LEARNING TO PLAY NICE: STRATEGY EVOLUTION IN THE NATIONAL HOCKEY LEAGUE[i]

T.K. Ahn
Department of Political Science
Florida State University
543 Bellamy Building
Tallahassee, FL 32306
Phone: 850 644-4540
Fax: 850 644 1367
e-mail: than@fsu.edu

Marco A. Janssen
School of Human Evolution and Social Change
& Department of Computer Science
Arizona State University
PO Box 872402
Tempe, AZ 85287-2402
Phone: 480 965 1369
Fax: 480 965 7671
Email: Marco.Janssen@asu.edu

Derek S. Reinders
Department of Political Science
Indiana University
1100 E. 7th Street
Bloomington IN 47405
Phone: 812 332 8343
Fax: 812 855 2027
e-mail: dreinders@indiana.edu

Jeffrey E. Stake
Indiana University School of Law
211 South Indiana Avenue
Bloomington, IN 47405
Phone: 812-855-4444
Fax: 812 855 055
Email: stake@indiana.edu

LEARNING TO PLAY NICE:
Strategy Evolution in the National Hockey League

Abstract

The effect of increased monitoring and rule-enforcement in National Hockey League (NHL) games is analyzed at two levels (player and team). The economic theory of crime predicts a reduction of rule breaking due to increased deterrence. No change is observed in behavior at the player level. At the team level, however, we find a change in composition in type of players. Private rule enforcers, the goons, become more costly and less necessary when official monitoring is increased. We observe a decrease in the salaries of the players with high level of goonnesss as our game theoretic model predicted.

Keywords: National Hockey League, monitoring, rule breaking, team composition, goons

1. INTRODUCTION

The economic theory of crime predicts that an increase in policing resources will lead to a decrease in the crime rate as potential criminals determine their optimal level of crime based on the expected punishment (Becker, 1968). Empirical studies to test this hypothesis have produced mixed results (Marvell and Moody, 1996; Di Tella and Schargrodsky, 2004). One of the main problems in detecting the impact of policing is the fact that police and crime influence each other simultaneously. When crime levels go up, governments may employ more police, and when police activity goes up, potential criminals may refrain from crime (Marvell and Moody, 1996). The endogenous processes make it difficult to estimate net effects. However, in extreme cases, these specification problems may be avoided. Di Tella and Schargrodsky (2004) estimate the impact of more police after a terrorist attack on the main Jewish center in Buenos Aires, Argentina in 1994. The government increased policing of Muslim and Jewish buildings, leading to increased monitoring in specific spatial locations. Di Tella and Schargrodsky (2004) estimate a significant reduction of crime rates due to more police.

The difficulty of measuring the real effect of increased monitoring on crime is one of the reasons why scholars use sport statistics to test the prediction that an increase in policing resources will lead to lower crime rates. McCormick and Tollison (1984), for example, applied the economic theory of crime to rules infractions in basketball. They analyzed the effect of the 1979 change from two to three referees in the Atlantic Coast Conference (ACC) basketball tournaments, using statistics during the period from 1954 to 1983, and found a reduction in the number of fouls called. In baseball, the effect of the designated hitter rule has been studied by comparing statistics in different leagues in the USA (see for example Trandel, 2004).

In this paper we test the economic theory of crime using data of the National Hockey

League (NHL). During the 1998-99 and 1999-2000 NHL seasons, the League conducted an experiment in which some games had two referees instead of one. Just as only some of the areas of Buenos Aires were monitored more heavily after the terrorist attack (Di Tella and Schargrodsky, 2004), only a subset of the NHL games experienced higher levels of monitoring during the experimental years, making it possible to measure a net effect of monitoring. The economic theory of crime predicts an increase in deterrence as follows. An increase in the number of referees will increase the level of monitoring, which will increase the likelihood a transgressor will be caught, which will increase the expected cost of an infraction, which will reduce the amount of rule breaking.

Several studies have found that an extra referee leads to better monitoring and therefore more penalty minutes for players who break the rules of the game (Allen, 2002; Levitt, 2002; Heckelman and Yates, 2003). These findings are consistent with the hypothesis that increased monitoring leads to increased detection of illegal behavior. But these studies did not find changes in behavior at the individual level and did not look for changes at the team level. That is, to date, the deterrence hypothesis has not been empirically supported for the NHL experiment.

In this article we analyze the NHL experiment on a longer time scale and at both the individual player level and the team level. We observe a significant drop in penalty minutes during the seasons having two referees compared to the seasons having one referee. We argue that the extra referee did change the pattern of behavior in hockey games, but not because the players adapted to the new rules. In fact, individual players' behavior did not change and, therefore, the dominant short term effect of the added referee was on the side of detection, not deterrence. Our empirical analysis of the NHL data includes not only the two seasons when one-referee and two-referee systems were both employed, but also subsequent seasons in which the two-referee system has been used exclusively. We find learning, adaptation at the team level, artificial selection of player types by the teams, changed composition of player types in teams, and, as a consequence of all these, changes in observed behavioral patterns which are consistent with the deterrence hypothesis.

The players of particular interest are the so-called “goons”. Goons are private enforcers or protectors. Goons punish the opposing team either for breaking the rules, when such transgressions are not detected and punished by the referees, or for attacking a protected player. This private enforcement by goons frequently takes the form of rough behavior that is itself a violation of the rules. Thus, if the referees catch the goon’s retaliation, it is costly to the team because the team suffers a penalty. Goons are also costly inasmuch as they do not have as many scoring and defensive skills as the other players.

With more referees on the ice, there is more public enforcement of the rules and there is less need for goons. At the same time, goons become more costly since their illegal actions will be caught more frequently. We would therefore expect the role of goons to diminish when there are more referees. We formulate a game theoretic model that shows the strategic interaction between teams as functions of the level of rule enforcement via referees. We have compiled a data set that contains NHL player level data from 1994 to 2002, which includes variables such as goals, assists, penalty minutes, salary, etc. At the aggregate level, the data show an overall decline of penalties in minutes per team per game since the introduction of the additional referee. The data also indicate that a large part of the decline is due to the selection of players by teams, which is shown by the differential patterns of salary change between goons and non-goons before and after the introduction of the additional referee.

The remainder of the paper proceeds as follows. First we discuss the NHL experiment, and the findings of other articles examining this experiment. Then we formulate a sequential-move goon game, and test the findings on NHL data. We end with some brief conclusions regarding the broader implications of our findings.

2. THE NATIONAL HOCKEY LEAGUE EXPERIMENT

During the course of two seasons the National Hockey League added a second referee to some, but not all, games. One of the goals of increasing the referees from one to two was to increase the chances of catching players that break the rules and to deter them from doing so. Since the players and game are faster today, it is hard for one referee to keep up with the play at both ends of the ice. Two referees can more effectively monitor the players, especially those trailing the puck and the play. In addition to enhancing the monitoring, the division of duties between two referees reduces the requirement for them to skate quickly, making it possible for the league to use older referees, those with the best judgment. Another potential benefit of better monitoring and fewer penalties is that the game can be completed in less time.

This change set up an interesting natural experiment, one that has been analyzed by various scholars (summarized below). Prior to 1998, the NHL employed one referee for each contest. Two linesmen also work each game, but linesmen do not call penalties. During the regular season between October 15, 1998 and February 28, 1999, all 27 NHL teams competed in 10 home games and 10 away games officiated by two referees. That means that about 20% of the games in the season were officiated by two referees instead of one. The assignment of one or two referees to the games was essentially random. In the 1999-2000 season, the number of games each team played with two referees increased to a minimum of 50. Since 2000, two referees have been used for each game.

The classic study of McCormick and Tollison (1984) found that more referees in basketball led to a reduction to the number of fouls called. Studies on the NHL experiments found that there is an increase in number of small penalties and a small decrease in the number of major penalties (Levitt, 2002; Allen, 2002; Heckelman and Yates, 2003: Depken and Wilson, 2004). This might be explained by the fact that an additional referee leads to fewer undetected infractions, and indirectly to less retaliation and fights. It also suggests that the extra referee had a monitoring effect and not a deterrence effect. The prediction of economic theory on the deterrence effect is not supported by previous studies.

Now that more time has passed, more data are available, data in which long-term effects can be seen. Figure 1 depicts the average penalties in minutes per team per game for 9 seasons from 1994-1995 to 2002-2003. These data show that the average penalty minutes – an indication of the extent to which rules are broken and such infractions are caught – has decreased significantly ($p<0.05$), from about 20 in 1994-1995 to about 15 in 2002-2003. In addition, the two experimental seasons of 1998-1999 and 1999-2000 appear to mark the critical point. Either policing became less effective, which seems unlikely with the addition of another referee, or players actually committed fewer infractions.

Thus, the findings of previous studies -- that monitoring became more efficient with the additional referee per game but players did not break the rules less – do not fit the long term trend in the data. Why did the previous studies fail to notice the decrease in penalty minutes with two referees? The previous studies looked at only the experimental seasons. A focus on only the official experimental years has the benefit of holding many factors constant, but it also has drawback of missing the effects of learning over time. In addition to increasing the number of observations, examining a longer period of time allows the observer to pick up changes in behavior at an organizational or team level from one year to the next which might not show up in a comparison of behavior across games within the season. Yet teams might be expected to respond over several seasons to the new rule change. They might, for example, change the type of players composing their teams or the relative ice-time of the different types of players. As noted above, when an additional referee makes monitoring more effective, players who commit a lot of rule infractions become both less useful and more costly.

This has had obvious consequences for the goons. As rule-enforcement by the officials became more effective, teams needed their own enforcers less. At the same time, the goons became more costly, since many of their retaliatory enforcement activities were more likely to be detected and punished by the officials. In the following section we will formulate the strategy of a team in a game theoretic framework, and provide a possible explanation for the trends shown in Figure 1.

[Figure 1 About Here]

3. THE GOON GAME

To analyze how the added referee might affect the teams' strategies in a game, we constructed a simple goon game in which the number of referees is modeled as the probability of detecting a transgression. The two teams in the game model are nicknamed "the Maple Leafs" and "the Red Wings" without any intention of portraying the real teams as having the tendencies depicted in the game model. The payoffs are relative values representing the changes in winning probabilities corresponding to certain strategy combinations[ii].

As can be seen from Figure 2, the first mover is the Maple Leafs. The Maple Leafs must decide whether to commit an infraction of the rules (signified by the word "violate" or "infract" or “transgress”) or to refrain from breaking the rules. In some situations, it makes sense to break the rules. For an easy example, if the opposing team, the Red Wings, has a player on a breakaway, he has a pretty good chance (perhaps nearing 25%, as judged by the frequency of scoring on a penalty shot)[iii] of scoring a goal, and that goal may well be the deciding factor in the game. In some such situations, a Maple Leafs player can use his stick to hook or trip the Red Wings player, thereby preventing the goal. In this example, preventing the goal in a close game improves the Maple Leafs’ chances of winning by 20%, and reduces the Red Wings' chances of winning by the same amount, if the infraction goes undetected by the referee and unretaliated by the Red Wings. The amount of difference the violation will make can of course be higher or much lower. In a game that is already 7 to 1, there is little chance that any move will affect the outcome.

But, of course, that is not the end of the calculations. This hooking violation will often result in a penalty, the probability of which is called p in the game. In such situations, the cheating player will have to spend some time in the penalty box and his team will have to play with fewer players during that time. The next node in the game models that event, with the referee calling a penalty (called "detection") for the violation of the rules. This creates a "power play" for the Red Wings, and the chances of scoring are statistically improved, although the success of teams on power plays varies widely across the league, as does the ability of teams to "kill" (fend off) power plays. Usually, if the referee has penalized the cheating Maple Leafs player, the Red Wings do not need to retaliate. Indeed, retaliation may result in a "stupid penalty" to the Red Wings, canceling the powerplay and nullifying the benefits they would have obtained by the penalty against the Maple Leafs. For that reason, our model does not give the Red Wings the option of retaliating when the referee has penalized the Maple Leafs, even though in theory of course that is also an option.

However, if the referee for whatever reason does not penalize the Maple Leafs player, the next move belongs to the Red Wings. The Red Wings must decide whether to privately enforce the rules. If the Red Wings retaliate, they will often do so by sending a goon out onto the ice to inflict some pain on the offending Maple Leafs player. Of course, even though this is hockey, this is a violation of the rules and may result in a penalty. The effects of the goon's retaliation are numerous. First, the goon may injure the Maple Leafs offender enough that he is unable to play, or play well, in the game. This effect may carry over to future games, and is signified by the -5% effect on the Maple Leafs winning games in the future. In addition, the Maple Leafs players may be somewhat intimidated during the present game and thus may be slightly less likely to win. This effect, too, may carry over to other games. One of the reasons teams reward defensemen for fighting is that it helps intimidate opposing players. Other opponents of the Red Wings in the future, having seen the brutal retaliation against the Maple Leafs, may be somewhat less likely to play aggressively against the Red Wings, giving the Red Wings a small advantage a in those later games.

On the flip side, if the Red Wings do not employ a goon to retaliate, they lose a little credibility, making it more likely to be the victim of cheating and aggressive play by other teams in the future, which is signified by the -5% in future games. Thus, the goon's brutality may result in an advantage for the Red Wings in this and future games. The effects on the future games, of course, do not affect the probability of winning the current game, but they can be translated to the equivalent value of changes in winning probability in the current game. That is, -5% for the Maple Leafs as consequence of having its one of its key players attacked by the Red Wings's goon implies that the Maple Leafs would take any action in the current game that might decrease its probability of winning the current game by 5% to avoid such an attack by the Red Wings.

If the Red Wings retaliate, the referee may impose a penalty, and does so with probability *q* in the game. When that happens, we have essentially the reverse of the situation in which the referee caught the Maple Leafs cheater. In this case, however, the cost to the Red Wings is slightly less because the Red Wings lose the use of a slightly less useful player. Moreover, there may be a long-term future-game advantage to the goon being caught in that the high number of penalty minutes the goon has suffered sends the clear signal that the Red Wings are willing to impose justice when the referees fail to do so.

[TABLE 1 ABOUT HERE]

The goon game (Figure 2) is constructed based on the payoffs shown in Table 1. The payoffs are divided into two sources, future games and the immediate game that the two teams are playing. The payoffs in the immediate games are the changes in the probability of winning. The payoffs in the future games are, as just noted, the benefits for a team that accrue in the future translated as equivalents of the changes in the winning probability in the current game. The two sources of payoffs are combined and incorporated into the payoffs of the Goon game shown in Figure 2. For convenience of analysis the utility payoffs are multiplied by 100; this linear transformation in the scale does not change the nature of the game at all.

[FIGURE 2 ABOUT HERE]

The parameter *p* in Figure 2 is the probability that the Maple Leafs' transgression will be caught by the referees. The parameter *q* is the probability that the referees will catch the Red Wings goon's reaction. The two probabilities are exogenously given for a given game (i.e. they are not strategic choices of the referees). We can think of the two probabilities as the League's policies that are fixed for a given game. Since a goon action is more conspicuous than other illegal actions intended to prevent the opponent from scoring a goal, we assume that q is greater than p.

[TABLE 2 ABOUT HERE]

By backward induction we can verify that the Red Wings will use a goon if $q < 20/(25\text{-a})$. The Maple Leafs' optimal strategy needs to be calculated separately for when the Red Wings use a goon ($q < 20/(25\text{-a})$) and when the Red Wings do not use a goon ($q>20/(25\text{-a})$). For both of the cases, the Maple Leafs' payoff of not transgressing is 0. If the Red Wings are ready to use a goon the Maple Leafs still transgress if $p < (6q+1)/(6q+3)$. Therefore, if $q < 20/(25\text{-a})$ and $p < (6q+1)/(6q+3)$, the subgame perfect equilibrium outcome is [the Maple Leafs infract, the Red Wings retaliate], which is the kind of game in which a lot violence is observed. If, on the other hand, $q < 20/(25\text{-a})$ but $p > (6q+1)/(6q+3)$, the subgame perfect equilibrium is [the Maple Leafs do not infract, the Red Wings use goons] and the outcome we actually observe is [the Maple Leafs do not infract]. That is, in this equilibrium, the Red Wings are ready to use a goon if the Maple Leafs' infraction is not detected by the referees but do not actually have to use the goon because the Maple Leafs do not infract.

When $q > 20/(25\text{-a})$, the Maple Leafs transgress if $p < 2/3$. Otherwise, the Maple Leafs play by the rules. Therefore, when $q > 20/(25\text{-a})$ and $p < 2/3$, the subgame perfect equilibrium is [the Maple Leafs infract, the Red Wings do not use a goon]. When $q > 20/(25\text{-a})$ and $p > 2/3$, the subgame perfect equilibrium is [the Maple Leafs do not infract, the Red Wings do not use a goon].

Our assumption is that the addition of a referee has changed the detection probabilities p and q to, say, p' and q' such that $p' > p$ and $q' > q$. These changes in detection probabilities can have different observable effects depending on how large were the original p and q and how large were the changes in those values due to the added referee. Figure 3 depicts possible equilibrium regimes of the goon game, which helps to understand the effects of rule change on behavior. Given the assumption that $q > p$, the feasible equilibrium space is limited to the triangle demarcated by the bold lines. The four zones in Figure 3 correspond to the four equilibria shown in Table 2. An increase of referees may lead to no observable effect in the use of goons, when increases in p and q are not enough to move the equilibrium zone. Suppose that the League was concerned with too much violence in the game (which corresponds to the equilibrium in zone IV) and added another referee to increase the detection probabilities p and q. This would cause the equilibrium to move somewhat northeast of the original one. However, there is no guarantee that the equilibrium actually moves to another zone, zones I, II, or III. In this case, the only change would be an increase in penalties which results from a combination of the same violent equilibrium and the enhanced detection capability. The findings of earlier studies using the outcomes of the experimental seasons correspond to this scenario. But the equilibrium switch could occur over time, not instantaneously, as we show in later sections of this paper. If that is the case, an examination of the two experimental seasons would fail to find the switch in equilibrium and fail to reveal how changes in rules affect the strategies of individuals and teams.

[FIGURE 3 ABOUT HERE]

If the increase monitoring does indeed disturb the old equilibrium, as the League probably intended, and the equilibrium moves to either zone I or zone III, there would be a substantial reduction of violence in games. Though the two equilibria differ from one another, the outcomes actually observed are the same. That is, in these two equilibria, there is no intentional rule infraction by the Maple Leafs to discourage the Red Wings' skilled players. But there is subtle difference which, at the same time, may have different ramifications for the fate of goons. In zone III, the Red Wings' readiness to use goons is in part responsible for the Maple Leafs' refraining from transgressions. Thus, though the usefulness of goons for the Red Wings is somewhat lower than that in zone IV, the Red Wings might still need to show that they have some goons ready for action to prevent the Maple Leafs' attack on the Red Wings' skilled players. On the other hand, in zone I, the Red Wings would not use goons even when the Maple Leafs violate the rules and the violation is not detected by referees. Here, what deters the Maple Leafs from such actions is the referees and referees alone. Thus, the goons are no longer useful for the Red Wings.

Deriving testable predictions from the game model requires some additional assumptions that are not directly testable by themselves. We assume that (1) the increase in the detection probability due to the addition of a referee was large enough to cause an equilibrium shift, (2) players have rather rigid types (skilled players or goons) that they cannot easily change, and (3) it takes time for the teams (coaches and managers) to implement the new equilibrium strategies by changing the levels of rule infraction and use of goons. These assumptions can be tested indirectly in several ways. In particular we propose the following as testable hypotheses:

(1) The games become less violent after the rule change

- a. The average penalty minutes per game decreases after the rule change.
- b. The average "goonness" of players decreases after the rule change.

(2) Over time, after the addition of a new referee, teams use the goons less.

- a. The disparity between the salaries of goons and non-goons is larger after the rule change than it was before the rule change.
- b. Goons have less ice time after the rule change than they had before the rule

change.

4. EMPIRICAL GOON ANALYSIS

Goons are popular players among hockey fans and various websites list the favorite goons in the NHL. However, “goon” is not an official position, and therefore we can only measure goons indirectly. The fan sites usually define goons by the number of major fights that players were involved in for a given season. The most notorious goons can still skate and use sticks (legally) far better than the rest of us, and the physical nature of the hockey game sometimes puts the most gentlemanly player into fights. Hence, neither willingness to fight nor inability to skate can be used as a simple measure of goonness. We have defined goonness as a composite measure derived from three numbers available for each player: penalty-in-minutes (PIM), goals, and assists. Specifically goonness is defined as PIM/(GOALS+ASSISTS+1). The number of goals and assists, of course, are among the most important indicators of how essential a player is to his team’s reason for being, which is winning hockey games. The PIM is taken as one indicator of how often a player is involved in retaliatory actions. One defect of this measure is that it does not do a very good job of distinguishing good defenders, who break up offensive plays of the other team but lack offensive skills themselves (goalies are the ultimate examples of this, so we have left them out), from goons, whose function is more private policing than defending. Unfortunately, there is no measure of defensive skill in hockey akin to steals and blocked shots in basketball or interceptions and tackles in football.

This section utilizes the goonness measure and other aspects of the data we gathered for the NHL seasons from 1994-1995 to 2002-2003 to study how the rule change in question, i.e., the added referee per game, affects the behavior of the players and the outcomes of the game.

Since many variables of interest are available only for the recent years, we are not able to perform the type of time series analysis we would prefer. We therefore build our evidence on an accumulation of arguments of partial data sources.

We will start by re-examining the hitherto prevalent view that the added referee contributed to better crime detection, but not prevention. The enhanced dataset that we use shows a substantial reduction in “crime” in hockey games, indicating adaptation to the new rules, with learning occurring over time rather than instantaneously. This finding leads us to inquire as to the mechanism by which the change occurred. Did the players learn immediately to refrain from illegal actions after such infractions became more costly? The previous studies discussed in section 2 show that this is not the case. This should not be a surprise. Players have certain skills, and may not adapt instantly to new selection environments. It is hard to change habits built over a lifetime of hockey and achieving offensive excellence in professional hockey is not an option for most goons and defensive defensemen. The mechanism of change appears to reside at a higher level. That is, the teams – the decision makers for teams such as coaches and management – seem to have learned that there are more situations in which the costs of utilizing goons have surpassed the benefits. Thus the new environment subjected the goons to a form of artificial selection, with teams using fewer goons and paying them less. It is the choice of the teams not to use goons, rather than the choice of the players not to be goons. And this change in team management has probably caused the substantial decrease in goon infractions.

4.1. Data base

We constructed a data base of player characteristics for eight seasons, namely from 1995-1996 to the 2002-2003 season. For each player, the usual statistics like games played, goals, assists, and penalty minutes were collected from online resources[iv]. In order to standardize the database, only regular season statistics were included because only select teams (and thus players) advance to the playoffs. Goalies were not included because they are not in a good position on the ice to act as enforcers and they are too valuable to lose by ejection for fighting. Goalies cannot be considered good candidates for goon status and therefore offer no utility in making comparisons.

We also estimated the salaries of the individual players from various sources. Those salary figures are not officially published but based on estimates from a number of online resources[v]. A total of 73% of the possible salary entries were estimated.

4.2 Penalty minutes per game

As has been discussed in Section 2 and illustrated in Figure 1, there is a significant decline in the amount of penalty minutes per game after the rule change compared to the seasons before the rule change. This supports hypothesis 1a.

4.3. Estimating the level of goonness

As noted above, "goon" is not an official recorded position, so we have to estimate indirectly the goonness of the players. Nevertheless we can check our estimates with unofficial goon lists on fan pages, and with the top 25 on the CBS "goon-o-meter"[vi].

We chose official recorded statistics like penalties in minutes, goals and assists, so that we were able to derive estimates of goonness for all players for all eight seasons. CBS's "goon-o-meter" provides different weights to different type of penalties related to the aggressiveness of the conduct, but this detailed information was not available to us. For example, penalty minutes derived for match penalty or gross misconduct is accorded a score of 25, while a major penalty for fighting scores 10 points, and elbowing, kneeing, slashing, spearing, unsportsmanlike conduct, and cross-checking are only weighted by one point per derived penalty minute. Figure 4 shows that the majority of the players have a goonness quotient of less than 3.

[FIGURE 4 ABOUT HERE]

Twenty-two of the goons in the CBS top 25 are wing players. Frequently, the job of the goon is to protect, or deter violence against, a star offensive player who has been hired for skills rather than strength or toughness. Even if a star has the strength and fighting ability to protect himself, it costs the team less for a goon to protect the star and serve the resulting penalty than for the star to protect himself and serve the penalty for doing so. Often these star players are centers, and occasionally wings. Since hockey players are usually substituted in groups, with defensive players grouped in pairs and offensive players grouped into “lines” of three, and since substitutions are made separately for the defensive pairs and offensive lines, it is easier to keep a goon with the player he is protecting if the goon plays with the star on the same offensive line. This explains the different goonness among types of players (around 2). Centers score frequently, and have on average a low goonness rating. With the exception of offensive defensemen, defensemen do not score or assist frequently, but derive many penalties in their role as defender. Although some goons play defensive positions, defensemen may have a high goonness rating without being a goon (around 5). Wings with high gooness ratings, on the other hand, are often goons. For this reason, wings on average have a goonness rating similar to that of the defensemen (around 5). We test statistically whether the goonness of the different positions before the rule change (up to the 1998-1999 season) differs from the goonness after the rule change (from 1999-2000 season). The difference of means shows that goonness has decreased in all positions, but the decrease is significant at the 5% confidence level only for the wings, not for the defenders and centers. The results provide support for our hypothesis (2)-b in particular among the wing players who are more likely to be goons than players in other positions.

When we look at the top 25 goonness rates of wings in the 2001/2002 season, it consists of 13 players of the top 25 from the CBS “goon-o-meter”. Given that there were 860 players in the NHL in that season, our goonness measure is to some degree validated by its overlap with the subjective estimator of the CBS “goon-o-meter”. There is a difference, however, in that players who are highly penalized but score reasonably well are not classified as a goon in our estimator, but do show up on the “goon-o-meter”. For example, Andrei Nazarov, second on the CBS list did not reach our goon list, despite serving 215 penalty minutes, because he scored 6 goals and contributed 5 assists.

With the constructed measure of goonness, we first test if the overall level of goonness has changed. Our hypothesis (1)-b states that as the teams learn to adapt to the rule change, the teams are going to use the goons less. But this behavioral change might not happen instantly. Rather, the effect of the rule change would be observed over time. To test if the level of goonness has gone down due to the rule change we divide the 8 seasons in our database into three groups. They are (1) before the rule change (seasons 1995-1996 to 1997-1998), (2) during the rule change (seasons 1998-1999 and 1999-2000), and (3) after the rule change (seasons 2000-2001 to 2002-2003).

[TABLE 3 ABOUT HERE]

Table 3 shows that average goonness in each of the eight seasons in our database with the number of players parenthesis. The table also shows (fourth column) the average goonness before, during, and after the rule change. It appears that the major reduction of the goonness level was achieved during the two experimental seasons. On average, the seasons during the rule change have about one unit smaller goonness per player compared to the seasons before the rule change. The level remains rather constant from During to After the rule change. Thus, while simply comparing the games with one and two referees did not reveal the effect of rule change, a comparison of the level of goonness over time does. The differences are significant between Before and During (p =.0001) and Before and After (p=.0000) in Wilcoxson ranksum tests, but not significant between During and After.

4.4. Goonness and Salary Change.

Previous studies of the effects of the NHL rule change have found that the effects are more on the side of detection than deterrence. But these studies used a single season in which games were played under both rules – some with one referee and some with two referees. Unless one assumes that the effects of a rule change are immediate, using only the official experimental period may not properly capture all of the effects of rule change. The decreased PIM over the several years since the rule change (Figure 1) indicates that the teams have adapted to the new environment. One of the testable implications is that the value of goons is lower than it used be before the introduction of two referees. In this subsection we test our hypothesis (2)-a, that the difference in salary increase between goons and non-goons is greater after the rule change than before the rule change.

There have been previous studies on the player and team characteristics that affect the salaries of the players. Not surprisingly, these analyses show that star players in teams with higher revenues get higher salaries than others (Idson and Kahana, 2000).

We test whether there is a difference in salary development between players classified as goons by our goonness measure, and other players. Figure 5 shows the percentage of salary change for the players with goonness scores less than 10 (Goon<10) and those with 10 or greater (Goon>10). The figure shows a trend of increasing difference in salary change between the two groups. Notice that before the rule change the difference is negligible: 0.2% point in the season 1995-6 and 2.2% point in the season 1996-7. Around the seasons of the experiment the difference rose to about 7% point, and after the rule change is finalized and two referees were used in all the games, the difference grew even greater. A visual examination of the trends seems to indicate, consistent with our hypothesis, that after the rule change the goons were less valued by the teams.

In Table 4 we test if the differences are significant. We conduct the tests by player position. Due to the small numbers of goons when the data is disaggregated by season, we bifurcate the seasons into before and after the experiment using the season 1999-2000 as the first season after the experiment. For this comparison we used a goonness rating of 10 as the goon threshold. At the 5% confidence level using a two-sample Wilcoxon rank-sum (Mann-Whitney) test, we find statistically significant differences in the salary developments only for wing players. That is, after the NHL experiment, wings who had a high goonness factor in the previous year obtained a significantly lower salary increase than the average player. We also conducted the same analysis using a goonness rating of 25 as the threshold. Again, the difference was the largest among the wing players, but the difference was not significant at 10% confidence level with a *p*-value of 0.15. Due to the long-term contracts and pre-determined salary schedules in contracts, we are not able to show the equivalence of the results in Table 3 in regressions. Thus, we consider the results in Table 4 to be limited support of our hypothesis (1)-a because it shows significance only for the wings.

[TABLE 4 ABOUT HERE]

4.5. Goonness and Ice Time

Our hypothesis (2)-b that goons have less ice time after the rule change is probably the most direct measure of our hypothesis that teams adapted to the rule change by selecting against goon players. There are a few problems, however, that make it difficult to test this hypothesis using the ice time data. First of all, despite our efforts, we could obtain average ice time data only for seasons since 1998 (National Hockey League, 1999-2003). Thus our analysis can use the ice time data for only one season that we classified as BEFORE the rule change. Second, even if we did have complete ice time data, we suspect strong selection bias. If our hypothesis is true, many goon players would have already been selected out by teams and, thus, would have no ice time at all. Therefore, analysis using active players only, especially for seasons after the rule change, would generate somewhat weaker effects. In spite of these problems our regression shows at least some support for the hypothesis.

We conducted a regression of the average ice time change between seasons at the individual player level using the variables “GOONNESS”, a dummy variable “AFTER” that is coded 1 for the seasons 1999-2000 and after, and an interaction variable between GOONNESS and AFTER. The dependent variable is ICETIME CHANGE, the difference between a player’s average icetime between season *t* and *t+1*. All the independent variables are measured at season *t*. Table 5 reports the results of the regression. Notice that the interaction variable has a negative coefficient that is significant at 10% confidence level ($p = 0.063$). This implies that having a high goon score has a larger negative effect on playing time after the rule change than before. A fixed-effect model with the same specification returns similar results in terms of the coefficients and their significance. When we did the regression by player position, the results were not significant, but came close to being significant for the wing players.

[TABLE 5 HERE]

5. CONCLUSION

This paper analyzes the effect of increased monitoring on the behavior of players and teams in the NHL. Our results suggest that players did not change their behavior, but teams did. Teams change their behavior by changing their composition, changing the mix of types of player on the team and the amount of ice time for those players. These results do not support the economic theory of crime at the player level, but do support the theory at the organizational or team level – provided, of course, that certain individual behavior may impose shared costs upon the group.

A lesson from our analysis is that the effect of a rule change may be obscured when we attempt to measure it immediately. Previous studies of the NHL experiment all focused on one or two seasons in which the experiment took place and compared the games with one and two referees. Our analysis shows that the behavioral change becomes measurable when a longer period involving multiple seasons is examined. As Campbell and Ross (1968) demonstrated in their classic study of the crackdown on speeding in Connecticut, researchers, when measuring the effects of a policy change, need to look at the targeted behavior over a sufficiently long time frame before and after the policy change.

REFERENCES

Allen, W.D. 2002. "Crime, Punishment, and Recidivism: Lessons From the National Hockey League." *Journal of Sport Economics* 3(1): 39-60.

Becker, G.S. 1968. "Crime and punishment: An economic approach." *Journal of Political Economy*, 76:

169-217.

Campbell, D.T. and Ross, H.L. 1968. "The Connecticut Crackdown on speeding: Time series data in Quasi-experimental analysis." *Law & Society Review* 3: 33-53.

Chiappori, P.-A., Levitt, S. and Groseclose, T. 2002. "A Test of Mixed Strategy Equilibria: Penalty Kicks in Soccer." *American Economic Review* 92: 1138-1151.

Crawford, S.E.S. and Ostrom, E. 1995. "A Grammar of Institutions." *American Political Science Review* 89(3): 582-600.

Cressman, R., Morrison, W.G. and Wen, F-F 1998. "On the evolutionary dynamics of crime." *Canadian Journal of Economics* 31(5): 1101-1117.

Depken II, C.A. and Wilson, D.P. 2004. "Wherein Lies the benefits of the second referee in the NHL." *Review of Industrial Organization* **24**(1): 51-72.

Di Tella R, and Schargrodsky, E. 2004. Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack, *American Economic Review* 94 (1): 115-133.

Heckelman J.C. and Yates, A.J. 2003. "And a Hockey Game Broke Out: Crime and Punishment in the NHL." *Economic Inquiry* 41(4): 705-712.

Idson, T.L. and Kahane, L.H. 2000. "Team effects on compensation: an application to salary determination in the national hockey league." *Economic Inquiry* 38(2): 345-357.

Levitt, S.D. 2002. "Testing the Economic Model of Crime: The National Hockey League's Two-Referee Experiment." *Contributions to Economic Policy and Analysis* (1/1), article 2.

Marvell, T.A., and Moody, C.E. 1996. Specification problems, police levels and crime rates, *Criminology* 34(4): 609-646.

McCormick, R.E. and Tollison, D.D. 1984. "Crime on the Court." *Journal of Political Economy* 92(2): 223-235.

National Hockey League. 2003. *The National Hockey League Official Guide and Record Book 2004*. Chicago: Triumph Books.

National Hockey League. 2002. *The National Hockey League Official Guide and Record Book 2003*. Chicago: Triumph Books.

National Hockey League. 2001. *The National Hockey League Official Guide and Record Book 2002*. Chicago: Triumph Books.

National Hockey League. 2000. *The National Hockey League Official Guide and Record Book 2001*. Chicago: Triumph Books.

National Hockey League. 1999. *The National Hockey League Official Guide and Record Book 2000*. Chicago: Triumph Books.

Trandel, G.A. 2004. Hit by Pitches: Moral Hazard, Cost Benefit, Retaliation, or Lack of Evidence? *Journal of Sports Economics* 5: 87-92.

FIGURES AND TABLES

Figure 1: Average penalties in minutes (PIM) per team per game for 9 seasons, where the average is depicted as solid dots, and the dotted lines show one standard deviation around the average PIM per team.

Figure 2. Sequential Goon Game with Imperfect Legal Enforcement. The first payoff entry is for the Maple Leafs and the second is for the Red Wings. q > p.

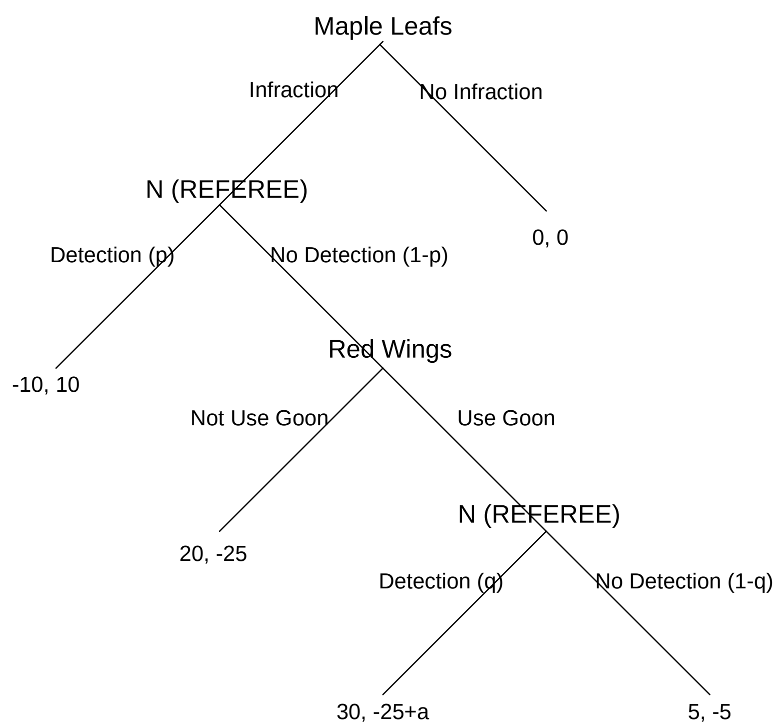
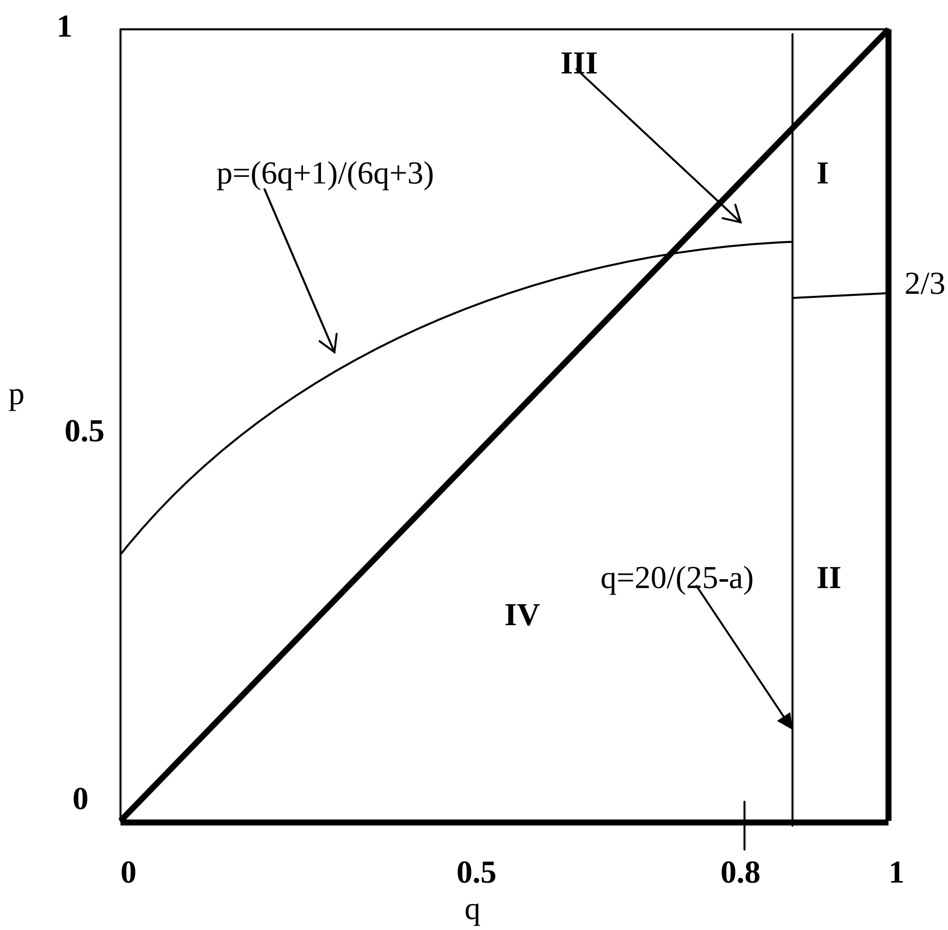


Figure 3. Equilibrium Zones of the Goon Game.



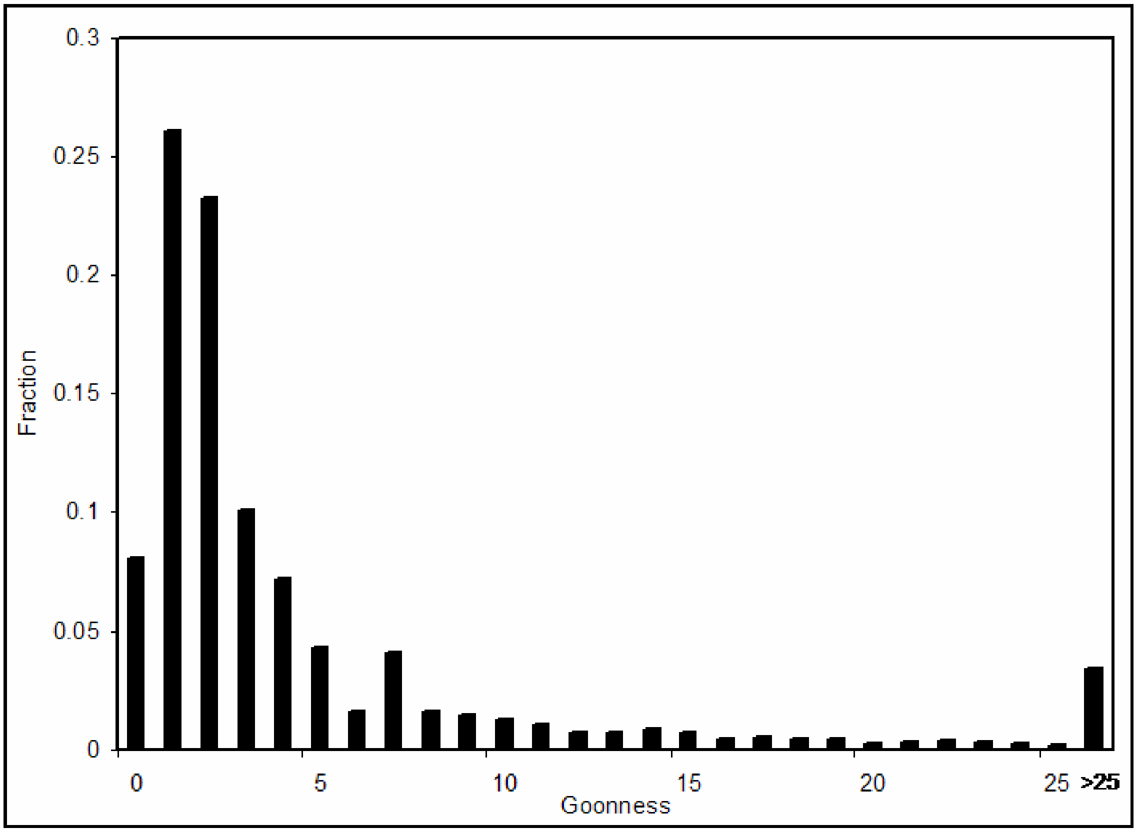
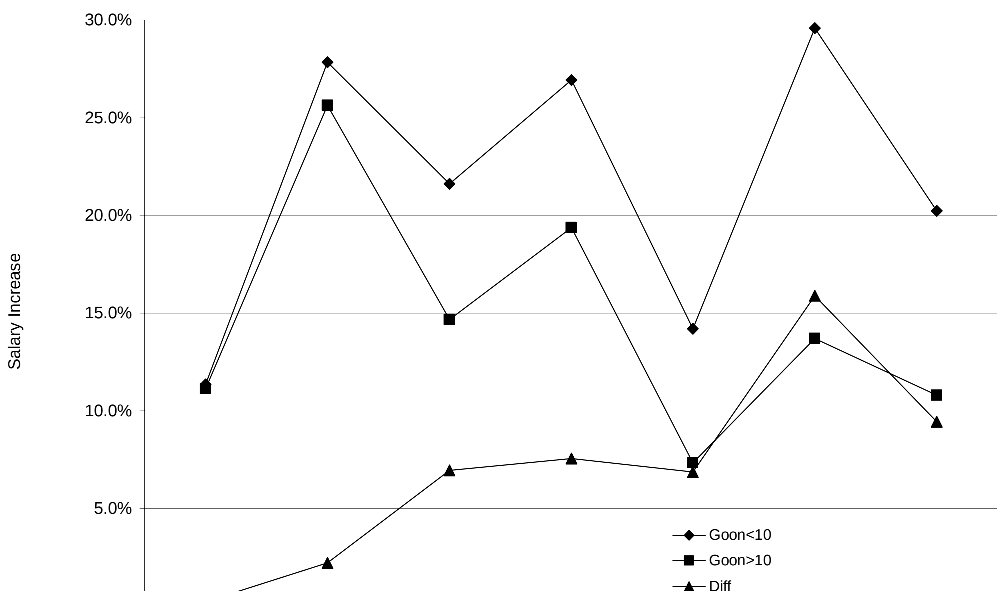


Figure 4. Distribution of Goonness – pooled across seasons



	1995-6	1996-7	1997-8	1998-9	1999-0	2000-1	2001-2
Goon<10	11.3%	27.8%	21.6%	26.9%	14.2%	29.6%	20.2%
Goon>10	11.1%	25.6%	14.7%	19.4%	7.3%	13.7%	10.8%
Diff	0.2%	2.2%	7.0%	7.5%	6.9%	15.9%	9.4%

Season

Figure 5. Goonness and Salary Change

Table 1. Payoffs of the Goon Game

	Immediate Game		Future Games	
	Maple Leafs	Red Wings	Maple Leafs	Red Wings
Maple Leaf Cheats	20	-20	0	0
Maple Leaf Does Not Cheat	0	0	0	0
Maple Leaf is Caught	-30	30	0	0
Red Wing Goon Retaliates	-10	10	-5	5
Goon is Caught	25	-25	0	5+a
Red Wings Do Not Retaliate	0	0	0	-5

Table 2. Equilibria of the Goon Game

	q	P	Equilibrium	Outcome
I	q>20/(25-a)	p>2/3	Maple Leafs Do Not Cheat Red Wings Do Not Use Goon	Maple Leafs Do Not Cheat
II	q>20/(25-a)	P<2/3	Maple Leafs Cheat Red Wings Do Not Use Goon	Maple Leafs Cheat Red Wings Do Not Use Goon
III	q<20/(25-a)	p>(6q+1)/(6q+3)	Maple Leafs Do Not Cheat Red Wings Use Goon	Maple Leafs Do Not Cheat
IV	q<20/(25-a)	P<(6q+1)/(6q+3)	Maple Leafs Cheat Red Wings Use Goon	Maple Leafs Cheat Red Wings Use Goon

Table 3. Average Goonness of Players over Before, During, and After the rule change

	Season	Average goonness	
Before the rule change	1995-6	4.84 (692)	5.30 (2139)
	1996-7	4.99 (714)	
	1997-8	6.03 (733)	
During the rule change	1998-9	4.81 (780)	4.27 (1617)
	1999-2000	3.78 (837)	
After the rule change	2000-1	4.10 (880)	4.24 (2644)
	2001-2	4.26 (879)	
	2002-3	4.36 (888)	

Table 4. Annual increase in salary by player position before and after the rule change
Defenders

	Goonness >=10	Goonness<10	Significance
Defenders			
Until 1998-1999	24.3%(88)	22.6% (442)	*p* = 0.37
After 1999-2000	12.6% (95)	20.3% (485)	*p* = 0.46
Centers			
Until 1998-1999	8.5% (10)	23.1% (327)	*p* = 0.50
After 1999-2000	4.9 (6)	25.0% (376)	*p* = 0.49
Wings			
Until 1998-1999	13.9% (90)	21.7% (548)	*p* = 0.92
After 1999-2000	8.9% (69)	19.7% (580)	*p* = 0.03*

Table 5: Goonness and Icetime Change*.

ICETIME CHANGE	Coefficient	Std error	*t*
GOONNESS	0.032	0.016	2.00
AFTERxGOONNESS	-0.035	0.019	-1.86
AFTER	-0.020	0.178	-0.11
Constant	0.059	0.156	0.38
N	2788		
R-squared	0.0019		

[i] The authors thank Elinor Ostrom and Roger Parks for helpful discussions on earlier versions of this paper. Financial support from the Resilience Alliance is appreciated.

[ii] These estimates are judgments of the authors. Two of the authors have been actively been involved in hockey as players, as well as observers of the NHL over many years.

[iii] For example, the 2002-2003 regular season witnessed 39 penalty shots, of which 9 went in.
http://www.sportschronicles.com/stats/nhl_penalty_shots.asp (accessed:8/25/03)

[iv] http://www.nhl.com/ [accessed:8-9-03] and http://www.hockeydb.com/ [accessed:8-9-03].

[v] http://moo.hawaii.edu:1749/hockey/misc/salary.html (season:1995-1996 [accessed: 8-9-03]), http://users.pullman.com/rodfort/PHSportsEcon/Common/OtherData/NHLSalaries/NHLSalaries.html (seasons:1996-1997, 1997-1998, 1998-1999 [accessed: 8-9-03]), http://www.seattleinsider.com/shared/sports/nhl/stats/salaries.html (season:1999-2000 [accessed: 7-20-02]), http://www.cris.com/~khallowe/powerplay/reports/reports.html (season:2000-2001 [accessed:7-20-

02]), http://www.faceoff.com/nhl/salaries/ (season:2001-2002 [accessed:7-20-02]),

http://www.hockeyzoneplus.com/$salai_e.htm (season:2002-2003 [accessed:8-9-03]).

[vi] http://cbs.sportsline.com/u/hockey/nhl/enforcers/ [accessed: 8-8-03].