Examining Regulatory Capture with High Frequency Data

Gregory DeAngelo* West Virginia University

Adam Nowak† West Virginia University

Imke Reimers‡ Northeastern University

March 17, 2016

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 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 (complaining), 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 (hitting below the knee), 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 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 impacts of a second referee on penalties per game using data from the 1999-2000 season with mixed results. Depken and Wilson (2004) find the second referee is possibly revenue enhancing: games with two referees have more scoring, and games with more scoring draw a larger television audience. We significantly expand the data used in all of these studies by using all games from 2000 to the present. While these studies are divided over the significance of the 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

[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 (see Buraimo et al., 2009 and Garicano et al., 2005), 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.

report the results of our estimation in Section 4. Finally, we conclude by discussing our results 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), Heckelman and Yates (2003), and **?** examine variation in the number of referees in NHL games (assuming that referees are homogeneous) and find mild evidence that increased monitoring decreases 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) have 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.

Given these findings, 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 Premier League and find referees are biased in favor of the home team, and 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 and Estimation Strategy

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*.[4] 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 our measure of the referee's game experience in seasons, $total.seasons = total.games/82$. 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.

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.[5] Thus, there is demonstrated discretion in the application of rules and

[4] 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.

[5] 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

it is this discretionary power that we exploit in measuring the extent of regulatory capture in our data.

Furthermore, the box scores do not specify which referee called a specific penalty. We follow Joseph Price (2010) by focusing on the total number of penalty minutes called by the referee *crew* in a game. For example, if there are three minor (two-minute) penalties, one double minor (four-minute) penalty, and one major (five-minute) penalty in a game, we record 15 penalty minutes. Similarly, we look at the total amount of experience for the referee crew, rather than each individual referee's experience. That is, our explanatory variable of interest is $total.seasons = total.seasons_1 + total.seasons_2$ for referees 1 and 2 in a game. If both referees in the crew each have 82 games of experience, $total.seasons$ would be equal to $2 = (82 + 82)/82$.

Table 1 shows descriptive statistics for all games beginning January 1, 2000. In the following regressions, penalties in minutes, *pims*, is the dependent variable. In order to remove the effect of any outliers, we only use those games with less than 100 penalty minutes; this filter removes 16 games from our data set. Table 1 shows that, on average, referees call 14.3 minutes of penalties in a game. As each minor penalty is a 2 minute penalty, 14.3 penalty minutes is the equivalent of 7.15 minor penalties. There is significant variation in the number of penalty minutes across games, with 10 games having no penalties at all, and other games having more than 60 minutes of total penalty time over the course of a 60 minute game.[6]

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. In our data set, 93.3% 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 an indicator for games with two referees in the estimation.

Finally, referee experience for the crew is also displayed in Table 1. The average number of total games of experience for a crew in the sample is equal to 832.1 total games or just over ten years of combined experience per crew, although there is considerable variation. Some referee crews feature two referees making their debut ($total.games = 2$) while other crews have nearly thirty seasons of combined experience.

With data covering a large span of time, we are able to estimate the impact of experience on the

[6]Because only 10 games have 0 *pims*, we are unable to estimate the effects of any censoring in the data above 0 using a Tobit or related methods. However, because the number of games with 0 penalties is negligible, we believe that the effect of any censoring is minimal. We thank an anonymous referee for pointing this out to us.

number of penalties that a referee calls. Figure 1 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 1 shows a clear decline in the number of penalties called per game as referee experience increases.

First-year referees call as many as 24 penalty minutes per game, which corresponds to almost ten penalties on average. Referees with one year of experience call 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 referee crews with zero years of experience.

Whether these differences are due to regulatory capture remains to be determined in the estimation. To that end, we estimate

$$pims_g = \beta_0 + \beta_1 total.seasons_g + \beta_2 two.ref_g + \mu_{year} + \mu_{month} + \mu_{ref1} + \mu_{ref2} + \mu_{team1,team2} + u_g, \quad (1)$$

where $pims_g$ denotes the number of penalty minutes assessed in game g. We are most interested in the effects of the referee crew's experience, $total.seasons_g$, and the number of referees, $two.ref_g$. The regression also includes a number of fixed effects to control for time trends as well as idiosyncrasies among referees and hockey teams. Specifically, μ_{year} and μ_{month} are fixed-effects for the year and month, respectively. The variables μ_{ref1} and μ_{ref2} are fixed effects for each ref. In games with only one referee, the effect is doubled as we set $\mu_{ref2} = \mu_{ref1}$. By including fixed effects for each referee, we are able to control for any referee specific views on penalties that do not change over time. The variable $\mu_{team1,team2}$ is a fixed-effect specific to each team pair. For example, the Colorado Avalanche and the Detroit Red Wings have a unique fixed-effect. We use the interaction of the teams rather than individual team effects because we want to control for any 'rivalry' effects that would increase intensity and, presumably, the number of penalty minutes.

In addition to the game level analysis, we also focus on the individual teams in each game and analyze if experience with a specific team influences the penalty minutes called. Table 2 displays summary statistics for the home and away teams in each game.[7] The "home games" ("away games") variable describes the total number of games the crew has had with the home team (away team) so far, including the current game. Correspondingly, "home seasons" ("away seasons") is the number

[7]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.

of games of experience with each team divided by 82.

Table 2 shows that some referees have a considerable amount of experience with specific teams. This is in part due to logistical factors that limit where a referee can be on a given day, although it is also affected by their overall experience. For example, the crew of Marc Joanette and Dave Jackson refereed a game between the Detroit Red Wings and the Montreal Canadiens game on October 17, 2015. Combined, these two had refereed 205 games with the Canadiens and 134 games with the Red Wings. In contrast, on February 11, 2002, Eric Furlatt was the only referee in a game between the same two teams; he had only refereed 5 games with the Canadiens and 1 with the Red Wings, before that game. Despite lacking team-specific experience, Furlatt had some overall experience as he had previously refereed 24 games. By exploiting differences in team experience as well as overall experience, we can determine if regulatory capture happens at the team level as opposed to a league level.

In order to test for this form of regulatory capture in the data, we examine the penalties for each team in each game and use team-specific experience. By doing so, we double the total number of observations as each game has two teams. For each team-referee pair in a game, we calculate the total seasons of experience with the team, $team.seasons$, analogously to the $total.seasons$ defined above. For instance, Marc Joanette and Dave Jackson have refereed 205 games involving the Montreal Canadiens; for this observation $team.seasons = 205/82 = 2.5$. The variable $team.seasons$ is then added to Eq (1) as an additional explanatory variable. For each game g and team t, we estimate

$$\begin{aligned} pims_{gt} &= \beta_0 + \beta_1 total.seasons_{gt} + \beta_2 two.ref_{gt} + \beta_3 team.seasons_{gt} + \beta_4 home_{gt} \\ &\quad + \mu_{year} + \mu_{month} + \mu_{ref1} + \mu_{ref2} + \mu_{team1,team2} + u_{gt}, \end{aligned} \tag{2}$$

where $home_{gt}$ is an indicator variable equal to 1 if team t is the home team. Also, $u_{gt} = u_g + v_{gt}$ is the error term specific to each game. It includes a random component common to both teams playing in the game and a random term specific to each team. Since we cannot rule out the presence of μ_g, we estimate equation 2 in a the seemingly unrelated regression. In results below, we report the correlation between u_{gHOME} and u_{gAWAY}.

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 referees tend to permit more infringing actions as they gain more experience in the league, as in equation 1. We further control for the number of referees in the game, and we include year, month, referee, and team rivalry fixed effects in our regressions to control for the games' overall levels of aggressiveness. In a later regression, we add to this the referees' experience with the individual teams on the ice in order to check whether a referee who has worked with a certain team more frequently is more susceptible to regulatory capture.

Table 3 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 season of experience, the expected number of penalty minutes decreases by 0.262. Given the longevity of some referees, this number is economically significant as well: referee teams with the mean level of experience (10 years) call 2.6 fewer penalty minutes per game than rookie referee teams or 1 fewer minor penalty.[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 1.038 more penalty minutes than games with only a single referee.[10] 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).

The unreported year and month fixed effects indicate a gradual decrease of penalty minutes over time. Interestingly, the only month with a statistically significant coefficient is October. Typically, the NHL season begins during October, which could affect the number of penalties called in three

[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.

[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 highly 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).

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 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 1 suggests that the effect of experience on the number of penalties called may not be linear. Columns 2 and 4 of Table 3 enter our variable of interest - the referee crew's total seasons of experience - in quadratic and log forms. Consistent with Figure 1, column 2 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 21.39 seasons of experience before they would begin to call more penalties. Since almost no referee crews have this much combined experience, we conclude that the effect of referee experience on expected penalty minutes is decreasing when accounting for quadratic effects. Column 4, using the natural log of total seasons called as the independent variable of interest, reaches similar conclusions. [BRAD WROTE "DON'T CARE" ABOUT "Solving for how long...." ON THE PIECE OF PAPER HE GAVE ME. PROB SHOULD CUT.']

Finally, column 3 includes the sum of total seasons for the referees (the referee crew's added experience) as well as the product of the total seasons for the two referees, defined as $total.seasons_1 \times total.seasons_2$. The estimate of this coefficient is positive and highly significant while the coefficient on the sum of the two referees' experiences is negative (and an order of magnitude larger). 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.

We then expand the analysis to include the referee crew's experience with each team, as described in equation 2. Table 4, shows the results from this specification, which follows DeAngelo and Owens (2015). It shows that the referees' experience with the individual team plays a large

role in determining how many penalty minutes are assessed. Column 1 indicates that each additional season of experience with a given team decreases the penalty minutes in a game by 0.972 whereas an additional season of overall game experience will decrease the penalty minutes called by only 0.068. These results remain roughly consistent across different specifications, showing that a referee's experience with a specific team significantly reduces the number of penalties assigned to that team, and that team-specific experience plays a larger role than overall NHL experience when determining *pims*.[11]

Overall, our results show that those referees who have spent more time on the ice with specific teams systematically call fewer penalties against that team than they do against teams that they have relatively less experience officiating.[12] The development of relationships with the teams appears to play a significant role in determining the number of penalties assessed, and this effect 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. 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

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

[11] We also look at the total experience of each referee-pair. For example, how many games have Dave Jackson and Wes McCauley monitored *together*. The results are similar to the reported results.

[12] Unfortunately, it is not possible to determine if teams take fewer illegal actions because they have more experience with a referee.

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, it does offer behavioral expectations about the 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.*

Depken, C. A. and Wilson, D. P. (2004). Wherein lies the benefit of the second referee in the nhl? *Review of Industrial Organization*, 24(1):51–72.

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.

Joseph Price, J. W. (2010). Racial discrimination among nba referees. *The Quarterly Journal of Economics*, 125(4):1859–1887.

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.

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.

6 Tables and Figures

Table 1: Descriptive Statics

Variable	N	Mean	St. Dev.	Min	Max
Year	16,731	2007.558	4.646	2000	2015
Penalty minutes	16,731	14.283	10.582	0	98
Referee crew's experience (games)	16,731	832.077	428.081	2	I 2,410
Referee crew's experience (seasons)	16,731	10.147	5.221	0.024	29.390
Two referees	16,731	0.933	0.249	0	1

Table 2: Team Specific Experience

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

Table 3: Penalties in Minutes and Referee Game Experience

	Dependent variable: penalty minutes			
	(1)	(2)	(3)	(4)
Two referees	1.038***	1.36**	1.261**	1.471***
	(0.431)	(0.441)	(0.435)	(0.448)
Experience	−0.262***	−0.599***	−0.494***	
	(0.003)	(0.106)	(0.076)	
Experience^2		0.014***		
		(0.004)		
Experience$_1$ × Experience$_2$			0.036***	
			(0.010)	
log(Experience)				−2.066***
				(0.186)
October	2.826***	2.809**	2.787**	2.845***
	(0.430)	(0.430)	(0.430)	(0.430)
Observations	16,554	16,554	16,554	16,554
R^2	0.112	0.113	0.113	0.113
Adjusted R^2	0.083	0.083	0.084	0.083
Residual Std. Error	14.810 (df = 16021)	14.562 (df = 16021)	14.562 (df = 16020)	14.800 (df = 16021)
F Statistic	3.808*** (df = 532; 16021)	3.824*** (df = 532; 16021)	3.824*** (df = 533; 16020)	3.827*** (df = 532; 16021)

Note: *p<0.1; **p<0.05; ***p<0.01. "Experience" is the total number of games the referee crew has officiated up to and including a particular game, divided by 82 (the number of games in a season). "Experience$_1$" and "Experience$_2$" are defined analogously for referee 1 and referee 2, respectively. 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. All regressions use year, month, referee and team-pair fixed-effects.

Table 4: Penalties in Minutes and Referee Team and Game Experience

	Dependent variable: penalty minutes			
	(1)	(2)	(3)	(4)
Two referees	0.071	0.073	0.073	0.085
	(0.200)	(0.200)	(0.200)	(0.200)
Team-specific experience	-0.972^{***}	-0.972^{***}	-0.972^{***}	-1.007^{***}
	(0.181)	(0.181)	(0.181)	(0.181)
Experience	-0.068^{***}	-0.088^{*}	-0.088^{*}	
	(0.019)	(0.051)	(0.051)	
Home team	-0.694	-0.694	-0.694	-0.694
	(0.089)	(0.089)	(0.089)	(0.089)
Experience^2		0.002		
		(0.004)		
$\text{Experience}_1 \times \text{Experience}_2$			0.002^{***}	
			(0.004)	
log(Experience)				-0.206^{***}
				(0.053)
October	1.422^{***}	1.420^{**}	1.420^{**}	1.405^{***}
	(0.166)	(0.166)	(0.166)	(0.166)
$\text{cor}(u_{gHOME}, u_{gAWAY})$	0.102^{***}	0.102^{***}	0.102^{***}	0.102^{***}
Observations	33,108	33,108	33,108	33,108
R^2	0.094	0.094	0.094	0.094
Adjusted R^2	0.079	0.079	0.079	0.079
Residual Std. Error	8.098 (df = 32,574)	8.098 (df = 32,573)	8.098 (df = 32,574)	8.098 (df = 32,574)
F Statistic	6.317^{***} (df = 533; 32,574)	6.305^{***} (df = 534; 32,573)	6.305^{***} (df = 534; 32,573)	6.321^{***} (df = 533; 32,574)

Note: *p<0.1; **p<0.05; ***p<0.01. "Experience" is the total number of games the referee crew has officiated up to and including a particular game, divided by 82 (the number of games in a season). "Experience_1" and "Experience_2" are defined analogously for referee 1 and referee 2, respectively. 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. All regressions use year, month, referee and team-pair fixed-effects.

Figure 1: Average Penalty Minutes per Game by Years of Experience

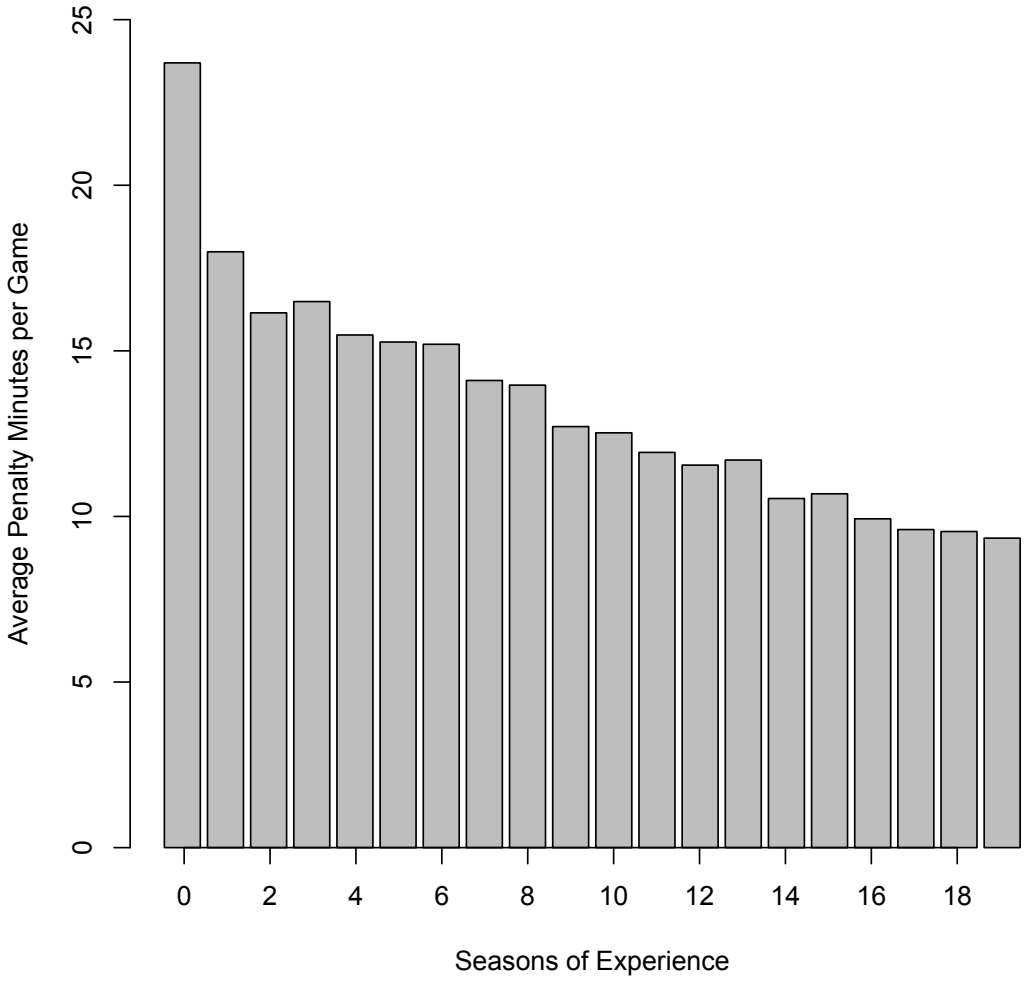

Figure 1 displays the average *pims* by experience for all referees who begin refereeing October 1, 1996 or after. For games with two referees, the total number of *pims* is the total called by both referees.