared will increase the time of the relay with 0.001 seconds. By running a race outdoors rather than indoors, *TIME* will drop by 3.54 seconds, ceteris paribus. A prelim time will be 2.24 seconds faster than a time from a regional race, ceteris paribus, and a final race time will be 3.86 seconds faster than a time from a regional race, ceteris paribus. Finally, for every additional member changed in the relay team from race to race, *TIME* will increase by 0.83 seconds, ceteris paribus.

The only significant coefficient with a positive relationship to the dependent variable is the *NUMDIFF* variable, which tells us that for every additional different runner that combines a team from race to race, the slower the time will be. Just by looking at the different year variables, the regression results suggest that times decreased a little each year except for in year 2006. A hypothesis test that the year dummy variables were jointly equal to zero yielded an F-test statistic of 3.94 and a p-value of 0.004, stating that at least one of the variables are not equal to zero.

The constant variable is only an intercept term, and will therefore not bring too much information to the analysis. Yet, based on the assumption that all other variables are zero, the time for the relay would be about 1249.47 seconds, which would be a time over 20 minutes.

One problem with the basic OLS model was that of possible heteroskedasticity. Heteroskedasticity occurs when the error does not have a constant variance, and is a common problem in cross-sectional data. When having heteroskedasticity in the data, the t-statistic becomes invalid as the standard error is either under- or overestimated.

As an extension to the basic OLS model, the model was estimated controlling for heteroskedasticity by using robust standard errors. Robust model controls the model for heteroskedasticity to some level, but without re-estimating the parameter estimates. Robust regression can be used both when suspecting heteroskedasticity and when dealing with outliers in the data set. What it does is that it gives lower weight to outliers that to the ordinary observations. By looking at the scatter diagrams of my dataset, I knew that this was the case for some of my observations as well. A White's test supports the existence of heteroscedasticity with a test statistic of 267.73 and a p-value of 4.4e-30.

The result from this model shows basically the same thing as did the basic model, except that all variables are now statistically different from zero at the 5% level (See table 1). Since the standard deviation of the team members' rankings and the standard deviation squared are now significant, I can explain these results more specifically. The standard deviation of the team members' ranking has a positive relationship with time, which was the similar of what I expected the relationship to be, and the parameter estimate suggests that one unit increase in the standard deviation of the team members' ranking will increase the time of the relay by 0.06 seconds, ceteris paribus. For the standard deviation squared, the results show an inverse relationship with time, suggesting that a one unit increase in the standard deviation squared will decrease the time of the relay with 0.007 seconds, ceteris paribus.

The last extension I did in the basic OLS model was to estimate a regression with team fixed effects. By including for team fixed effects, the model controls for any unobserved and omitted variables that are constant over time but are different between cases, and capture these in the fixed team effects. More specifically, the team fixed effects control for things such as the coach, the program, the school, and the team's history. These are factors that we know will be relevant when considering problems of negative peer effects, but since there is no good way of measuring them, the team fixed effects are used instead. The reference team in this model is Arizona State University.

The output from the model shows again that all the original variables are significant, but only 10 out of the 20 dummy variables for the different teams are significant (See table 2). The actual results between the teams varied considerably, but of the 10 significant teams, three were faster than Arizona State University, namely Louisiana State University, Sam Houston State University, and University of Southern California. An example can be shown with Louisiana State University, which is one of the top schools in the nation when it comes to the 4x400m relay. The parameter estimate suggests that on average, LSU was 1.27 seconds faster than Arizona State University, ceteris paribus. The other seven schools with significant effects had a positive parameter estimate relative to Arizona State University during the sample period, meaning that on average these teams' relay times were slower than the reference team, ceteris paribus.

This model provides a very important result for this paper. The positive parameter estimate for *SDRANK*, the negative parameter estimate for $SDRANK^2$, and both of them being significant, and the positive parameter for *NUMDIFF*, shows that there are net negative peer effects on the relay teams, and that it does affect the team's performance. These results imply that the standard deviation of the team member's ranking does reduce team productivity in an absolute sense, that is, the time of the relay is slower when net negative peer effects exist.

The second part of the results is an expansion from the basic OLS model. Here the variable *PRANK* is the dependent variable instead of *TIME*. This variable is generated from the time that the relay teams actually ran. The time of the relay is converted to a set of points using the same conversion table as for the individual rankings. The points for the average ranking of the team are then subtracted from the set of points from the actual relay time. One is thereby left with the difference, in points, between the actual performance and the expected performance of a particular team. This difference is denoted *PRANK*. Using *PRANK* as the dependent variable allows for testing whether race and team characteristics correlate with the team over or under achieving. Also, the results from this regression investigate the relative performance of the relay teams, not the absolute performance as in the basic OLS regression.

For this model, both the basic model and that with the team fixed effects are estimated (See table 1). Here, all the main variables are again significant at the 5%

level, and 11 out of the 20 teams show significant fixed effects. The results suggest that a marginal change in quality improves team production, but that team member disparity reduces a team's relative performance. This means that with a greater disparity between the team members quality, the performance of the team decreases in a relative sense, that is, the team does not perform as well as it was expected to do.

Finally, a probit model was estimated with a dummy variable that takes a value of one if the team is under-performing as the dependent variable. This model suggests both an alternative way of looking at the dilemma of negative peer effects and teams' performance as well as a confirmation of the previous results. The probit model shows how the independent variables influence the probability of a team performing worse than their expected outcome. In the probit regression, the same independent variables are included as in the basic model.

The results from the probit model shows that both the standard deviation of the teams' ranking as well as the standard deviation squared became insignificant again, along with the dummy variables for the year variables and the intercept term. However, the probit model provides the marginal probabilities of the various independent variables included.

These are interpreted in the following fashion; a marginal change in the average ranking of a team, evaluated at the sample mean, decreases the probability that the team under-performs by 41.6%. That is, there is a greater chance that the team will perform just as well as they are expected to, or maybe better. However, since the standard deviation parameter came out as insignificant, the results suggest that the disparity of the team member quality does not impact the probability that the team will under-perform.

CHAPTER 6

CONCLUSION

This paper shows that a higher disparity in quality on a relay team leads to a reduction in team performance due to negative peer effects. The parameters of interest are the team members' average ranking and the standard deviation of team members' rankings, as well as the number of different runners on a team from one race to the next. The team members' rankings are used to measure the quality of the runners, and are related to the performance of the teams. The standard deviation of team member quality is used to capture any peer effects on the relay teams. The findings are based on team level data rather than at the individual level, therefore, it is not possible to identify which individual is causing the negative peer effects.

Starting with a basic OLS regression the model is extended to control for heteroscedasticity as well as team fixed effects. Next, a new dependent variable measuring whether a team under or overachieves, given its inputs, is created. While the first OLS model investigates the absolute performance of the relay teams, the second set of models investigate the relative performance of relay teams. The results from these two models show that net negative peer effects do exist and that they meaningfully affect the performance of the relay teams, both in an absolute and a relative sense.

As a third robustness check of these results, a probit model is estimated. This model shows that a marginal change in the average quality of a team decreases the probability that the team under-performs, but the standard deviation of the team member's quality does not impact the probability that the team under-performs.

These results suggest that greater disparity in team member quality increases NCAA relay team times, which suggests the existence of negative peer effects. These negative peer effects can arise because of free-riding or shirking, but could also arise from other non-monetary reasons such as jealousy or mistrust. This is important for coaches to be aware of as to maximize the performance of a relay team. The results support the existing literature, as similar evidence already has also been shown in other sports such as baseball. The results are also supportive of the team cohesiveness hypothesis, which states that greater wage (quality) disparities within a team can decrease team production.

An interesting extension to this research would be if one can look at individual as well as team level negative peer effects. For example, if information about individual splits from each race and each runner were available, one would be able to learn more about the characteristics of the runners, and who would be more prone to create negative peer effects. If it then were possible to evaluate the same runners in different types of races and compare these performances with the individual's personal bests in the event, one might learn more specifically what causes people to not only "create" negative peer effects, but also how such effects impact the other team members.

APPENDIX A

IAAF RULE 170: RELAY RACES[2]

[2] Source: International Association of Athletic Federations Competition Rules 2006-2007

1. Lines 5cm wide shall be drawn across the track to mark the distances of the stages and to denote the scratch line.

2. Each take-over zone shall be 20m long of which the scratch line is the centre. The zones shall start and finish at the edges of the zone lines nearest the start line in the running direction.

3. The scratch lines of the first take-over zones for the 4x400m (or the second zones for the 4x200m) are the same as the start lines for the 800m.

4. The take-over zones for the second and last take-overs (4x400m) will be the 10m lines either side of the start/finish line.

5. The arc across the track at the entry to the back straight showing the positions at which the second stage athletes (4x400m) and third stage athletes (4x200m) are permitted to leave their respective lanes, shall be the same as the arc for the 800m event, described in Rule 163.5.

6. 4x100m and, where possible, 4x200m relay races, shall be run entirely in lanes. In 4x200m (if this event is not run entirely in lanes) and 4x400m relay races, the first lap, as well as that part of the second lap up to the line after the first bend (breakline), will be run entirely in lanes. Note: In the 4x200m and 4x400m relay races, where not more than 4 teams are competing, it is recommended that only the first bend of the first lap should be run in lanes.

7. In relay races of 4x100m and 4x200m, members of a team other than the first runner may commence running not more than 10m outside the take-over zone (see paragraph 2 above). A distinctive mark shall be made in each lane to denote this extended limit.

8. In the 4x400m relay race, at the first take-over, which is carried out with the athletes remaining in their lanes, the 2nd runner is not permitted to begin running outside his take-over zone, and shall start within this zone. Similarly, the 3rd and 4th athletes shall begin running from within their take-over zones. The 2nd athletes in each team shall run in lanes as far as the nearer edge of the breakline marked after the first bend where athletes may leave their respective lanes. The breakline shall be an arced line, 5cm wide, across the track, marked at each end by a flag at least 1.50m high, positioned outside the track, 30cm from the nearest lane line. Note: To assist athletes identify the breakline small cones or prisms, 5cmx5cm and no more than 15cm high, preferably of different colour from the breakline and the lane lines, may be placed on the lane lines immediately before the intersection of the lane lines and the breakline.

9. The athletes in the third and fourth legs of the 4x400m relay race shall, under the direction of a designated official, place themselves in their waiting position in the same order (inside to out) as the order of their respective team members as they complete 200m of their legs. Once the incoming athletes have passed this point, the waiting athletes shall maintain their order, and shall not exchange positions at the beginning of the take-over zone. If an athlete does not follow this Rule, his team shall be disqualified. Note: In the 4x200m relay race (if this event is not run entirely in lanes) the athletes in the fourth leg shall line up in the order of the start list (inside to out).

10. In any relay race, when lanes are not being used, including when applicable, in 4x200m and 4x400m, waiting athletes can take an inner position on the track as incoming team members approach, provided they do not jostle or obstruct another athlete so as to impede his progress. In 4x200m and 4x400m, waiting athletes shall maintain the order in accordance with paragraph 9.

11. Check-Marks. When all or the first portion of a relay race is being run in lanes, an athlete may place one check-mark on the track within his own lane, by using adhesive tape, maximum 5cmx40cm, of a distinctive colour which cannot be confused with other permanent markings. For a cinder or grass track, he may make a check-mark within his own lane by scratching the track. In either case no other check-mark may be used.

12. The baton shall be a smooth hollow tube, circular in section, made of wood, metal or any other rigid material in one piece, the length of which shall be 28 to 30cm. The circumference shall be 12 to 13cm and it shall not weigh less than 50gm. It should be coloured so as to be easily visible during the race.

13. The baton shall be carried by hand throughout the race. Athletes are not permitted to wear gloves or to place substances on their hands in order to obtain a better grip of the baton. If dropped, it shall be recovered by the athlete who dropped it. He may leave his lane to retrieve the baton provided that, by doing so, he does not lessen the distance to be covered. Provided this procedure is adopted and no other athlete is impeded, dropping the baton shall not result in disqualification.

14. In all relay races, the baton shall be passed within the take-over zone. The passing of the baton commences when it is first touched by the receiving runner and is

completed the moment it is in the hand of only the receiving runner. In relation to the take-over zone, it is only the position of the baton which is decisive, and not the position of the bodies of the athletes. Passing of the baton outside the take-over zone shall result in disqualification.

15. Athletes, before receiving and/or after handing over the baton, should keep in their lanes or zones, in this last case until the course is clear to avoid obstruction to other athletes. Rule 163.3 and 4 shall not apply to these athletes. If an athlete willfully impedes a member of another team by running out of position or lane at the finish of his stage, his team shall be disqualified.

16. Assistance by pushing or by any other method shall result in disqualification.

17. Once a relay team has started in a competition, only two additional athletes may be used as substitutes in the composition of the team for subsequent rounds. Substitutions in a relay team may be made only from the list of athletes already entered for the competition whether for that or any other event. Once an athlete, who has started in a previous round, has been replaced by a substitute, he may not return to the team. If a team does not follow this Rule, it shall be disqualified.

18. The composition of a team and the order of running for a relay shall be officially declared no later than one hour before the published first call time for the first heat of each round of the competition. Further alterations must be verified by a medical officer appointed by the Organizing Committee and may be made only until the final call time for the particular heat in which the team is competing. If a team does not follow this rule, it shall be disqualified.

APPENDIX B

VARIABLE NAMES AND DESCRIPTIONS

Varname	Description	Source
Time	The final time of the 4x400m relay	Trackshark.com
Averank	The average quality of the team	Trackshark.com and IAAF.org
$Averank^2$	The average quality squared	Calculated by author
Sdrank	The standard deviation of the team members' quality	Calculated by the author
$Sdrank^2$	The standard deviation squared	Calculated by the author
Outdoors	If the race is run outdoors	Trackshark.com
Indoors	If the race is run indoors	Trackshark.com
Finals	If the race is a final race	Trackshark.com and NCAA.com
Prelims	If the race is a prelim race	Trackshark.com and NCAA.com
Regionals	If the race is a regional race	Trackshark.com and NCAA.com
Numdiff	The number of different runners on a particular team from one race to the other	Calculated by the author
Prank	The difference between the actual quality and the expected quality of the team	IAAF.org, and calculated by the author
Prankhigher	If the difference between the actual and expected quality of the team is negative	Calculated by the author
Allstar	If a runner's ranking is more than one and a half standard deviation above the mean	Calculated by the author

APPENDIX C

DESCRIPTIVE STATISTICS

Varname	Obs	Mean	Std. Dev	Min	Max
Time	280	185.115	3.654	179.59	202.43
Averank	280	1109.581	30.206	1051.75	1174.375
$Averank^2$	280	1232079	67342.12	1106178	1379157
Sdrank	280	35.653	15.498	11.247	74.293
$Sdrank^2$	280	1510.535	1224.271	126.5	5519.5
Outdoors	280	0.957	0.202	0	1
Indoors	280	0.042	0.202	0	1
Finals	280	0.342	0.475	0	1
Prelims	280	0.3	0.459	0	1
Regionals	280	0.357	0.480	0	1
Numdiff	280	0.685	0.995	0	4
Prank	280	-1.25	46.200	-231.375	59.187
Prankhigher	280	0.471	0.500	0	1
Allstar	280	0.103	0.305	0	1

APPENDIX D

ECONOMETRIC RESULTS

Table 1: Empirical Results

Variable	(1) OLS Time	(2) Robust OLS	(3) Team Fixed Effects	(4) OLS Prank	(5) Team Fixed Effects
Averank	-1.834	-1.834	-3.512	26.553	51.631
	(4.92)**	(4.72)**	(7.68)**	(4.80)**	(7.59)**
$Averank^2$	0.001	0.001	0.002	-0.012	-0.023
	(4.74)**	(4.55)**	(7.55)**	(4.80)**	(7.60)**
Sdrank	0.061	0.061	0.196	-0.888	-2.846
	(1.52)	(2.21)*	(4.89)**	(1.50)	(4.79)**
$Sdrank^2$	-0.001	-0.001	-0.002	0.011	0.026
	(1.43)	(2.32)*	(3.27)**	(1.45)	(3.23)**
Outdoors	-3.544	-3.544	-2.831	-46.920	-58.610
	(3.66)**	(4.99)**	(3.59)**	(3.26)**	(5.00)**
Finals	-3.861	-3.861	-2.940	59.201	45.594
	(9.38)**	(9.91)**	(8.70)**	(9.70)**	(9.07)**
Prelims	-2.241	-2.241	-1.945	33.743	29.478
	(5.97)**	(6.56)**	(6.93)**	(6.06)**	(7.07)**
Numdiff	0.831	0.831	0.761	-12.345	-11.270
	(5.04)**	(4.77)**	(5.45)**	(5.05)**	(5.42)**
Year 2003	-3.064	-3.064	-3.964	45.903	58.525
	(2.26)*	(5.33)**	(3.63)**	(2.29)*	(3.60)**
Year 2004	-4.023	-4.023	-5.807	61.281	87.081
	(3.10)**	(8.27)**	(5.42)**	(3.18)**	(5.46)**
Year 2005	-4.516	-4.516	-6.254	68.666	93.864
	(3.52)**	(8.52)**	(6.01)**	(3.61)**	(6.07)**
Year 2006	-4.111	-4.111	-4.854	62.416	72.537
	(3.13)**	(8.29)**	(4.47)**	(3.20)**	(4.49)**
Constant	1,249.478	1,249.478	2,182.171	-14,788.200	-28,724.673
	(6.04)**	(5.79)**	(8.55)**	(4.82)**	(7.57)**
Observations	280	280	280	280	280
R-squared	0.59	0.59	0.81	0.44	0.74
F-test	32.44		34.57	17.47	22.75

* Significant at 5%; ** Significant at 1%
Absolute value of t statistics in parentheses

Table 2: Empirical Results Continued

Variable	(6) Probit	(7) Probit, d(P=1)/dX
Averank	-1.043	-0.4161564
	(4.47)**	
Averank^2	0.000	0.0001893
	(4.51)**	
Sdrank	0.009	0.0036129
	(0.32)	
Sdrank^2	-0.000	-0.0001667
	(1.17)	
Finals	-2.407	-0.7234414***
	(7.58)**	
Prelims	-1.184	-0.4355738***
	(5.02)**	
Numdiff	0.455	0.1813419
	(3.61)**	
Year 2003	-6.942	-0.6943084***
	(0.05)	
Year 2004	-7.014	-0.9856827***
	(0.05)	
Year 2005	-7.160	-0.9971806***
	(0.06)	
Year 2006	-7.712	-0.9453778***
	(0.06)	
Constant	581.119	
	(.)	
Observations	268	268
R-squared		
X^2-test	150.88	

* Significant at 5%; ** Significant at 1%
Absolute value of t statistics in parentheses
*** Is for discrete change of dummy variable from 0 to 1
Outdoors dropped; predicts failure perfectly

Table 3: Empirical Results, Model 3 and 5 with Team Fixed Effects from Table 1
Reference Team: Arizona State University

	(3) Team Fixed Effects	(5) Team Fixed Effects
Baylor	-0.960	14.427
	(1.16)	(0.17)
Brigham Young	1.982	-30.113
	(2.63)**	(2.69)**
Clemson	-1.666	23.673
	(1.29)	(1.24)
Florida	1.105	-16.686
	(1.99)*	(2.02)*
Georgetown	-0.937	13.940
	(0.78)	(0.78)
Hampton	3.118	-46.618
	(3.19)**	(3.21)**
Illinois	6.814	-98.641
	(8.42)**	(8.20)**
Kentucky	1.584	-22.812
	(1.42)	(1.37)
LSU	-1.273	19.994
	(2.11)*	(2.23)*
Minnesota	-0.142	2.475
	(0.20)	(0.24)
Miss State	-0.302	4.731
	(0.40)	(0.42)
Nebraska	3.121	-44.431
	(2.29)*	(2.20)*
Oregon	-0.862	13.782
	(1.17)	(1.26)
Sam Houston State	-6.215	92.994
	(4.44)**	(4.47)**
South Carolina	-1.060	15.537
	(1.55)	(1.53)
TCU	0.851	-12.370
	(1.04)	(1.01)
Texas Tech	1.963	-28.498
	(2.96)**	(2.89)**
UCLA	3.122	-48.301
	(4.30)**	(4.47)**
USC	-2.111	31.525
	(2.48)*	(2.49)*

* Significant at 5%
** Significant at 1%
Absolute value of t statistics in parentheses
Model numbers correspond to those in Table 2

REFERENCES

Alchian, Armen A., and Harold Demsetz. "Production, Information Costs, and Economic Organization". <u>The American Economic Review</u>. 62.5 (1972): 777-795.

Baiman, Stanley, and Joel S. Demski. "Economically Optimal Performance Evaluation and Control Systems". <u>Journal of Accounting Research</u>. 18 (1980): 184-220.

Barron, John M. and Kathy Paulson Gjerde. "Peer Pressure in an Agency Relationship". <u>Journal of Labor Economics</u>. 15.2 (1997): 234-254.

Carmichael, Fiona, Dennis Thomas, and Robert Ward. "Team Performance: The Case of English Premiership Football". <u>Managerial and Decision Economics</u>. 21.1 (2000): 31-45.

Depken, Craig A. "Wage Disparity and Team Productivity" Evidence from Major League Baseball". <u>Economic Letters</u>. 67.1 (2000): 87-92.

Drago, Robert, and Geoffrey K. Turnbull. "The Incentive Effects of Tournaments with Positive Externalities among Workers". <u>Southern Economic Journal</u>. 55.1 (1988): 100-106.

"DSS Panel Data." Data and Statistical Services. 2006. Princeton University Library. 5 April 2007 <http://dss.princeton.edu/online_help/analysis/panel.htm>.

Falk, Armin, and Andrea Ichino. "Clean Evidence on Peer Effects". <u>Journal of Labor Economics</u>. 24.1 (2006): 39-57.

Forster, Edgar, and Heinz Steinmuller. "An Alternative View on Moral Hazard". <u>The Journal of Risk and Insurance</u>. 45.3 (1978): 531-534.

Green, Jerry R., and Nancy L. Stokey. “A Comparison of Tournaments and Contracts”. The Journal of Political Economics. 91.3 (1983): 349-364.

Gould, Eric D., and Eyal Winter. “Rational Peer Effects Between Workers: Evidence from Professional Baseball”. Mimeo, The Hebrew University (2005).

Grossman, Sanford J., and Oliver D. Hart. “An Analysis of the Principal-Agent Problem”. Econometrica. 51.1 (1983): 7-45.

Groves, Theodore. “Incentives in Teams”. Econometrica. 41.4 (1973): 617-631.

Holmström, Bengt. “Moral Hazard and Observability”. The Bell Journal of Economics. 10.1 (1979):74-91.

---, “Moral Hazard in Teams”. The Bell Journal of Economics. 13.2 (1982): 324-340.

IAAF International Association of Athletics Federations. 1996. International Association of Athletics Federations. Jan 10 2007 <http://www.iaaf.org>.

Jacob, Nancy L., and Alfred N. Page. “Production, Information Costs, and Economic Organization: The Buyer Monitoring Case”. The American Economic Review. 70.3 (1980): 476-478.

Jones, Gareth R., and Jennifer M. George. “The Experience and Evolution of Trust: Implications for Cooperation and Teamwork”. The Academy of Management Review. 23.3 (1998): 531-546.

Kotlyarenko, Dmitry, and Ronald G. Ehrenberg. “Ivy League Athletic Performance: Do Brains Win?”. Journal of Sports Economics. 1.2 (2000): 139-150.

Lazear, Edward P., and Sherwin Rosen. “Rank-Order Tournaments as Optimum Labor Contracts”. The Journal of Political Economy. 89.5 (1981): 841-864.

Lederman, Doug. “NCAA Finds Fraud at TCU”. Inside Higher Ed Sept 23, 2005. Feb 14, 2007 <http://insidehighered.com/news/2005/09/23/tcu>.

Marschak, Jacob, and Roy Radner. Economic Theory of Teams. New Haven: Yale University Press, 1972.

McAfee, R. Preston, and John McMillan. “Optimal Contracts for Teams”. International Economic Review. 32.3 (1991): 561-577.

Mirrlees, James A. “The Theory of Moral Hazard and Unobservable Behaviour: Part I”. The Review of Economic Studies. 66.1 (1999): 3-21.

Nalebuff, Barry J., Joseph E. Stiglitz. “Prizes and Incentives: Towards a General Theory of Compensation and Competition”. The Bell Journal of Economics. 14.1 (1983): 21-43.

NCAA.org. 2005. National Collegiate Athletic Association. Dec 20 2006 <http://www.ncaa.org>.

NCAAsports.com. 2003. National Collegiate Athletic Association. Dec 20 2006 <http://www.ncaasports.com>.

Schmidt, Martin B., and David J. Berri. “On the Evolution of Competitive Balance: The Impact of an Increasing Global Search”. Economic Inquiry. 41.4 (2003): 692-704.

Sanderson, Allen R., and John J. Siegfried. “Thinking about Competitive Balance”. Journal of Sports Economics. 4.4 (2003): 255-279.

Trackshark.com: College Track & Field. 2002. Ed. Tom Borish. Dec 23 2006 <http://www.trackshark.com>.

BIOGRAPHICAL INFORMATION

Lisa Elin Haglund received her Bachelor of Business Administration in International Business German, and her Bachelor of Arts in German from University of Texas at Arlington in May 2005, and her Masters of Arts in Economics in May 2007 from the same university. She intends to continue her studies and pursue a PhD in economics.