Examining Regulatory Capture with High Frequency Data

Gregory DeAngelo* West Virginia University

Adam Nowak† West Virginia University

Imke Reimers‡ Northeastern University

December 31, 2015

Abstract

Regulatory capture has garnered significant attention, but poses a difficult empirical exercise since most relationships between regulators and regulated parties occur behind closed doors. In this research we overcome this problem by analyzing an environment where the behavior of both the regulator and regulated parties are publicly available. Specifically, we utilize data from the *National Hockey League (NHL)* to examine the impact of general experience as a referee as well as experience refereeing a particular team on the assignation of penalties. We find that gaining general experience as a referee significantly reduces the number of penalties that a referee assigns. However, as a referee gains experience refereeing a specific team, they significantly reduce the number of penalties assessed to this team relative to teams that they have less experience refereeing, confirming that regulatory capture is observed amongst referees and teams in the *NHL*.

JEL Codes: G18; G28; L51; K42
Key Words: regulatory capture, general experience, enforcement

*Department of Economics, College of Business and Economics, PO Box 6025, Morgantown WV 26506-6025; gregory.deangelo@mail.wvu.edu

†Department of Economics, College of Business and Economics, PO Box 6025, Morgantown WV 26506-6025; email: adam.nowak@mail.wvu.edu;

‡Department of Economics, Northeastern University, 304 Lake Hall, 360 Huntington Ave. Boston, MA 02115; email: i.reimers@neu.edu

1 Introduction

To (re-)calibrate incentives and conduct, regulatory controls are enacted by a governing body. To ensure that compliance is being achieved, a regulator is charged with overseeing the actions of the regulated party. More often than not this results in a relationship between the regulatory authority and the firms or individuals that are regulated. This is the case whether we look at financial institutions, firms attempting mergers, or criminals and the police. A continuing relationship could be beneficial to the efficiency of any regulatory activity because the regulator can acquire new information about the regulated party, which can be be incorporated into revised regulatory policies (see Baron and Besanko (1984)). On the other hand, a continuing relationship could lead to regulators becoming biased in their assessment of a situation. For example, there is often concern that bank regulators could be "schmoozed" by owners or managers of the bank in return for less stringent regulation. This is highlighted in Dal Bó (2006). Whether this concern of regulatory capture - regulatory agencies acting in the interest of special groups rather than in the public interest - is indeed valid is an empirical question.

Due to significant difficulties in observing the relationships between regulators and market participants over time, it is often difficult to produce any verifiable evidence of regulatory capture.[1] Attempts to determine the amount of regulatory capture have nevertheless been made. Hilton (1972), Leaver (2009) and DeAngelo and McCannon (2015) examine the role of squawking, finding that the likely objective of regulators is to minimize complaints on the part of those being regulated. Ades and Tella (1997), Dal Bó and Tella (2003), and Dal Bó and Rossi (2004) explore the role of market competition and trade conditions in the production of corruption. A healthy literature also investigates utility companies and regulators to determine a relationship between utility prices and regulatory pressure in several different forms - consumer advocacy groups, executive appointments with stipulations, legislative pressure, and the fear of court intervention (see Boyes and McDowell (1989), Smart (1994), and Spiller and Tiller (1997)). While the empirical research on regulatory capture has grown in recent years, much of this work relies on institutional systems to differentiate potential differences in observed outcomes, which are aggregated. Thus, the empirical research on regulatory capture is ripe for detailed individual-level behavior amongst regulators and regulated parties.

We empirically estimate the presence and extent of regulatory capture using a novel data set:

[1] Some of the interactions between regulators and regulated parties are through verbal (or perhaps non-verbal) communication behind closed doors, making observability in a meaningful way impossible.

penalties in the *NHL*. While there certainly are rules that govern the referees' actions in a game, considerable discretion remains. This is especially true in the assignation of more egregious offenses, such as targeting/illegal check to the head, fisticuffs (fighting), elbowing, checking from behind, clipping, etc.[2] Given that considerable discretion exists when referees are assigning penalties, it is possible that interpersonal relationships could guide the decisions of the referees.[3] Thus, as is seen in more traditional regulatory settings, it is possible that referees are more stringent with some players than they are with others.

We exploit two variations in our data that permit us to disentangle the impact of relationships and experience on the presence of regulatory capture. First, referees have varying experience performing the tasks associated with the role of refereeing. Some referees could have recently joined the ranks of professional hockey referee while other referees have been in the league for a considerable amount of time. The introduction of new referees permits us to observe how experienced regulators compare with inexperienced regulators - both with certain teams and more generally in any regulatory setting. Second, referees have varying experience refereeing specific teams. For example, approximately half of a referee's responsibilities will involve the referee overseeing one division (comprised of 7-8 teams), resulting in repeated interactions between specific teams and referees.

Previous studies have also investigated the role of NHL referees as monitors who detect crime. Levitt (2002) and Heckelman and Yates (2003) investigate the impact of using either one or two referees per game using data from the 1999-200 season. We expand their data significantly by using all games from 1996 to the present. While these studies (as well as ours) do not find a significant effect of the *number* of referees in a game, we find that the referees' *experience* does significantly impact the number of penalties assessed. In particular, every season of referee experience decreases the expected penalty time in a game by over 30 seconds. This coefficient takes on greater meaning considering that the average referee remains in the NHL for more than 10 years.

In the next section we review the relevant literature on regulatory capture. We then describe the data to be used in our analysis and the methods used to collect this data in Section 3. We report the results of our estimation in Section 4. Finally, we conclude by discussing our results

[2]For example, the NHL rulebook with regards to fisticuffs notes that "A minor (roughing), major or a major and a game misconduct, at the discretion of the Referee, shall be imposed on any player involved in fisticuffs." Similar wording is present in the other penalties as well.

[3]While other research has examined the behavior of referees in sports (seeSutter and Kocher (2004), Buraimo et al. (2009), Garicano et al. (2005), and Page and Page (2010)), the present research focuses on interpersonal relationships amongst participating actors, rather than the role of external, social pressure (e.g. home team bias) on the quality of refereeing.

with particular attention to their external validity in Section 5.

2 Literature Review

The idea of viewing referees as monitors or regulators is not new, especially in the economics of crime literature. Previous studies have investigated the idea that more monitors will decrease crime. McCormick and Tollison (1984) examine the impact of the number of monitors in college basketball when games are refereed by either two or three referees; Hutchinson and Yates (2007) confirm this result using corrected data. Levitt (2002) and Heckelman and Yates (2003) examine variation in the number of referees in NHL games (assuming that referees are homogeneous) and find mild evidence that increased monitoring will decrease crime. However, none of these studies investigate the behavior of individual referees over time.

Variation in penalties called is related to the supply and demand for such monitoring. Stigler (1971) provides one of the earliest examples of an economic analysis into the determinants of the supply and demand for regulation. In particular, he emphasizes that the supply of regulation is not necessarily provided by an altruistic regulator. In a survey on regulation, Dal Bó (2006) even goes so far as to say that the theory in Stigler (1971) views politicians as pursuing a selfish objective. In Peltzman (1976), regulation arises when politicians trade off benefits to producers and consumers. Surprisingly, regulation does not arise from any normative analysis of consumer welfare. The takeaway in both Stigler (1971) and Peltzman (1976) is that regulators should be viewed as economic agents who take actions in order to maximize their utility.

As mentioned above, monitors cannot always be viewed as altruistic. Boot and Thakor (1993) has shown that inept bank regulators will manipulate the decision to close banks when the regulator is unable to effectively monitor the bank. Regulatory capture can develop when relationships form between the regulator and the regulated. When a producer has unknown marginal costs, Baron and Myerson (1982) show that the regulator will necessarily allow the producer to collect some rents. However, Tirole (1986); Laffont and Tirole (1993) find that monitors who can learn the true marginal cost can be bribed into reporting higher costs.

Given these results, it is not surprising to find that referees do not always act as an ideal regulator or monitor. Previous research has focused on the idea of referee bias. For instance, Dawson et al. (2007) examine the English premiere leagues and find referees are biased in favor of the home team. Torgler (2004) finds evidence that native language plays a role in the actions of soccer referees at the world cup. Webster et al. (2012); Frank and Gilovich (1988) find that

NFL teams with darker colors are more likely to receive penalties. Abrevaya and McCulloch (2014) document that NHL referees tend to make *reverse calls* in order to even out the number of penalties on both teams.

3 Data Description

Penalty calls result in the player being removed from the playing surface for two (minor), four (double minor), five (major), or ten minutes (misconduct), typically while their team competes with a one person deficit. The length of the penalties follows strict rules conditional on the penalty that has been assessed, whereas calling the penalty is at the referee's discretion. For example, whether a check results in a penalty depends crucially on the timing of the hit and the position of the puck and the player. A player can check a player that recently maintained possession over the puck, before passing the puck. How much time is *too long* between the player passing the puck and being checked is not a crystallized notion, however, and can vary from one referee (or player) to another. As another example, contact between a non-goalkeeper and goalkeeper is often impermissible and results in a penalty. However, several exceptions to this rule exist and significant discretion is placed in the hands of the referee.[4] Thus, there is demonstrated discretion in the application of rules and it is this discretionary power that we exploit in measuring the extent of regulatory capture in our data.

We collect data on referees and penalties for all regular season NHL games from January 1, 1996 to December 11, 2015 using the NHL box scores on *espn.com*.[5] Using game level data, we are able to count the total number of games that referee n has appeared in as of time t. We define this variable as *total.games*. For example, Wes McCauley has refereed 100 games as of January 1, 2001. We then divide *total.games* by 82 (the number of games in a regular season) in order to create

[4]Rule 78 of the NHL Rule Book discusses contact between an attacking player and the goalkeeper. Subsection (c) offers considerable insight into the role of discretion by the referee in noting that "In all cases in which an attacking player initiates other than incidental contact with a goalkeeper, whether or not the goalkeeper is inside or outside the goal crease, and whether or not a goal is scored, the offensive player will receive a penalty (minor or major, as the Referee deems appropriate)." However, three exceptions to this rule exist: (1) In exercising his judgment under subsections (a) and (b) above, the Referee should give more significant consideration to the degree and nature of the contact with the goalkeeper than to the exact location of the goalkeeper at the time of the contact, (2) If an attacking player has been pushed, shoved, or fouled by a defending player so as to cause him to come into contact with the goalkeeper, such contact will not be deemed to be contact initiated by the attacking player for purposes of this rule, provided the attacking player has made a reasonable effort to avoid such contact, and (3) A goalkeeper is not "fair game" just because he is outside the goal crease. The appropriate penalty should be assessed in every case where an attacking player makes unnecessary contact with the goalkeeper. However, incidental contact will be permitted when the goalkeeper is in the act of playing the puck outside his goal crease provided the attacking player has made a reasonable effort to avoid such unnecessary contact

[5]We omit NHL playoff games from the analysis as there is a possibility that the nature of penalties in regular and post-season games are different.

our measure of the referee's experience in seasons, *total.seasons*. We report our results as effects of seasons rather than games of experience to facilitate the interpretation of the coefficients. Of course, some referees may have started officiating NHL games before 1996. In order to mitigate the effect of left censoring of our measure of referee experience, we calculate the total games refereed beginning January 1, 1996 but restrict our analysis to games beginning four years later, on January 1, 2000.

Furthermore, the box scores do not specify which referee called a specific penalty. We therefore focus on the total number of penalty minutes called in a game. More importantly for the interpretation of our results, we look at the total amount of referee experience, rather than each individual referee's experience. That is, $total.seasons = total.seasons_1 + total.seasons_2$ for referees 1 and 2 in a game. If two referees each have a single season of experience, *total.seasons* would be equal to 2.

Table 1 shows descriptive statistics of the variables of interest for all games beginning January 1, 2000. Penalties in minutes, *pims*, is the dependent variable of interest. Figure 1 shows a histogram of the *pims* per game in the sample. In order to remove the effect of any outliers, we only use those games with less than 100 *pims*; this filter removes 16 games from our data set. On average, referees call 14.3 minutes of penalties in a game, from an average of about 5.75 penalties. There is significant variation in this number, however, with some games having no penalties at all, and others having penalties assessed during the majority of the game. The average number of total games each referee pair has in the sample is 832.1 total games, or just over ten years of experience per referee team. The total number of refereed games also shows considerable variation, with some referee pairings being complete rookies while other pairings have over thirty seasons of combined experience.

In addition to game level analysis, we also focus on the individual teams in each game and analyze if experience with a specific team influences *pims*. Table 2 displays summary statistics for the home and away teams in each game.[6] The variable *pims.away* (*home.away*) is the number of *pims* for the away team. *home.games* (*away.games*) is the sum of the total number of games both referees have with the away team including the current game, and *home.seasons* (*away.seasons*) is the number of games of experience scaled by 82. Table 2 shows that some referees have a considerable amount of experience with specific teams in part due to their overall experience as

[6]Because team data and game data were pulled from different tables in the boxscores, there are some missing values in the team values that are not missing in the game values. Therefore, the number of observations in Table 2 is less than the number of observations in Table 1.

well as logistic factors that limit where and when a referee can be.

There is another interesting source of variation in the regulatory setting of NHL hockey games. During the 1998-1999 season, the NHL decided to increase the number of referees in a game from one to two. This change was gradual, with some games still being officiated by just one referee early in the dataset. Figure 2 shows the percentage of games with two referees by season. In our data set, 89.7% of games use a two referee system, and the remaining games have just one referee. We control for the number of referees officiating a game by including the indicator $two.ref$ in the estimation, even though previous research (see Levitt (2002)) did not find a significant effect of the number of referees on the probability of penalties being called.

With data covering a large span of time, we are able to estimate the impact of experience on the number of penalties that a referee calls. Figure 3 shows the average $pims$ by season. It is apparent from the figure that the average number of penalties per game is decreasing. We suggest one reason for this decrease is the regulatory capture of referees. Figure 4 displays the average number of penalty minutes called per game for all referees based on the number of seasons of experience in the data. In order to control for the left-censoring in the data, only those referees who first show up after the 1996 season are used. Figure 4 shows a clear decline in the number of penalties called per game as the referee gains more experience. First-year referees call as many as 24 penalty minutes per game, which corresponds to almost ten penalties on average. After one year of experience, the penalty minutes decrease to about 18, and even after that, there is a steady decrease in penalty minutes as referees gain more experience. Referee teams with 18 years of experience call less than half as many penalties as referees with zero years of experience. Whether these differences are due to regulatory capture remains to be determined in the estimation.

Finally, we address the fact that the two referees officiating a game can vary significantly in their experience. It could be that one referee has just joined the league while the other has officiated games for several years. A game that is called by two seasoned referees is likely called differently than one that has one rookie referee. We address this by looking at the experience level of the less experienced referee. The bottom panel of Table 1 shows the distribution of the "younger" referee's experience. For example, $game40$ is equal to 1 if the less experienced referee has 40 or fewer games of experience. The variable $game1$ is equal to 1 if the less experienced referee is making his debut. The table shows that most games are officiated by two relatively experienced referees. Only 9.6% of games feature at least 1 referee who has 40 or fewer games of experience.

As we are interested in the effect of the regulators' experience and relationship with players,

we segment the data by the referees' experience level and report the descriptive statistics for our variables of interest in Table 3. When both referees have fewer than 20 games of experience (panel A), the number of penalty minutes is significantly larger than when only one referee is relatively inexperienced (panel B) as the t-statistic for the difference in means is larger than 12. When both referees are experienced (panel C), the number of penalty minutes drops even further. This offers preliminary evidence that referees who have little experience are more likely to call penalties. Table 3 also shows that our referees become more experienced over time, as the games with two inexperienced referees happened earlier than those with two very experienced referees.[7] This is most likely a result of a general move towards referees being more experienced in the league, since we do include a four-year lead time in our referee collection. Still, an analysis of the presence of regulatory capture needs to take into account overall time trends. We do this in the following section.

4 Results

The analysis in this paper includes different specifications of regressions on the game level. Our outcome variable of interest is the penalty minutes called in a game. Since referees have some discretion over which penalties to (not) assess, penalty minutes can be used as a measure of regulatory capture.[8] We start our analysis by determining whether those referees who have previously interacted with a team tend to permit more infringing actions than referees who have not yet had many interactions with a specific team. We account for a referee's experience as a referee by using the referees' overall tenure in the *NHL*, which should capture the referee's experience performing the tasks associated with the job. We further control for the number of referees in the game, and we include year, month, and referee fixed effects in our regressions. In a later regression, we add to this the referees' experience with the teams on the ice, assuming that a referee who has worked with a certain team more frequently might be more susceptible to regulatory capture.

Following the empirical specification in DeAngelo and Owens (2015), Table 4 displays the results of these regressions. Column 1 is a base model where penalty minutes are regressed on the total number of seasons that the referee team has called. The results indicate a statistically significant relationship where referees with more experience are less likely to call penalties. For each additional

[7] On average games with two inexperienced referees were called in the year 2002.7, compared to 2007.6 when both referees were experienced.

[8] DeAngelo et al. (2015) examines the relationship between penalties and fighting in the *NHL*, finding that fights are the response to an absence of appropriate refereeing

season of experience, the expected number of penalty minutes decreases by 0.473. Given the longevity of some referees, this number is economically significant as well: referee teams with the mean level of experience (10 years) call five fewer penalty minutes per game than rookie referee teams.[9]

We also control for the number of referees officiating the game. The positive and significant coefficient on the indicator that the game has two referees suggests that games that utilize a two referee system have 1.531 more penalty minutes than games with only a single referee.[10] This value is larger than the values in Levitt (2002) and Heckelman and Yates (2003), who examine the effect of a second referee on penalties in the 1999-2000 season. In unreported results, we find that referees who call games in tandem have significantly less experience than those who call games alone. We interpret these results as the NHL pairing more experienced referees with less experienced referees, a feature which would explain the small values found in Levitt (2002).

Of course, column 1 does not account for idiosyncrasies across years or across referees. To address this, we include year and month fixed effects in column 2, and we further include referee fixed effects in column 3.[11] In both cases, the impact of referee experience diminishes but remains significant, at least at the five percent level. Still, one more year of referee experience decreases the number of penalty minutes per game by 0.995.

The unreported year and month fixed effects indicate a gradual decrease of penalty minutes over time. On an interesting note, the only month with a significant coefficient was October. Typically, the NHL season begins during October, which could affect the number of penalties called in three ways. First, games early in the season could be seen as less important since the playoffs are in the distant future. As a result, players may refrain from being overly tough on their opponents, leading to fewer penalties. Second, at the beginning of the season teams could be uncoordinated and not coalescing, therefore committing more penalty-worthy infractions - especially those that result from a lack of focus. Third, and most interestingly with regard to this paper, players and referees likely have not seen each other, and certainly have not been involved in the same games, for several months, so that the relationships between the referees (regulators) and the players had

[9]Not surprisingly, penalties are associated with fewer goals and a decreased probability of winning. Of the 16,731 games in our sample, the team with fewer penalties won 64.8% of the time. This difference is statistically significant as the t-statistic is above 18.

[10]Of course, this could be a reverse causality issue, which is common in the economics of crime literature whereby locations with greater police intensity have higher crime rates - see DeAngelo and Hansen (2014).

[11]For games where there is only a single referee, the indicator variable for that referee is equal to 1. For games with two referees, both indicator variables for each referee are equal to 1. In unreported results, the conclusions are not altered if we set the value of the indicator variable equal to 0.5 for each referee in two referee games.

time to "deteriorate" as the referees forget previous relationships with players. Our coefficients on the month of October are large and significantly positive throughout, suggesting that perhaps the relationships with players indeed play a role in determining how many penalties are assessed.

Figure 4 suggests that the effect of experience on the number of penalties called may not be linear. Columns 4 and 5 of Table 4 experiment with quadratic and log forms of *total.seasons*. Consistent with Figure 4, column 4 indicates that the expected number of penalty minutes is decreasing in the number of games called by the referees, but that this effect is diminishing. Solving for how long it takes until the expected penalty minutes begin to increase, we find that a referee team would need 21.75 seasons of experience before they would begin to call more penalties. In our data, no referees have this much game experience. Therefore, we conclude that the effect of referee experience on expected penalty minutes is decreasing when accounting for quadratic effects. Column 5 uses the natural log of total seasons called as the independent variable of interest and reaches similar conclusions.

Finally, column 6 includes the sum of total seasons for the referees as well as the product of the total season for the two referees, defined as $cross.total.season = total.season_1 \times total.season_2$. The estimate of this coefficient is positive while the coefficient on *total.seasons* is negative. Thus, we find that the number of penalties called is decreasing in referee experience, but that this effect is smaller when two veteran referees are paired together, rather than one very experienced referee and a rookie referee.

The results in Table 4 show that the experience of both referees plays an important role in determining the number of penalties called in a game, but it is still unclear whether the effect is driven by the referee pairing, or by the individual referees that are put together on the ice. In order to determine if the above results are driven by non-linear effects for referees with little to no experience, we use the variables $game1$ to $game40$ as explanatory variables in addition to the total number of years that the referees have been in the league. Recall that $game40$ is an indicator that is turned on if the less experienced referee has called less than 40 games prior to the game in question, and so on. This implies that the interpretation of the coefficients is cumulative. Due to collinearity issues, we only include year fixed effects.

The regression results are displayed in Table 5. Column 1 indicates that if a game has at least 1 ref with fewer than 40 games of experience, the expected number of penalty minutes increases by 2.143. The coefficient on $two.ref$ remains significant and positive, while the coefficient on *total.seasons* is smaller than the coefficient on *total.seasons* in Column 1 of Table 4. Thus, it

appears that inexperience is driving more penalty minutes and that this will diminish over time.

The results remain robust as we add more information about the individual referees. Columns 2-5 indicate the cumulative effects of the experience of the less experienced referee. Column 2 indicates that if the less experienced referee has 20 or fewer games, the expected number of penalties increases by 3.643. Column 2 also indicates that there is no difference in expected penalty minutes between games where the less experienced referee has 21 to 40 games of experience and games where he or she has more than 40 games of experience. Thus, it appears as though 20 games is an appropriate cutoff between labeling referees *experienced* or *not experienced*.

Columns 3-5 show the remaining effects when including $game15$ to $game1$. The results indicate that there is a significant change in expected penalty minutes at the breakpoints of 10 and 20 games. Using results in Column 5, games where the less experienced referee has 20 or fewer games of experience have 3.130 more expected penalty minutes, and games where the less experienced referee has 10 or fewer games of experience have $3.130 + 3.068 = 6.198$ more expected penalty minutes. In all situations, the number of referees officiating the game is statistically significant, indicating that two referees assess approximately 1 penalty minute more than a single referee.[12]

As a final example of regulatory capture in the data, we examine the penalties for each team in each game and use team-referee experience. By doing so, we double the total number of observations as each game has two teams. For each team and referee, we calculate the total games of experience, *team.seasons* and scale by 83; thus, *team.seasons* is *total.seasons* but for each unique team-referee pair. For instance, Wes McCauley has refereed 10 games involving the Detroit Red Wings. We add *team.seasons* for both referees if the game has two referees. The results, when expanding the analysis for each team and including *team.seasons*, are displayed in Table 7. The results indicate that the individual team experience plays a large role in determining *pims*. Column 2 indicates that each additional season of experience with a given team decreases *pims* by 0.922 whereas an additional season of overall game experience will decrease *pims* by 0.061. These results remain roughly consistent across different specifications, showing that referee experience with a specific team significantly reduces the number of penalties assigned to that team, by almost one penalty minute per season.[13]

Overall, our results show that those referees who have spent more time on the ice with specific

[12]It is possible that some of the results could be driven by other structural differences between one and two referees systems. In order to address this issue, table 6 reports the same results when using only those games with two referees.

[13]We also look at the total experience of each referee-pair. For example, how many games have Dave Jackson and Wes McCauley done together. The results are in Table 8. The results are similar to the results in other tables.

teams systematically call fewer penalties against that team than they do against teams that they have relatively less experience officiating. The development of relationships with the teams appears to play a significant role in determining the number of penalties assessed, which persists over time. The interpretation of these results as an instance of regulatory capture is supported by the fact that referees call more penalties early in the season - when they have not interacted with the players for several months prior to the games. Most importantly, the presence of regulatory capture tips the scales in favor of the less penalized team, as those teams that are assessed fewer penalties have a higher likelihood of winning the game.

5 Conclusion

Concern about the relationship between regulators and regulated parties has garnered significant concern from several government enforcement agencies in the fields of banking regulations, antitrust cases, and criminal justice. The concern, of course, is that the interactions between regulated parties and regulators could be complicated by attempts to evade regulation by asking favors (e.g. not reporting oversights), which might come with a form of compensation. Unfortunately, observing these relationships in sufficient detail is not usually possible.

Turning to an observable relationship, we examine the relationship between professional athletes and referees using data from the NHL. Referees are entrusted with a significant amount of discretion that they use in discerning whether a violation of the rules has occurred. In an ideal world, referees behave as if they are automatons and simply strictly interpret the rules while applying them to players on the ice. This is, however, short sighted, as players and referees likely form relationships through repeated interactions. We leverage these relationships by making use of the fact that both new players and referees enter the league, thereby allowing us to determine how more experienced referees behave relative to rookie referees. We find significant regulatory capture, as rookie referees call significantly more penalties than veteran referees.

This research contributes to a growing body of literature that is focused on regulatory capture. While our empirical exercise focuses on a sports environment, we believe that it offers behavioral expectations about the expected interactions between the regulating and regulated parties. Notably, we find significant evidence that regulator behavior is altered by experience in their role, but also with individual-level experience, which could be cause for concern. Future work would benefit from exploiting a regulatory environment where the regulator's behavior is less apparent to observers, as this would examine an environment where there is the greatest concern for regulatory capture.

References

Abrevaya, J. and McCulloch, R. (2014). Reversal of fortune: a statistical analysis of penalty calls in the national hockey league. *Journal of Quantitative Analysis in Sports*, 10(2):207–224.

Ades, A. and Tella, R. D. (1997). National champions and corruption: Some unpleasant interventionist arithmetic. *Economic Journal*, 107(443):1023–42.

Baron, D. P. and Besanko, D. (1984). Regulation and information in a continuing relationship. *Information Economics and Policy*, 1(3):267–302.

Baron, D. P. and Myerson, R. B. (1982). Regulating a monopolist with unknown costs. *Econometrica: Journal of the Econometric Society*, pages 911–930.

Boot, A. W. and Thakor, A. V. (1993). Self-interested bank regulation. *The American Economic Review*, pages 206–212.

Boyes, W. and McDowell, J. (1989). The selection of public utility commissioners: A reexamination of the importance of institutional setting. *Public Choice*, 61(1):1–13.

Buraimo, B., Forrest, D., and Simmons, R. (2009). The 12th man?: Refereeing bias in english and german soccer. *Journal of the Royal Statistical Society: Series A*, 173(2):431–449.

Dal Bó, E. (2006). Regulatory capture: a review. *Oxford Review of Economic Policy*, 22(2):203–225.

Dal Bó, E. and Rossi, M. (2004). Corruption and inefficiency: Theory and evidence from electric utilities. *University of California, Berkeley, mimeo*.

Dal Bó, E. and Tella, R. D. (2003). Capture by threat. *Journal of Political Economy*, 111(5):1123–54.

Dawson, P., Dobson, S., Goddard, J., and Wilson, J. (2007). Are football referees really biased and inconsistent?: evidence on the incidence of disciplinary sanction in the english premier league. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170(1):231–250.

DeAngelo, G. and Hansen, B. (2014). Life and death in the fast lane: Police enforcement and traffic fatalities. *American Economic Journal: Economic Policy*, 6(2):231–57.

DeAngelo, G., Humphreys, B. R., and Reimers, I. (2015). Examining law and order with community and specialized enforcement. *Working Paper*.

DeAngelo, G. and McCannon, B. (2015). Public outcry and police behavior. *BE Journal of Economic Analysis and Policy*, Forthcoming.

DeAngelo, G. and Owens, E. (2015). Learning the ropes: General experience, task-specific experience, and the output of police officers. *Working Paper*.

Frank, M. G. and Gilovich, T. (1988). The dark side of self-and social perception: black uniforms and aggression in professional sports. *Journal of personality and social psychology*, 54(1):74.

Garicano, L., Palacios-Huerta, I., and Prendergast, C. (2005). Favoritism under social pressure. *The Review of Economics and Statistics*, 87(2):208–216.

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.

Hilton, G. (1972). The basic behavior of regulatory commissions. *American Economic Review*, 62:47–54.

Hutchinson, K. P. and Yates, A. J. (2007). Crime on the court: A correction. *Journal of Political Economy*, 115(3):515–519.

Laffont, J.-J. and Tirole, J. (1993). *A theory of incentives in procurement and regulation*. MIT press.

Leaver, C. (2009). Bureaucratic minimal squawk behavior: Theory and evidence from regulatory agencies. *American Economic Review*, 99(3):572–607.

Levitt, S. D. (2002). Testing the economic model of crime: The national hockey league's two-referee experiment. *Contributions in Economic Analysis & Policy*, 1(1).

McCormick, R. E. and Tollison, R. D. (1984). Crime on the court. *The Journal of Political Economy*, pages 223–235.

Page, K. and Page, L. (2010). Alone against the crowd: Individual differences in referees ability to cope under pressure. *Journal of Economic Psychology*, 31(2):192–199.

Peltzman, S. (1976). Toward a more general theory of regulation.

Smart, S. (1994). The consequences of appointment methods and party control for telecommunications pricing. *Journal of Economics and Management Strategy*, 3(2):93–109.

Spiller, P. and Tiller, E. (1997). Decision costs and the strategic design of administrative process and judicial review. *Journal of Legal Studies*, 26(2):347–70.

Stigler, G. J. (1971). The theory of economic regulation. *The Bell journal of economics and management science*, pages 3–21.

Sutter, M. and Kocher, M. G. (2004). Favoritism of agents the case of referees home bias. *Journal of Economic Psychology*, 25(4):461–469.

Tirole, J. (1986). Hierarchies and bureaucracies: On the role of collusion in organizations. *Journal of Law, Economics, & Organization*, pages 181–214.

Torgler, B. (2004). The economics of the fifa football worldcup. *Kyklos*, 57(2):287–300.

Webster, G. D., Urland, G. R., and Correll, J. (2012). Can uniform color color aggression? quasi-experimental evidence from professional ice hockey. *Social Psychological and Personality Science*, 3(3):274–281.

Table 1: Descriptive Statics

Statistic	N	Mean	St. Dev.	Min	Max
year	16,731	2007.558	4.646	2000	2015
pims	16,731	14.283	10.582	0	98
total.games	16,731	832.077	428.081	2	2,410
total.seasons	16,731	10.147	5.221	0.024	29.390
two.ref	16,731	0.933	0.249	0	1
game1	16,731	0.003	0.052	0	1
game5	16,731	0.013	0.114	0	1
game10	16,731	0.027	0.161	0	1
game15	16,731	0.039	0.194	0	1
game20	16,731	0.051	0.221	0	1
game40	16,731	0.096	0.295	0	1

Table 2: Team Specific Experience

Statistic	N	Mean	St. Dev.	Min	Max
pims.away	16,554	13.377	8.544	0	76
pims.home	16,554	12.679	8.316	0	76
home.games	16,554	58.628	30.970	1	205
away.games	16,554	58.231	30.761	1	197
home.seasons	16,554	0.716	0.378	0.012	2.500
away.seasons	16,554	0.711	0.375	0.012	2.402

Table 3: Individual Referee Experience

PANEL A: All referees have fewer than 20 games of experience

	N	Mean	St. Dev.	Min	Max
year	55	2002.691	3.344	2000	2010
pims	55	25.455	18.074	4	89
total.games	55	12.509	6.691	2	32
two.ref	55	0.218	0.417	0	1

PANEL B: One referee has fewer than 20 games of experience

Statistic	N	Mean	St. Dev.	Min	Max
year	861	2006.436	5.385	2000	2015
pims	861	19.110	13.985	0	93
total.games	861	457.177	299.434	2	1,219
two.ref	861	0.950	0.218	0	1

PANEL C: Neither referee has fewer than 20 games of experience

Statistic	N	Mean	St. Dev.	Min	Max
year	15,870	2007.619	4.595	2000	2015
pims	15,870	14.021	10.302	0	98
total.games	15,870	852.416	424.615	21	2,410
two.ref	15,870	0.933	0.251	0	1

Table 4: Penalties in Minutes and Referee Game Experience

Table 4

	pims					
	(1)	(2)	(3)	(4)	(5)	(6)
two.ref	1.531***	1.008***	0.995***	1.709***	1.850***	1.441***
	(0.340)	(0.341)	(0.341)	(0.363)	(0.365)	(0.356)
total.seasons	−0.473***	−0.137***	−0.138***	−0.528***		−0.341***
	(0.016)	(0.022)	(0.022)	(0.071)		(0.052)
total.seasons.squared				0.016***		
				(0.003)		
total.seasons)					−1.483***	
					(0.166)	
cross.total.season						0.031***
						(0.007)
October10		6.437***	6.401***	6.291***	6.259***	6.339***
		(0.284)	(0.285)	(0.285)	(0.285)	(0.285)
Constant	17.655***	17.565***	16.408***	17.156***	17.057***	16.768***
	(0.314)	(0.402)	(0.597)	(0.610)	(0.603)	(0.602)
year and month fixed-effects		Y	Y	Y	Y	Y
referee fixed-effects			Y	Y	Y	Y
Observations	16,731	16,731	16,731	16,731	16,731	16,731
R^2	0.050	0.121	0.125	0.126	0.127	0.126
Adjusted R^2	0.050	0.119	0.120	0.121	0.122	0.120
Residual Std. Error	10.314 (df = 16728)	9.930 (df = 16708)	9.929 (df = 16632)	9.920 (df = 16631)	9.917 (df = 16632)	9.924 (df = 16631)
F Statistic	440.659*** (df = 2; 16728)	104.080*** (df = 22; 16708)	24.175*** (df = 98; 16632)	24.317*** (df = 99; 16631)	24.650*** (df = 98; 16632)	24.148*** (df = 99; 16631)

Note: *p<0.1; **p<0.05; ***p<0.01. Table 4 displays regression results when total penalty minutes per game are regressed on *total.seasons*, *total.seasons.squared* and *log.total.seasons*. *total.seasons* is an experience variable that is the sum of total game experience for each referee in a particular game, divided by 82. For each referee, total game experience at time t is calculated as the total number of games refereed since 1996. Because the data was collected beginning 1996, the regression results above use only games after December 31, 1999 in order to avoid a censoring problem.

Table 5: Effects of the Less Experienced Referee

Table 5

	pims				
	(1)	(2)	(3)	(4)	(5)
game40	2.143***	0.210	0.208	0.207	0.207
	(0.289)	(0.395)	(0.395)	(0.395)	(0.395)
game20		3.643***	3.130***	3.130***	3.130***
		(0.508)	(0.795)	(0.795)	(0.795)
game15			−1.139	−1.139	−1.139
			(0.997)	(0.997)	(0.997)
game10			2.669***	3.069***	3.068***
			(0.857)	(0.981)	(0.981)
game5				−0.807	−1.060
				(0.965)	(1.027)
game1					1.212
					(1.686)
two.ref	1.092***	1.076***	1.067***	1.071***	1.070***
	(0.351)	(0.351)	(0.350)	(0.351)	(0.351)
total.seasons	−0.147***	−0.143***	−0.143***	−0.144***	−0.144***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Constant	16.708***	16.676***	16.674***	16.683***	16.687***
	(0.591)	(0.590)	(0.590)	(0.590)	(0.590)
Observations	16,731	16,731	16,731	16,731	16,731
R^2	0.083	0.086	0.087	0.087	0.087
Adjusted R^2	0.078	0.081	0.081	0.081	0.081
Residual Std. Error	10.160 (df = 16636)	10.145 (df = 16635)	10.142 (df = 16633)	10.142 (df = 16632)	10.142 (df = 16631)
F Statistic	16.098*** (df = 94; 16636)	16.517*** (df = 95; 16635)	16.292*** (df = 97; 16633)	16.133*** (df = 98; 16632)	15.975*** (df = 99; 16631)
Year Fixed Effects	Y	Y	Y	Y	Y

Note: *p<0.1; **p<0.05; ***p<0.01. Table 4 displays regression results when total penalty minutes per game are regressed on *total.seasons* and *log.total.seasons* and indicators for game experience of the less experienced referee. *total.seasons* is an experience variable that is the sum of total game experience for each referee in a particular game, divided by 82. The remaining variables are equal to 1 if the referee has fewer than the stated experience; as such, the effects are cumulative. For each referee, total game experience at time t is calculated as the total number of games refereed since 1996. Because the data was collected beginning 1996, the regression results above use only games after December 31, 1999.

Table 6: Penalties in Minutes and Referee Game Experience: Two Refs

	Table 6					
	pims					
	(1)	(2)	(3)	(4)	(5)	(6)
total.seasons	-0.522^{***}	-0.129^{***}	-0.130^{***}	-0.511^{***}		-0.329^{***}
	(0.016)	(0.023)	(0.023)	(0.076)		(0.055)
total.seasons.squared				0.015^{***}		
				(0.003)		
total.seasons)					-1.574^{***}	
					(0.194)	
cross.total.season						0.030^{***}
						(0.007)
October		6.107^{***}	6.068^{***}	5.975^{***}	5.950^{***}	6.016^{***}
		(0.293)	(0.293)	(0.294)	(0.294)	(0.293)
Constant	19.878^{***}	30.012^{***}	29.145^{***}	30.022^{***}	30.203^{***}	29.615^{***}
	(0.187)	(1.047)	(1.107)	(1.119)	(1.119)	(1.113)
year and month fixed-effects		Y	Y	Y	Y	Y
referee fixed-effects			Y	Y	Y	Y
Observations	16,028	16,028	16,028	16,028	16,028	16,028
R^2	0.061	0.138	0.143	0.145	0.145	0.144
Adjusted R^2	0.061	0.137	0.138	0.139	0.139	0.138
Residual Std. Error	10.450 (df = 16026)	10.021 (df = 16004)	10.016 (df = 15926)	10.008 (df = 15925)	10.006 (df = 15926)	10.011 (df = 15925)
F Statistic	$1,047.869^{***}$ (df = 1; 16026)	111.418^{***} (df = 23; 16004)	26.333^{***} (df = 101; 15926)	26.389^{***} (df = 102; 15925)	26.707^{***} (df = 101; 15926)	26.255^{***} (df = 102; 15925)

Note: *p<0.1; **p<0.05; ***p<0.01. Table 4 displays regression results when total penalty minutes per game are regressed on *total.seasons*, *total.seasons.squared* and *log.total.seasons*. *total.seasons* is an experience variable that is the sum of total game experience for each referee in a particular game, divided by 82. For each referee, total game experience at time t is calculated as the total number of games refereed since 1996. Because the data was collected beginning 1996, the regression results above use only games after December 31, 1999 in order to avoid a censoring problem.

Table 7: Penalties in Minutes and Team Experience

Table 7

	pims					
	(1)	(2)	(3)	(4)	(5)	(6)
team.seasons	−1.983***	−0.922***	−0.900***	−0.897***	−0.917***	−0.897***
	(0.169)	(0.179)	(0.179)	(0.179)	(0.174)	(0.179)
two.ref	−0.171	0.009	0.006	0.011	0.014	0.011
	(0.196)	(0.200)	(0.200)	(0.200)	(0.200)	(0.200)
total.seasons	−0.192***	−0.061***	−0.063***	−0.107**		−0.107**
	(0.018)	(0.019)	(0.019)	(0.051)		(0.051)
total.seasons.squared				0.004		
				(0.004)		
log.total.seasons					−0.199***	
					(0.054)	
team.cross.total						0.004
						(0.004)
October		1.246***	1.242***	1.238***	1.228***	1.238***
		(0.184)	(0.184)	(0.184)	(0.184)	(0.184)
Constant	15.328***	13.668***	13.919***	13.989***	13.901***	13.989***
	(0.264)	(0.320)	(0.431)	(0.444)	(0.430)	(0.444)
year and month fixed-effects		Y	Y	Y	Y	Y
referee fixed-effects			Y	Y	Y	Y
Observations	33,108	33,108	33,108	33,108	33,108	33,108
R^2	0.023	0.048	0.052	0.052	0.052	0.052
Adjusted R^2	0.022	0.048	0.049	0.049	0.049	0.049
Residual Std. Error	8.343 (df = 33104)	8.234 (df = 33084)	8.229 (df = 33008)	8.229 (df = 33007)	8.229 (df = 33008)	8.229 (df = 33007)
F Statistic	254.136*** (df = 3; 33104)	73.206*** (df = 23; 33084)	18.164*** (df = 99; 33008)	17.991*** (df = 100; 33007)	18.195*** (df = 99; 33008)	17.991*** (df = 100; 33007)

Note: *p<0.1; **p<0.05; ***p<0.01. Table 4 displays regression results when total penalty minutes per game are regressed on *total.seasons*, *total.seasons.squared* and *log.total.seasons*. *total.seasons* is an experience variable that is the sum of total game experience for each referee in a particular game, divided by 82. For each referee, total game experience at time t is calculated as the total number of games refereed since 1996. Because the data was collected beginning 1996, the regression results above use only games after December 31, 1999 in order to avoid a censoring problem.

Table 8: Penalties in Minutes and Ref Experience For Ref Pairs

	Table 8					
	pims					
	(1)	(2)	(3)	(4)	(5)	(6)
two.ref	−3.835***	−0.766**	−0.779**	−0.480	−0.272	−0.764**
	(0.268)	(0.334)	(0.335)	(0.343)	(0.300)	(0.336)
total.seasons	−0.447***	−0.138***	−0.140***	−0.364***		−0.147***
	(0.015)	(0.023)	(0.023)	(0.062)		(0.028)
total.seasons.squared				0.010***		
				(0.003)		
log.total.seasons					−0.849***	
					(0.129)	
cross.total.season						0.006
						(0.013)
October		6.494***	6.456***	6.424***	6.455***	6.454***
		(0.283)	(0.284)	(0.284)	(0.284)	(0.284)
Constant	19.204***	18.202***	17.013***	17.679***	17.434***	17.032***
	(0.176)	(0.356)	(0.569)	(0.594)	(0.581)	(0.571)
year and month fixed-effects		Y	Y	Y	Y	Y
referee fixed-effects			Y	Y	Y	Y
Observations	16,731	16,731	16,731	16,731	16,731	
R^2	0.058	0.120	0.124	0.125	0.125	0.124
Adjusted R^2	0.058	0.119	0.119	0.120	0.120	0.119
Residual Std. Error	10.271 (df = 16728)	9.932 (df = 16708)	9.931 (df = 16632)	9.927 (df = 16631)	9.929 (df = 16632)	9.931 (df = 16631)
F Statistic	514.632*** (df = 2; 16728)	103.843*** (df = 22; 16708)	24.128*** (df = 98; 16632)	24.055*** (df = 99; 16631)	24.213*** (df = 98; 16632)	23.885*** (df = 99; 16631)

Note: *p<0.1; **p<0.05; ***p<0.01. Table 4 displays regression results when total penalty minutes per game are regressed on *total.seasons*, *total.seasons.squared* and *log.total.seasons*. *total.seasons* is an experience variable that is the sum of total game experience for each referee in a particular game, divided by 82. For each referee, total game experience at time t is calculated as the total number of games refereed since 1996. Because the data was collected beginning 1996, the regression results above use only games after December 31, 1999 in order to avoid a censoring problem.

Figure 1: Histogram of Penalties in Minutes

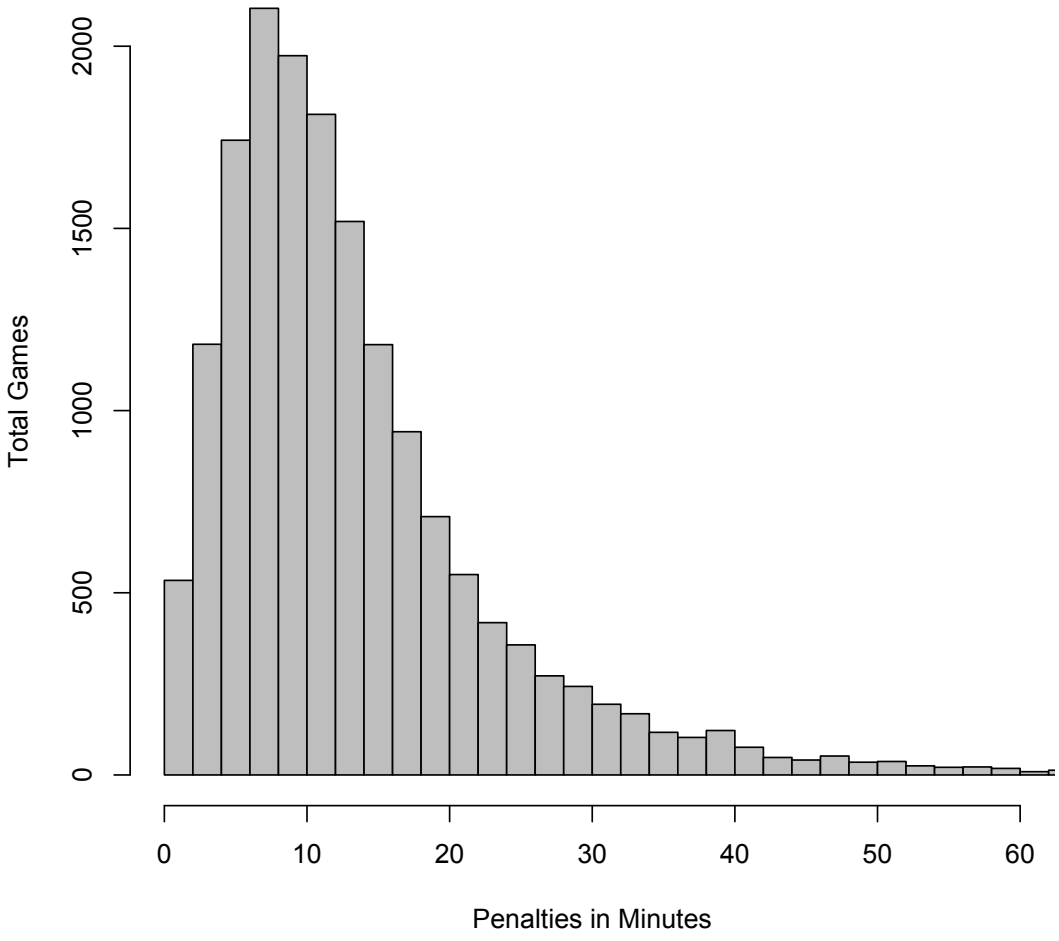

Figure 4 displays the average *pims* by experience for all referees who begin refereeing October 1, 1996 or after. For games with two referees, the total number of *pims* is the total called by both referees.

Figure 2: Percentage of Games with Two Referees per Season

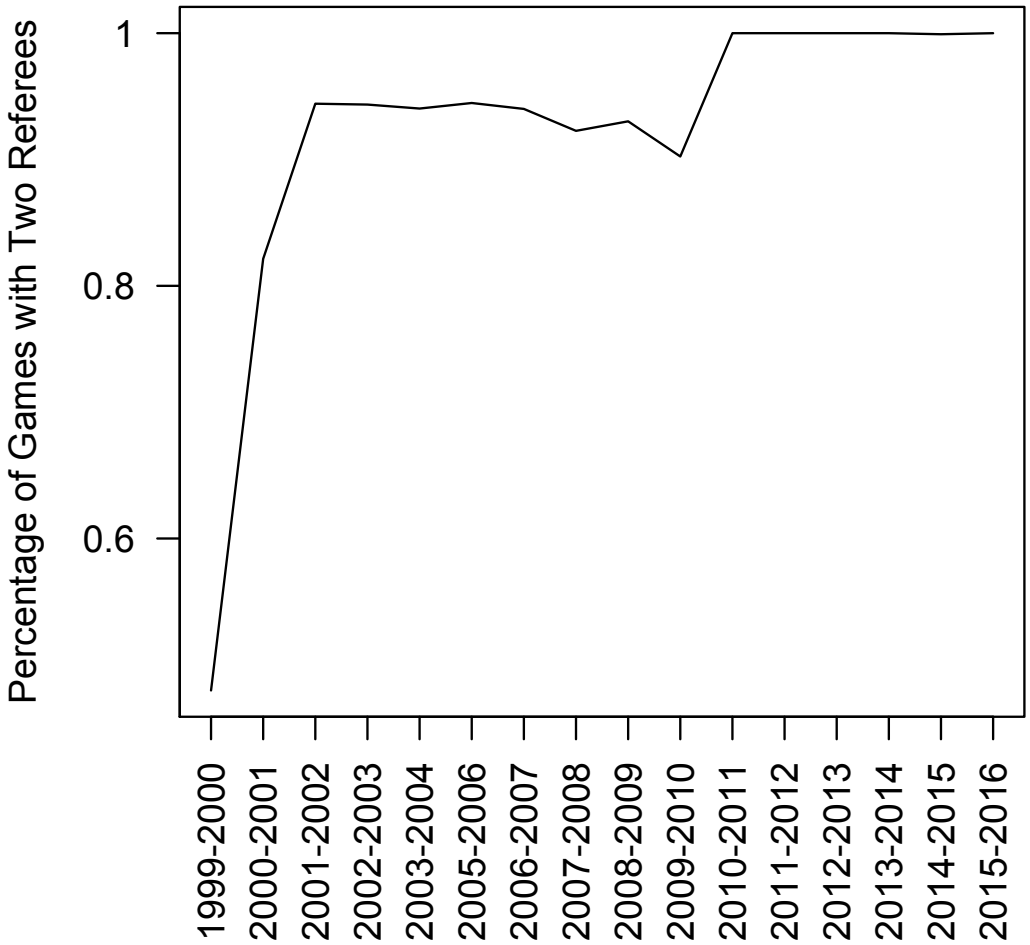

Figure 2 displays the percentage of games with two referees by season.

Figure 3: Average Penalty Minutes per Game per Season

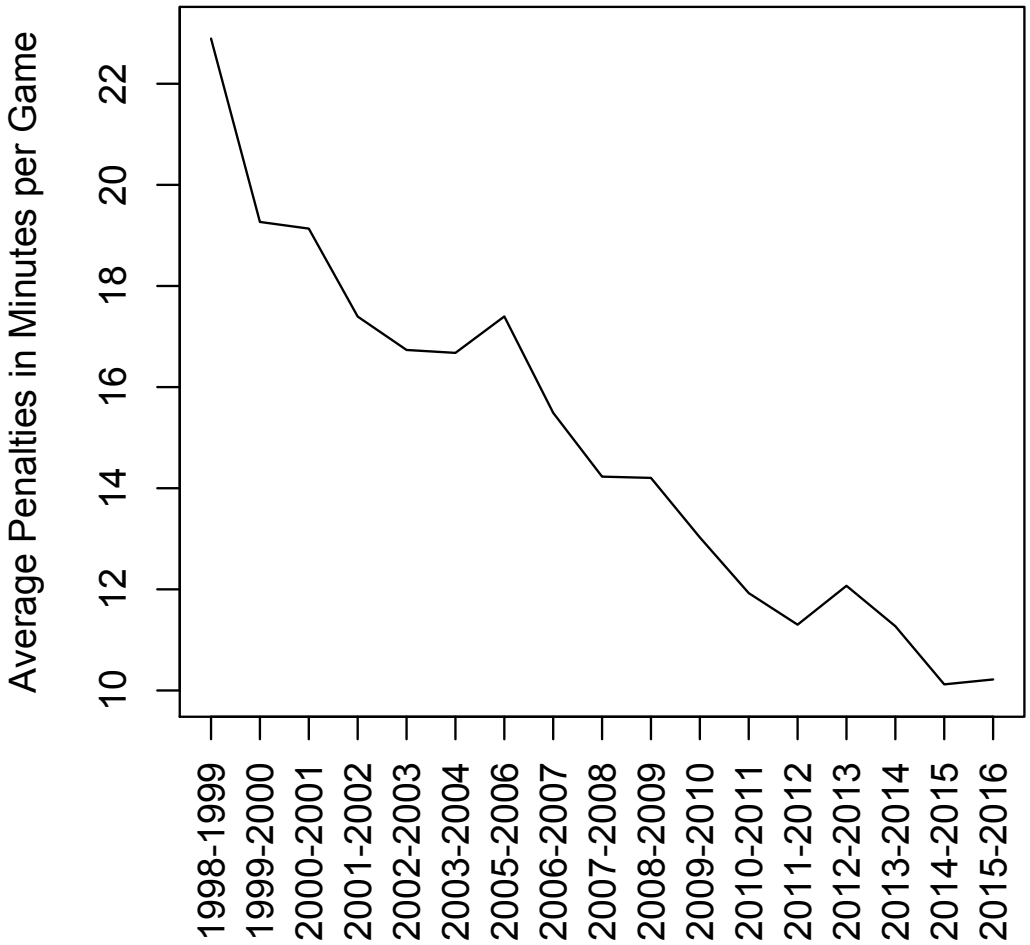

Figure 3 displays the average *pims* by season.

Figure 4: Average Penalty Minutes per Game by Years of Experience

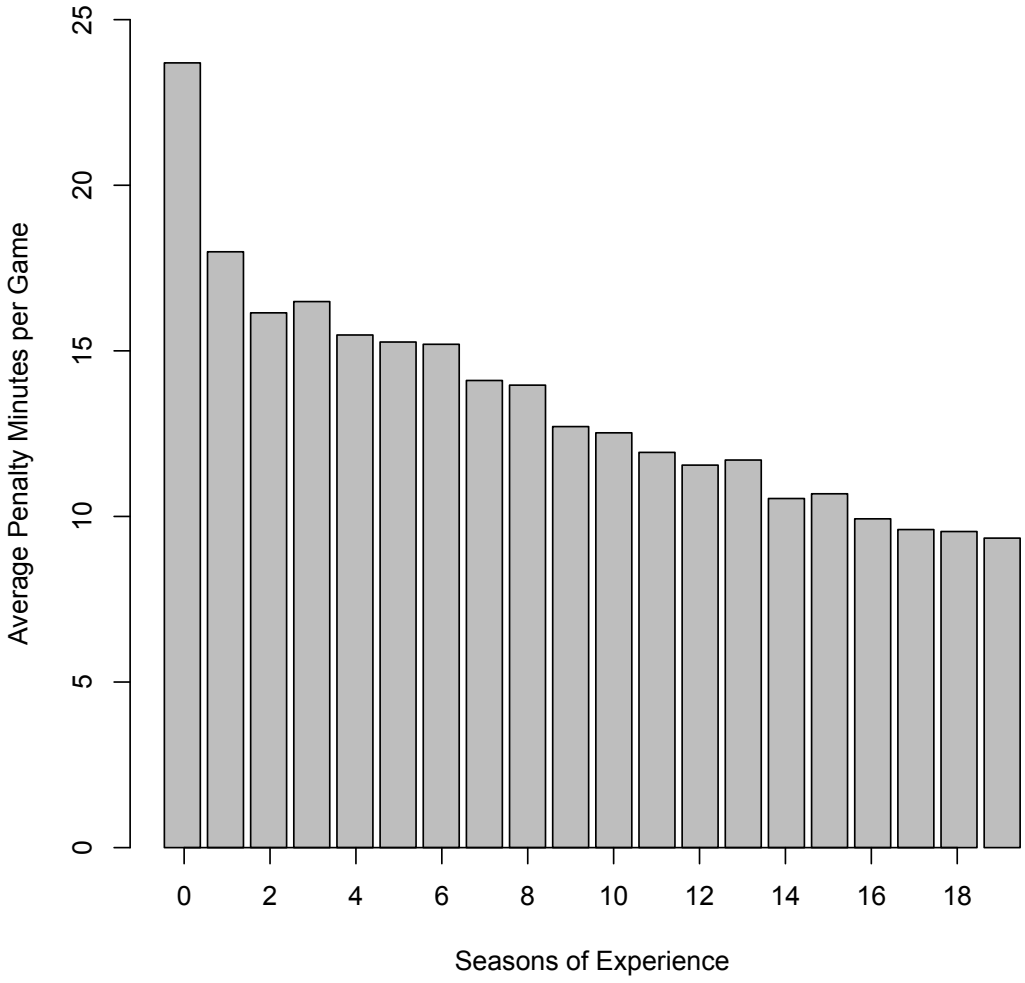

Figure 4 displays the average *pims* by experience for all referees who begin refereeing October 1, 1996 or after. For games with two referees, the total number of *pims* is the total called by both referees.

Figure 5: Correlation of Variables Post 2005

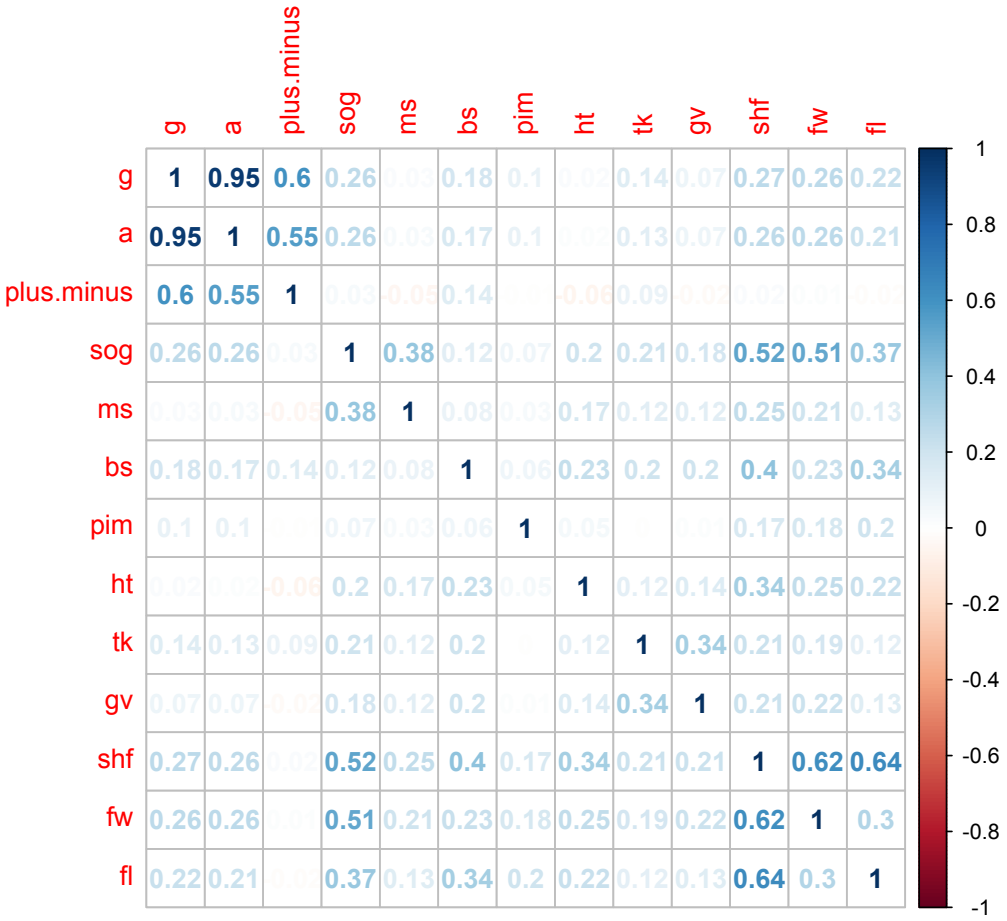

Figure 5 displays the correlation for team statistics following the 2004-2005 season. We only use this period as not all variables are included in the box score before this season.