# Community and Specialized Enforcement: Complements or Substitutes?

Gregory DeAngelo*
West Virginia University

Brad R. Humphreys†
West Virginia University

Imke Reimers‡
Northeastern University

March 23, 2016

**Abstract**

Specialized enforcers are charged with the responsibility of upholding laws intended to discourage proscribed behavior. Community enforcement discourages behavior that is prohibited by codified laws and behavior violating social norms. Previous research utilizes temporally and spatially aggregated crime and enforcement data, which obscures the relationship between the timing and commission of anti-social and enforcement actions. We utilize a unique environment (the National Hockey League) with detailed event-level data to analyze the impact of enforcement proxies: penalties and fights. Our evidence suggests that penalties and fights share a state-dependent relationship. When specialized enforcement efforts are lacking, community enforcement acts as a substitute. Conditional on the occurrence of community enforcement, however, specialized enforcement further deters anti-social behavior.

JEL Codes: H4; K42; L51; Z2
Key Words: deterrence; public and private law enforcement; community enforcement

*The authors are grateful to Joel Waldfogel, Bryan McCannon, Mark Anderson, Pierre Chausse, Phil Curry, Dan Nagin, Jacob Shapiro, discussants at the Southern Economic Association 2015 Conference, Western Economics Association 2016 International Conference, and participants at the Northeastern University Economics reading group.
Department of Economics, College of Business and Economics, PO Box 6025, Morgantown WV 26506-6025; email: gregory.deangelo@mail.wvu.edu

†Department of Economics, College of Business and Economics, PO Box 6025, Morgantown WV 26506-6025; email: brhumphreys@mail.wvu.edu; phone: 304-293-7871.

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

# 1 Introduction

Throughout history societies have utilized many mechanisms, policies, and practices to enhance welfare and deter anti-social behavior. Individuals' responses to these interventions are inherently interesting and important. For example, the introduction of automotive safety measures like seat belts and air bags often increases the level of anti-social behavior, in this case aggressive driving (Peltzman, 1975). To guide these interventions, a substantial theoretical literature, primarily based in game theoretic models, examines the effectiveness of various policies and practices designed to encourage pro-social behavior, and the conditions under which this will emerge as an equilibrium outcome (Deb and González-Dıaz, 2014). This literature identifies two general approaches to deterring anti-social behavior: specialized enforcement (government-sponsored police activities that uphold the law) that monitors individuals and punishes those engaged in anti-social behavior, and community enforcement that discourages anti-social behavior through threats to withhold future cooperation by members of the community and other actions.[1]

In a related line of research, the order-without-law literature generates interesting predictions about the deterrence effect of community norms on proscribed behavior. Acemoglu and Jackson (2014) highlight the importance of understanding substitutability and complementarity between private and public law enforcement as an important step in this line of research.[2] However, empirical research on the economics of crime often suffers from an important omitted variables problem: community enforcement of laws impact outcomes in crime data (e.g. Uniform Crime Report data and other publicly available sources of crime data), but cannot be directly observed. In this research, we analyze data from a unique environment where both community and specialized enforcement can be jointly and separately identified.

While data limitations have traditionally presented difficulties for applied researchers, the interplay between these two types of enforcement have received substantial attention in the theoretical

[1]Community enforcement can take many forms. Non-coercive organizations such as community watch programs ((Bennett et al., 2008)), business improvement districts (Cook and MacDonald, 2011), and civic organizations such as Mothers Against Drunk Drivers (MADD) were formed to discourage anti-social behavior. Alternatively, more coercive organizations, like mafias and gangs, also police anti-social behavior (Sobel and Osoba, 2009; Roth and Skarbek, 2014). In all of these cases the groups enforce well established legal rules or community norms (see Hume (1793); Bicchieri and Muldoon (2011)) established through relationships built on reciprocity.

[2]While previous empirical research that examines the relationship between community and specialized enforcement concludes that these forms of enforcement are complements, this may reflect sorting, and not a causal relationship. That is to say, individuals with a higher willingness-to-pay for deterrence will Tiebout (1956) sort into locations offering private security in addition to specialized enforcement.

literature. Acemoglu and Wolitzky (2015) develop a theoretical model of interaction among agents containing both specialized and community enforcement. The model identifies conditions under which anti-social behavior is optimally punished by only specialized enforcers and conditions under which both specialized enforcement and community enforcement are needed to maintain societal order. One key prediction establishes that a single enforcer is not optimal in cases where all or part of society benefits from imperfections in the monitoring structure. The model is general, and extends ongoing research related to community enforcement in repeated games (Ellison, 1994), economic enforcement of laws and norms (Greif, 1993; Acemoglu and Jackson, 2014), and penal codes in general repeated games (Abreu, 1988).

The deterrence effect of specialized and community enforcement has also received empirical attention - primarily using aggregated data - in economics of crime literature.[3] While data on the activity of specialized enforcement have become increasingly more available at finer levels of temporal and spatial disaggregation, data on community enforcement are rarely observed at finer levels of granularity.

We analyze the determinants of anti-social behavior in a setting where both specialized enforcers and community enforcement exist, high-stakes interactions occur repeatedly among clearly identified groups of individuals, and a benefit from monitoring imperfections in the enforcement of anti-social behavior exists. Utilizing data from professional ice hockey games in the National Hockey League (NHL), we observe all occurrences of anti-social behavior (hitting) and whether the behavior led to sanctioning by either specialized or community enforcers.[4] Our empirical setting contains both specialized enforcers (in the form of referees), and community enforcement mechanisms (in the form of fights taking place between members of competing teams) to maintain order.[5]

This setting permits a detailed analysis of the relationship between specialized enforcement and

[3]As Nagin (2013) notes, the reduction in criminal behavior attributable to the threat of punishment is called *deterrence* by economists, but *specific deterrence* by criminologists. Additionally, criminologists include the effect of punishment on recidivism in their definition of *specific deterrence*. Alternatively, *general deterrence* refers to the crime prevention effects of the threat of punishment. For a more detailed discussion of criminal deterrence, see Nagin (2013).

[4]A full list of terms that relate to hockey is provided in the appendix. We use "hitting" and "checking" synonymously. While not all hitting in hockey is illegal or anti-social, it is true that all illegal hitting or anti-social behavior in hockey involves physical contact between at least two competing players. We assume hits represent an important form of anti-social behavior in this setting, and we explain in Section 4 why hits are a good proxy for anti-social behavior.

[5]We discuss the evolution of fighting in Section 3 and show in Section 4 that fights are effectively used as another mechanism to control the game.

community enforcement. Data on the timing of anti-social behavior and both types of enforcement are observable in game logs for over 6800 NHL games played between 2007 and 2013. We observe the exact time of occurrence of all key events, and identify the specific individuals directly involved in all of these events, as well as individuals on the ice who are not directly involved. The detailed information on the timing of anti-social events and enforcement provides a unique environment for analyzing the dynamics of the interaction between specialized enforcement and generalized community enforcement.

Figure 1 summarizes this interaction by examining the response of fighting to changes in the number of penalties assessed. Time zero corresponds to the exact moment when the first fight starts in a hockey game, while 1 (-1) corresponds to the 60 seconds period immediately after (before) the first fight occurred, and so on. It is clear that fights follow periods of lower than average penalty assessment. This relationship would, however, not be evident if we examined temporally aggregated data, as specialized enforcers in turn respond to community enforcement by increasing the number of penalties in the aftermath of a fight. The ability to identify these patterns represents a unique aspect of our empirical analysis, since other data sources do not contain information at this level of temporal disaggregation.

The formal analysis uses data at three levels of aggregation. We first analyze the interaction of specialized and community enforcement and their effect on anti-social behavior at a temporally-disaggregated (minute-of-play) level. Next, we assess the effectiveness of the two types of enforcement in deterring anti-social behavior of different members of society: aggressors (those directly involved in a penalty or fight), direct observers (teammates of the aggressors), and indirect observers (members of the opposing team). Finally, we compare our estimates to the results we would have obtained using traditional, temporally aggregated crime data by analyzing the interaction of fights and penalties at a temporally-aggregated (game) level.

We first estimate models explaining the occurrence of fights as a function of the number of penalties called in the previous minutes, and the number of penalties called in the minutes following a fight. We find that penalties and fights operate as substitutes: penalties are unusually low in periods preceding fights, although referees increase their enforcement activity immediately after a fight occurs. We then estimate the effect of penalties and fights on hits, both individually and in conjunction, in regression models including lagged hits, penalties and fights. This analysis shows that both fights and penalties deter anti-social behavior, and that the order in which fights and penalties occur influences their effectiveness.

We then disaggregate the data further to the player-minute-level to account for unobservable player-level heterogeneity. This allows us to examine the differential response of players to both types of enforcement, analogous to citizens responding differently to community and specialized enforcement. Our results indicate that penalties are most effective at discouraging miscreants from committing future anti-social acts, but do not necessarily extend to observers. Fights, on the other hand, discourages anti-social behavior in a more systematic manner that involves coaches removing players that have a higher inclination to engage in egregious behavior.[6]

Finally, we compare our results to more traditional crime data by analyzing the interaction between, and the effects of, the two types of enforcement at the temporally aggregated game-level. Since we cannot rely on the exact temporal order of enforcement actions to establish causal relationships in this setting, we use a measure of how well-rested the game's referees are as an instrument to explain the number of penalties assessed in a game. Although we confirm the validity of the instrument, the results of this IV analysis are not consistent with the temporally-disaggregated results, likely due to aggregation bias. Thus, we show that the temporal disaggregation of crime data is essential in capturing critical details about the behavioral responses of individuals committing anti-social behavior to different types of enforcement.

## 2 Context

This research is motivated by, and contributes to, several ongoing and related lines of research in the area of crime and economics. In his seminal paper on the economics of crime, Becker (1968) mentioned the idea that private expenditure, as well as publicly provided policing and courts, could be used to deter crime and provide security. Subsequent research focused on the role that private activities can play in deterring anti-social behavior (Benson and Mast, 2001; Cook and MacDonald, 2011; DeAngelo and Smith, 2015). The private solutions in this line of research can be interpreted as community enforcement, or alternatively as examples of privately provided security. Our empirical analysis addresses one key empirical question in this literature, which has been understudied due

[6] We consider player behavior to be "egregious" if the individual intentionally behaves in a manner that inflicts avoidable harm on another player.

to data limitations: how much deterrence is generated by public and private enforcement, both independently and in conjunction?

Since the community enforcement examined here takes place spontaneously, with no formal institutions or organizations promoting or planning these actions, this research also contributes to the line of research focused on understanding the emergence of order without law (Ellickson, 1991; Leeson, 2007, 2009). This literature focuses on how order - in the form of institutions, practices, or social customs - can emerge to deter anti-social behavior in situations where no formal, controlling bodies or organizations (e.g. police or courts) exist. In most papers, the environment is characterized by no clearly codified set of laws or regulations to serve as a formal, public method of encouraging pro-social behavior on the part of self-interested agents. The general approach in this literature is to identify specific situations where order emerges without public enforcement of laws, examine the institutions and practices that emerge, and assess the effectiveness of these institutions and practices in encouraging cooperation. In our context, the community enforcement activity - fights between hockey players on rival teams - emerges spontaneously based on the conditions in the game.

Most of the current empirical literature on the effectiveness of enforcement types uses highly aggregated data as temporally and spatially disaggregated data are not generally available.[7] Using temporally and spatially aggregated data, these studies estimate the elasticity of crime incidence with respect to changes in police manpower and other, similar indicators of deterrence. This approach ignores any role played by community enforcement, since these activities are difficult to quantify and observe. It also assumes that measures of police manpower capture the actual impact of policing at a micro level. Durlauf et al. (2010) discuss issues associated with estimating causal effects of deterrence on crime using aggregate data.

Actions occurring in NHL games represent a bridge between the ideal disaggregated crime data and the temporally and spatially aggregated data used in much of the empirical research on the economics of crime.[8] Players in the NHL face numerous choices throughout the course of a game,

[7] Chalfin and McCrary (2014) review the literature on the deterrence of crime.

[8] Alternatives to aggregated data (e.g. *Uniform Crime Reports (UCR)*, *National Incident-Based Reporting System (NIBRS)*) have become more prevalent. For example, daily or even minute-level data on reported crimes or arrests have become available - see http://www.crimemapping.com/map/ca/sanfrancisco for an example of crime data from San Francisco. However, none of these data report unobserved criminal behavior, nor are they capable of merging information about community and specialized enforcement.

such as passing the puck or taking a shot on goal, checking an opposing player or not, and so on. For the purposes of our analysis, we focus on activities that can involve anti-social behavior. Specifically, we examine the decision to check another player or not.

In order to deter anti-social behavior, a centralized authority - the NHL - places referees on the ice to call penalties against individual players that commit acts prohibited by NHL rules.[9] We interpret penalties called on a specific player for a specific action as specialized enforcement. The impact of specialized enforcement can be understood by analyzing the tendency of the penalized player, his teammates, and his opponents, to deliver hits after the penalty was assessed.

As discussed above, social order can be imposed via community enforcement like private security, conceal and carry laws, community watch programs, business improvement districts, shaming, ostracism, etc. Importantly, hockey provides such a tool, as fighting can be employed by either team to discourage members of the opposing team from engaging in anti-social behavior. Since we observe specific acts of community and specialized enforcement, we can perform unique, dynamic empirical analyses of the effect of both order and law on anti-social behavior.

Our empirical analysis defines hits as anti-social behavior. Hitting or checking is a defensive technique intended to disrupt players on the other team. While some hitting is not illegal in hockey, it reflects the amount of physical contact occurring between players and thus serves as a proxy for anti-social behavior. 25% of all observed injuries in our data set occur within 10 seconds of a hit, underscoring the anti-social nature of this action.[10] Many forms of reckless hitting can be penalized, and referees are responsible for distinguishing legal from illegal hits. Players' opinions on the appropriateness of a hit may differ from referees' assessments. This ambiguity may create a void in specialized enforcement of social order, providing an incentive for community enforcement of social order to emerge.

This paper extends the existing line of research on enforcement and punishment of crime in the context of sports (McCormick and Tollison, 1984). Much of the empirical analysis in this line of research focuses on exogenous variation in the number of referees or officials as an empirical proxy for changes in the level or funding of policing. Examples include McCormick and Tollison (1984),

[9]For a full list of the current rules see http://www.nhl.com/rulebook. Note that some penalties are assessed against the team (e.g. too many players on the ice) and not against specific players. These penalties are removed from our data set.

[10]It is uncommon in hockey for play to be suspended when a player is injured, so that these injuries are likely quite severe.

Allen (2002), Levitt (2002), and Heckelman and Yates (2003). This literature generally concludes that increases in specialized enforcement decrease anti-social behavior. However, past research in this area has not explicitly considered both specialized enforcers and community enforcement, nor has it analyzed data at low levels of temporal aggregation to investigate the dynamics of monitoring and enforcement.

## 3 Institutional Background

Fighting occurred in ice hockey games from the sport's invention. The roots of ice hockey date to the late 1800s, when English-Canadian rugby players, in search of a winter activity, developed the sport. Gopnik (2012) describes early versions of the sport as "an improvisational game played on a frozen street, in part a brutal game of rugby played at high speed, in part a form of soccer on ice." Given the evolution and growth of hockey from these roots, it is not surprising that hockey shares several characteristics with rugby. One characteristic identified by Gopnik (2012) is a mentality that players should settle it themselves rather than rely on officials to maintain order in games.[11]

While early hockey included an attitude that players should resolve issues on their own, this did not include fighting in the manner that currently exists in the sport. Hockey historian Bill Fitsell points out that early violence in hockey typically involved stick swinging[12], an especially dangerous event given the size and weight of hockey sticks. Fitsell has also been quoted as saying that fighting between players was "few and far between in the early years of the game." However, stick swinging was often more violent than fighting, and even resulted in fatalities in some cases.[13] To temper the violence associated with stick swinging, fighting grew to become a part of the sport to introduce a credible form of retribution for stick swinging offenders.

While stick swinging almost never occurs in the modern game, fighting remains a part of hockey, although there is considerable variation in the occurrence of fighting over time.[14] During the early

[11] Vigneault (2011) notes that fighting in hockey could have alternatively grown out of lacrosse, the traditional sport of First Nations Native Americans, where fighting was quite common, since many athletes played hockey during winter months and lacrosse during warmer weather months. Alternatively, Vigneault (2011) argues that fighting could have grown out of Irish hurling and Scottish shinty, where fighting was also common.

[12] Stick swinging involves a player using their hockey stick as a weapon by striking another player, typically in the head, with a chopping motion.

[13] In the early years of hockey, stick swinging violence was thought to be a product of ethnic differences. As noted in a New York Times article, in 1905 Alcide Laurin was "...killed instantly after being hit with a fist and a stick ...". Only two years later, Owen McCourt was also fatally injured after being struck in the head by a stick.

[14] While fighting is an illegal part of hockey, it is also an accepted behavior, since players are typically not ejected

1960s, fights occurred roughly once every five games. By the late 1980s, an average of more than one fight per game took place. In the last five years, the number of fights per game has fallen to one occurrence every two games.

Despite variation in the incidence of fighting over time, the role of fighting has remained largely unchanged. Fights are typically initiated by a specific player, known as an *enforcer*. Enforcers discourage anti-social behavior, in the form of both legal and illegal actions, during games. While illegal actions can be deterred by both referees and enforcers, legal actions that could result in injury can only be deterred by enforcers.[15] In this manner enforcers encourage safety by introducing reciprocal violence (in the form of fighting) toward individuals that choose to commit anti-social acts. League officials have made public comments supporting this idea.[16] The end result of the actions of enforcers and fighting is that aggressing players are forced to consider whether undertaking legal actions that could result in injury to another player are "worth it," given the non-trivial likelihood that the aggressing player could be confronted by the opposing team's enforcer.

Some argue that fighting exists entirely to increase television ratings or attendance (Paul, 2003), but personnel involved in the sport assert that fighting serves the purpose of discouraging behavior that can cause further harm (Lebrun, 2013). Brandon Prust, one of the league's "enforcers", once stated "I get that there are plenty of people who don't like fighting. Trust me when I say that everyone in the league takes head injuries very seriously now. But in a fight, there are no cheap shots. If you take away fighting, there's no real consequences for guys taking runs[17] at each other

from the game for fighting - instead, those individuals involved in the fight receive a five minute penalty, which is equivalent to incarceration. Moreover, while the players involved in the fight are incarcerated, the teams still play at equal strength.

[15]As an example of a legal action that can cause injury, consider player that receives a "suicide pass". A "suicide pass" is a pass that requires the receiving player to look down or away from play in order to successfully receive the pass. While teammates would almost never want to initiate such a pass intentionally, they are sometimes unavoidable in the course of play. Since the recipient of the pass must divert his attention in order to receive the puck, this leaves the receiving player vulnerable to a body check from an opposing defensive player, which can result in injury. The defensive player - who could check the recipient of the pass - has discretion over his actions. The checking player could body check the recipient, potentially injuring the unsuspecting recipient. Alternatively, the checking player could skate through the recipient's stick, eschewing a body check while stealing the puck from the intended recipient. In either case, the checking player has (a) stopped the recipient from proceeding forward with possession of the puck and (b) not committed an infraction that is punishable by the referees. By skating through the stick instead of initiating a body check, the recipient is not vulnerable to injury. Given that referees cannot dis-incentivize the checking player, the enforcer acts as a deterrent to the checking player by providing a credible threat to the checking player by in the form of retaliatory physical violence.

[16]The commissioner of the National Hockey League dubs fighting in hockey as the thermostat in hockey that helps cool things down. (Prime Time Sports Management Conference, Toronto, CA)

[17]"Taking runs" is hockey jargon for aggressively hitting another player, especially players in particularly vulnerable or unsuspecting positions.

...If they take fighting out, and guys aren't worried about answering the bell, I guarantee more people will get hurt from an increase in open-ice body checks." (Stone, 2015) Others have echoed Prusts's sentiments. Professional hockey player Jarome Iginla noted that "fighting holds players accountable for their actions on the ice." (Tosi, 2013). Finally, the management of hockey teams also touts the value of fighting in hockey. Bobby Smith - former professional player and general manager of several professional hockey teams - advocates that fighting provides a service to the game by reducing traumatic spinal cord injuries and concussions.(Smith, 2016) In fact, a 2013 study titled Bodychecking Rules and Concussion in Elite Hockey conducted a study that analyzed approximately 1,410 NHL games during 30 randomly selected weeks between the 2009-10 and 2011-12 seasons. It found that 8.8 concussions or suspected concussions occurred per 100 NHL games. Approximately 9 percent of these injuries were attributed to fighting and the remaining 91 percent were caused by a variety of means, the most common of which were Bodychecking with head contact and Bodychecking with no head contact. (Donaldson et al., 2013) Taking this seriously, we define hitting as our proxy for anti-social behavior.

## 4 Data Description

We make use of event-level data from all regular season National Hockey League (NHL) games played in the 2007/8 through 2012/13 seasons. The NHL makes available game logs (formerly the "Ice Tracker" feature on the NHL web page) that list every event that occurred during games, recorded at the minute and second of game time when the event occurred.[18] These records identify 15 types of events, including shots, turnovers, goals, penalties, hits, stoppage of play, and other outcomes. In addition, the game logs identify which players were involved in each event, and which players were on ice at the time the event occurred. In total, we observe 6,859 games, containing 2,042,037 unique events. On average, we observe an on-ice event every 12.7 seconds, or close to 300 events per game, providing us with a rich environment for disentangling the effect of penalties and fights on anti-social behavior.

Penalties are employed more often than fighting to maintain social order in NHL games. Table 1 contains summary statistics on the number of events per game for the 6,859 games in the sample.

[18]There was a rule change after the 2012/13 season which made fighting more costly. We limit our data to a time period where the cost of fighting was relatively constant.

The average game contains 8.8 penalties and 0.5 fights. This reflects the fact that fighting is a relatively costly mechanism for a team to use to deter hitting, since fights can result in injuries to critical personnel. Moreover, referees are officially charged with the task of policing the game and maintaining social order, and teams can effectively rely on this mechanism in many cases.

There is considerable variation in the number of penalties called, while the number of fights per game is relatively stable, with over 95 percent of the games having two or fewer fights. There are several potential explanations for this difference. First, assessed penalties closely follow the number of infractions that are being committed during a game, and the number of committed infractions varies considerably across games. Alternatively, there could be idiosyncratic differences in the propensity for a referee to assess a penalty.

Table 1 also contains summary statistics for other key variables used in our analysis, including our outcome variable, the number of hits per game. On average, we observe 43 hits per game, with an injury occurring once every four games. Additionally, there are approximately 6 goals scored per game, 55 shots on goal, and 57 face-offs.

Breaking the events of interest down further reveals the variation that we exploit in the empirical analysis. First, consider the number of hits by minute. While half of the minutes in our data do not contain a hit, and another third contain only one hit, we also observe a number of more aggressive minutes, with almost 3000 minutes in our data set containing four hits or more. Of course, there is less variation at the minute level in penalties and fights, as these generally occur less frequently than hits.[19]

# 5 Empirical Analysis

The empirical analysis focuses on identifying and analyzing the relationship between penalties, fights, and hits occurring during NHL games in order to determine how specialized enforcement and community enforcement interact to deter anti-social behavior. Throughout our analysis we draw analogies to the empirical crime literature under the assumption that this environment proxies for more general interactions among individuals in society.

We observe the exact timing of specific events occurring during hockey games. This level of

[19]Minute-level summary statistics are equivalent to game level summary statistics divided by 60.

temporal disaggregation and detail permits analysis of the dynamic interaction between penalties, fights and hits. We first aggregate the data to the minute level to analyze the timing of the two types of enforcement and their effect on hits. Then, we disaggregate the data further by looking at how different players react to enforcement.

### 5.1 Minute-Level Analysis: Penalties and Fights

As previously noted, Figure 1 in Section 1 displays the temporal nature of the relationship between fights and penalties. A clear decline in specialized enforcement occurs in the 3 minutes immediately prior to the first instance of community enforcement, as an average of only 0.1 penalties are called in those minutes, compared to 0.18 penalties per minute before that. We see an even larger increase in penalties shortly after the first fight, relative to earlier play.[20] The outcomes summarized in Figure 1 reflect average behavior over more than 2,600 hockey games (all games in which a fight occurs). In general, during these games, penalties first declines, often followed by a fight, which is, in turn, followed by an increase in penalties.

This temporal relationship between penalties and fights is supported by regression models which control for other factors that might affect the incentive to commit penalties. Our first regression model estimates the emergence of fights in response to variation in the assessment of penalties. This is estimated in the following reduced-form model:

$$fight_{jt} = X_{jt}\beta + \sum_{m=1}^{5} \gamma_m \times penalties_{(t-m)} + \epsilon_{jt}. \tag{1}$$

The unit of observation is a one minute interval of play. The dependent variable $fight_{jt}$ is an indicator that is 1 if there was a fight in minute $t$ of game $j$. We are interested in the coefficients on the number of penalties per minute in the five minutes before the fight, $\gamma_m$. A positive coefficient would suggest that we are more likely to see a fight if the referees assessed penalties $m$ minutes before the minute of the fight, and negative coefficients would suggest that fights follow periods of low specialized enforcement.

The analysis controls for all observable external factors. $X_{jt}$ contains other factors that might affect the players' behavior in a hockey game. These variables include each player's average time spent on ice up to that point in the game, indicator variables for minutes when a power play occurs

[20]Note that these penalties do not include penalties assessed for fighting.

(one team having more players on the ice than the opposing team due to penalties), each team's total shots on goal and goals scored by each team up to that minute of the game, the number of hits in that minute, the number of injuries in the game leading up to that minute, and the players' average time spent on ice, when the teams were at even strength, in power play, and in penalty kill situations.[21]

We also include game fixed effects and minute indicators. These variables control for heterogeneity in the incentive to commit infractions at specific points in time during the game and across games. For example, players might be more fatigued toward the end of the game and, as a result, increase their tendency to commit an anti-social act.

Table 2 shows that periods of low referee activity are more likely to be followed by fights than periods of high referee activity.[22] Across the three specifications that we examine in Table 2, we find statistically significant and negative coefficients for penalties assessed, especially in the two minutes prior to a fight. Thus, the reaction to reduced penalties happens quite promptly. This supports Ellickson's (1991) assertion that society may experience more order when there is a lack of law.

The emergence of fighting serves to restore order. It also signals to specialized enforcement that anti-social behavior is at an unacceptably high level. Once this signal has been sent, specialized enforcers might react by increasing their enforcement presence. We test for this response in the following reduced-form model:

$$penalties_{jt} = X_{jt}\beta + \sum_{m=1}^{5} \theta_m \cdot \mathbb{1}[t - t_{fight} = m] + \epsilon_{jt}. \tag{2}$$

Again, $X_{jt}$ controls for all observable factors that might affect player behavior. The dependent variable $penalties_{jt}$ is the number of penalties called by referees in minute $t$ of game $j$. Here, we are interested in the coefficients on the five minutes after the fight, $\theta_m$, as they provide insight into whether referee activity rose more than warranted by changes in game situations immediately after a fight. As fights may arise endogenously after periods of unusually low referee activity (see Table

[21] Note that goals scored by each team up to that minute of play captures how closely contested the game is at each minute in the game. The goal difference in a game can drive anti-social behavior if, for example, increasing hitting results in an opponent reducing the level of aggression in their play and increases a team's likelihood of scoring a goal and subsequently winning the game.

[22] As previously mentioned, monitoring imperfections in specialized enforcement can exist. In the context of hockey, this could involve referees missing a penalty that should have been assessed or intentionally not calling a penalty in order to "make up" for a previously missed call.

2), we drop all observations before a fight takes place when examining referee activity after a fight.

Table 3 shows the parameter estimates and standard errors for the key parameters of interest in Equation (2): the indicator variables for each of the five discrete one minute periods of play after a fight occurs. In column (1) penalties are expressed in levels, in column (2) penalties are expressed in logs, and column (3) reports results of a Poisson estimation.

The parameter estimates on the minute indicator variables after a fight occurs are positive and statistically different from zero in the first minute after a fight occurs. Penalties called increase by 0.43 penalties per minute, or 271% (= $e^{1.312} - 1$), although the effect fades quickly.[23] Penalties briefly increases after community enforcement occurs, but reverts to a normal level shortly thereafter.

### 5.2 Minute-Level Analysis: Do Penalties and Fights Deter Hits?

Figure 1 and Tables 2 and 3 establish that fights follow periods of significantly fewer penalties, and that fights are followed by increased penalties. This result holds even after controlling for all other events occurring in the game.

Most interestingly, the effect is robust to the amount of aggressive behavior in the game: the number of hits delivered by both teams immediately before and after a fight. Figure 2 provides a breakdown of the mean number of hits in each minute before and after a fight. In the three minutes leading up to a fight, the number of hits per minute increases from 0.78 to 0.93. In the minutes after a fight, hitting decreases to around 0.67 hits per minute. Additionally, Figure 3 examines the relationship between hits and egregious penalties.[24] A similar trend to the one observed between hits and fights emerges. Specifically, in the two minutes prior to an egregious penalty being assessed, hitting increases from 0.75 hits per minute to 0.8 hits per minute. In the 3-4 minutes after an egregious penalty is assessed, the number of hits per minute decrease by approximately 40 percent.

The patterns from Figures 2 and 3 - increased hitting before a fight or penalty, and decreased

[23] Penalties have a significant impact on the outcome of a game. In this sample, if a team has one fewer penalty than their opponent, they have an 86% higher likelihood of winning the game.

[24] We examine egregious penalties due to the frequency with which penalties are assessed in a game. By focusing on egregious penalties, which are the types of penalties that are associated with player injuries, we reduce the number of penalties that are used as objects of interest. Additionally, this allows us to more cleanly identify pre- and post-penalty periods.

hitting after a fight or penalty - suggest that fights and penalties fulfill the same purpose of reducing hits. Whether the two types of enforcement are effective in deterring anti-social behavior - separately and together - requires a more detailed analysis. To this end, we estimate a series of regression models where we focus on how the number of hits in a given minute varies in the periods immediately following penalties and fights, and how this is affected by whether a fight (penalty) closely follows a penalty (fight):

$$\begin{aligned} hits_{jt} &= X_{jt}\beta + \alpha_p \mathbb{1}[0 < t - t_{\text{penalty}} \leq 5] + \alpha_f \mathbb{1}[0 < t - t_{\text{fight}} \leq 5] \\ &+ \alpha_{fp}\mathbb{1}[t_{\text{fight}} < t_{\text{penalty}} \leq t_{\text{fight}} + 5] \times \mathbb{1}[0 < t - t_{\text{penalty}} \leq 5] \\ &+ \alpha_{pf}\mathbb{1}[t_{\text{penalty}} < t_{\text{fight}} \leq t_{\text{penalty}} + 5] \times \mathbb{1}[0 < t - t_{\text{fight}} \leq 5] + \epsilon_{jt}, \end{aligned} \tag{3}$$

where $t = 1, \ldots, 60$ identifies the minute of play in game $j$, and $X_{jt}$ includes lagged hits and penalties in addition to the controls discussed in Section 5.1. We are mainly interested in $\alpha_p, \alpha_f, \alpha_{pf}$, and $\alpha_{fp}$, the coefficients on the five-minute periods after penalties, after fights, after penalties that closely follow fights, and after fights that closely follow penalties, respectively.[25]

Table 4 shows the results from this regression. It provides several insights that a more aggregated analysis cannot offer. First, in the aftermath of a fight, hits by both teams are reduced significantly. Second, penalties also significantly reduce the number of hits for both teams, albeit to a smaller extent. While hits are decreased by 18 percent (0.06 hits per minute) in the five minutes after a fight, they are only decreased by 7.5 percent (0.02 hits per minute) after a penalty.

Third, the timing of the two types of enforcement affect deterrence. Specifically, when a fight happens shortly after a penalty, the deterrent effect of the fight is diminished. Using the coefficients from column (2), a fight that follows a penalty would be associated with 0.0111 (=0.0569-0.0458) fewer hits per minute in the five minutes following that fight. This suggests that if specialized enforcement is present, there is little to be gained from subsequent community enforcement. Penalties, on the other hand, are effective deterrents whether or not they are assessed shortly after a fight, as the coefficient on penalties after fights is not statistically different from zero in any specification.

[25] We chose five-minute periods after running a vector autoregression (VAR) model including lags up to 10 minutes, and interactions of those lags for fights, penalties, and hits. In the VAR regressions, any lagged effect of fights and penalties on hits vanished after five minutes. The VAR results are qualitatively identical, and quantitatively similar, but we report these more aggregated after-fight and after-penalty periods for ease of exposition.

We find a substitution effect between community enforcement and specialized enforcement not previously found in research using aggregated crime and policing data. If a fight closely follows a penalty, hitting does not decrease significantly compared to a penalty which is not followed by a fight. However, if penalties closely follow a fight, the effects of the two types of enforcement are reinforcing. In this case, fights may serve as a "wake up call", which signals to both players and referees that the game has become too aggressive. More penalties are expected because referees pay particularly close attention, and less hitting is expected, both because of the deterrent effect of the fight and penalties. These dynamics are only visible at the disaggregated, minute-level of play, and the importance of disaggregated data in detecting this interaction likely translates to more traditional crime data as well.

### 5.3 The Effect of Enforcement on Individual Behavior

The previous sections document important interactions between specialized enforcement and community enforcement. In particular, the relative timing of the two types of enforcement is important when evaluating their relative effectiveness. The temporally disaggregated nature of the analysis allows us to make causal inferences about the effect of enforcement on anti-social behavior because we observe whether the crime or enforcement occurred first. However, the above analysis still aggregates over all members of the community, even though there may be other forces that affect each individual's behavior.

The results in the previous section may reflect the effect of specific personnel match-ups during games, rather than the effect of fights and penalties. For example, less aggressive players might be systematically utilized immediately after a fight. To account for the possibility of heterogeneous patterns in personnel match-ups, we estimate another set of regressions in which we track individual players' behavior over time and include player-game indicators (as well as minute indicators) to capture unobserved player-level heterogeneity in aggression.

The player-level analysis also allows us to compare the effectiveness of enforcement on different players. It is quite possible that sanctioned individuals - players who are directly involved in fights or penalized - react differently to enforcement actions relative to bystanders. Previous research in crime and economics has established that specific and general deterrence effects can exist when community and specialized enforcement are present (see Benson and Mast (2001) and MacDonald et al. (2012)). However, disentangling the effects of punishment and deterrence has proven challenging (Chalfin and McCrary, 2014). The difficulty in separating these effects is further exacerbated by the presence of order that is driven by social institutions and norms.[26] We overcome these issues by observing the timing of all actions, as well as the identities of those involved.

We estimate Equation (3) at the player-minute-of-play level, with the addition of interaction terms between each enforcement variable and indicator variables identifying whether the player was directly involved in the enforcement, whether he was a teammate of the player who was directly involved, or if he was on the opposing team. The dependent variable is the number of hits a player makes in each minute of play when on the ice. We also include interaction variables for each type of enforcement with an indicator that is turned on if the player has previously been assessed a penalty in the game in order to account for the possibility that an individual who has previously been "caught" may react differently to enforcement than someone who has only observed enforcement.[27]

Table 5 shows the results for these player-level regressions.[28] Considering the individual effects of the types of enforcement, penalties appear to be more effective than fights in deterring anti-social behavior. While each specification features a negative coefficient on the indicator for the five minutes following a fight, this coefficient is small and not statistically significant. This suggests that personnel decisions might play a role in reducing aggression after fights, although this result may be a consequence of the structure of the data: most of the player-minute cells have zero fights.

The effectiveness of a penalty depends on the individual's "proximity" to the penalized player. Penalized individuals significantly reduce subsequent hitting, while teammates of the penalized player actually increase their hitting even after controlling for penalty kill situations. Opponents of the penalized player, on the other hand, decrease hitting slightly.

When analyzing the effect of fights and penalties that closely follow each other, we find the same patterns as reported in section 5.2 above, although the coefficients are not statistically significant.

---

[26] In his seminal work, Ellickson (1991) considers cattle farmers and land owners enforcing social norms. Greif (1993) discusses social norms in a market environment between medieval Maghribi traders and their overseas agents. In more recent work Skarbek (2012) notes that prison gangs can provide deterrence of anti-social behavior, while MacDonald et al. (2012) discuss the role of community policing.

[27] The economics of crime literature notes that perceptions about the likelihood of being detected for criminal activities is significantly influenced by a person's interaction with the legal system. Lochner (2007), analyzing the behavior of young men, notes that those who engage in criminal behavior and are undetected revise their likelihood of being detected downward while those detected revise their probability upward.

[28] No players are observed on the ice during all 60 minutes of the game. Therefore, we only follow those players who were on the ice at any time during a minute of play. We also drop goal keepers as they spend most of their playing time in a small area in front of the goal and are less likely to commit infractions than other players.

When a fight follows a penalty, the effect of the fight is smaller than it is if the fight does not follow a penalty, regardless of who is assessed the penalty. But when a penalty is given to an individual or his teammates after a fight, the individual seems to decrease hitting more.

Finally, prior offenders may have different incentives to commit anti-social acts. They may be more inclined to commit anti-social acts due to their role on the team, or their style of play. Our results show that while prior offenders do not seem to be deterred by fights, their reactions to penalties seem stronger than those of individuals who have not previously been assessed a penalty.

# 6 Placebo Tests and Game Level Analysis

Our analysis shows that reductions in, or the absence of, specialized enforcement leads to increases in decentralized enforcement. However, the substitution between the two types of enforcement is unidirectional since fighting occurs in the absence of penalties, but the reverse does not occur. While fights work better independently of penalties, penalties can deter further anti-social behavior either independently or in conjunction with fights. Moreover, penalties seem to work by deterring the offender and the offender's opponents, while they do not necessarily deter the offender's teammates.

These results have important policy implications, and their validity, as well as the role of the disaggregated data used to generate these results, should be assessed. To this end we first perform placebo tests in which we simulate fights in games that do not experience fights. next, we compare our results to those we would obtain if we only had game-level data, which resembles the temporally aggregated data typically used in empirical economics of crime research.

## 6.1 Placebo Tests

The large number of observations in this data set may generate some concern that we find statistically significant parameter estimates on variables of interest simply by chance. We address this issue by performing placebo tests in which we simulate fights in games where no fights actually occurred.[29] We estimate the temporal distribution of fights across minutes of play by calculating the probability that a fight occurred in each minute of game time for the 30% of games in the sam-

[29] We perform placebo tests on fights because fights occur relatively infrequently; no fights occurred in 69.76 percent of the games in this sample. Penalties are assessed too frequently to identify sufficient time periods without penalties to conduct placebo tests.

ple with an observed fight - roughly 110,000 game-minute observations. We then add a simulated fight to 30% of the fightless games in the sample as a placebo test. The temporal distribution of these placebo fights exactly matches the temporal distribution of actual fights. The placebo test estimates Equation (3) using only data from games with no fights, actual game outcomes (including actual penalties) and the timing of the placebo fights. Since the placebo fights are randomly placed in games where no fight actually occurred, finding a statistically significant effect of placebo fights on subsequent anti-social behavior would suggest that the statistically significant relationships reported above represent a statistical artifact of this very large data set, and not an actual relationship between enforcement actions and anti-social behavior.

Table 6 shows the results of the placebo tests. It has the same general format as Table 4. The coefficient on the post-placebo-fight period, the parameter of primary interest, is not statistically different from zero throughout all model specifications. The fact that we do not see a statistically significant relationship between placebo fights and hitting indicates that the results reported in section 5.2 reflect a true deterrent effect of fights on anti-social behavior. In addition, the effect of (real) penalties on hits in Table 6 remains essentially identical to the results in Table 4, suggesting that penalties and fights operate independently. If their effects were inter-related, or if penalties in games without fights had a systematically different effect on hitting than penalties in games with fights, we would see coefficients that are different from those reported in Table 4.

### 6.2 Game-Level Analysis: Penalties and Fights

Although some change in the level of temporal aggregation of data used in empirical crime and economics research has occurred - especially as micro-level data have become more available - the majority of the empirical research on order with and without law utilizes temporally aggregated data (Durlauf et al., 2010). We now ask what the results would look like if we had conducted the analysis using traditional, temporally aggregated data. In order to answer this question, we aggregate the total number of fights, penalties, and hits to the game level (60 minutes of play).

Figure 4 summarizes the relationship between fighting and penalties in the temporally aggregated data. At the game level, specialized enforcement and community enforcement share a strong, positive relationship, with a correlation coefficient of 0.265. The figure suggests that fights and penalties function as complements, rather than substitutes, in these temporally aggregated data.

Of course, this unconditional correlation omits potentially important factors, as more aggressive games likely feature more penalties *and* more fights. We formally examine the relationship between fights and penalties by estimating an instrumental variables (IV) model explaining variation in fights per game and treating penalties called as an endogenous explanatory variable. In this conditional analysis we exploit the fact that referees have direct control over the penalties called during a game, while their effect on the number of fights is more indirect, working either through penalties or as a consequence of individual referees' reputations among players. Importantly, referee assignments in the NHL are quasi-random, and some referees are naturally more stringent than others in terms of enforcing the rules of hockey as set out in the NHL rulebook.[30]

Our instrument for penalties called is similar to the instrument used in DiTella and Schargrodsky (2013). Since the act of officiating involves considerable physical and mental effort, it is likely that a referee who has officiated many games in a relatively short period of time might experience more mental and physical fatigue, resulting in this referee missing more calls than one who is relatively well-rested.[31] We use the time (in days) between the referees' previous games and the current game as an instrument for the number of penalties called in a game when explaining observed variation in fights.[32] Formally, we estimate the parameter of

$$fights_j = X_j'\beta + \alpha \ln(pen\hat{a}lties_j) + \epsilon_j, \tag{4}$$

where $\ln(pen\hat{a}lties_j)$ is the fitted value of

$$\ln(penalties_j) = X_j'\beta + \sum_{i=1}^{2} \delta_d \mathbb{1}[\text{days off}_{ij} = d] + \nu_j.$$

[30] Although the details of the contract between the NHL and the NHL Officials Association - the referees' union - are not public, it is evident from observed patterns of assignment of referees to games that referees will officiate a small majority of games in a particular region (e.g. the Northeast), but also referee a large number (approximately 20 games) in another region (e.g. the west Coast of the United States). For example, some referees work as few as 7 percent of their total games in the Eastern Conference and others work as many as 82 percent of their games in the Eastern Conference. However, the standard deviation of Eastern Conference games refereed is 19 percent, implying that 64 percent of our referees work between 31 to 69 percent of their assigned games in either the Eastern or Western conference.

[31] Referee fatigue can exists in two ways. First, in 77 percent of games, they officiate most games within a week of each other. Additionally, referees simultaneously observe the movement of the players on the ice and their interactions in order to ensure that behavior is in accordance with the rules while trying to avoid being inadvertently involved with the play by obstructing a player's movement or being struck by the puck.

[32] Each game is officiated by two referees. We keep track of both.

In this regression model, $d$ denotes the number of days between the current game $j$ and the $i^{th}$ referee's last game, and $X_j$ includes each team's shots on goal, goals scored, whether the game went into overtime (was tied after 60 minutes of play), and identifiers for the home and away team (including their standing in their respective conferences), for the season and month of the game, and indicators that identify whether the competing teams play in the same division or conference. We further interact the teams' current conference standings with the month of the game to control for the importance of the game with respect to qualifying for post-season play in each season.

Table 7 shows the results of these IV regression models. The first-stage F-statistic (Stock-Yogo statistic) of 26.37 suggests that our instrument provides sufficient explanatory power.[33] The OLS regression model (columns 1 and 3) suggests that a one-percent increase in penalties assessed leads to a 0.97 percent increase in fights per game. This relationship is similarly strong (but estimated less precisely) in the instrumental variable regression (columns 2 and 4).

Penalties and fights appear complementary when we explore the temporally aggregated data. Recall that the analysis using temporally disaggregated data indicates that the two types of enforcement operate like substitutes. Even with a valid instrument, a traditional data set predicts the wrong sign on the relationship between specialized and decentralized enforcement.

### 6.3 Game-Level Effect of Penalties and Fights on Hits

Even if the game-level analysis predicts the relationship between penalties and fights incorrectly, temporally aggregated data may still be useful if estimates of the effectiveness of the two types of enforcement using these data match the results from temporally disaggregated data. We generate results mirroring the spirit of studies using aggregate data by estimating

$$\ln(hits_j) = X_j'\beta + \alpha \ln(penalties_j) + \sum_{i=1}^{I} \gamma_i \mathbb{1}[fights_j = i] + \sum_{i=1}^{I} \delta_i \ln(penalties_j) \times \mathbb{1}[fights_j = i] + \epsilon_j, \tag{5}$$

where $X_j$ contains the same game-level characteristics as in Section 6.2, and $hits_j$, $penalties_j$, and $fights_j$ describe the total numbers of hits, penalties and fights occurring in game $j$, respectively.

[33] The instrument seems valid. The referees' time off before the game has a statistically significant effect on the number of penalties called in a game, while time off does not significantly affect the number of fights observed during a game. The correlations between days off and penalties called are small but positive at 0.06 (first referee) and 0.07 (second referee). On the other hand, the correlation between days off and fights is even smaller and negative at -0.015 (first referee) and -0.010 (second referee).

Again, OLS regressions likely suffer from endogeneity of the enforcement variables, since particularly aggressive games will have more hits *and* more penalties and fights. This is analogous to the reverse causation problems present in the empirical police-crime literature (Levitt, 2002; DeAngelo and Hansen, 2014). We again use the days since the referees' last game worked to instrument for the number of penalties called in a game. In addition, we use referee identifiers as instruments for the number of fights in the game. Some referees may have a reputation for being tough on "trouble makers," whereas their ability to make correct calls is more likely to vary on a day-to-day basis.

Table 8 contains OLS and estimates of the regression from Equation 5. The OLS results show a statistically significant, negative correlation between penalties and hits, with a doubling of penalties leading to an approximately 4 percent decrease in hits. At the mean, this would imply that an increase from 8 to 16 penalties would lead to a decrease in hits from 43 to 41 hits per game.

While the OLS regressions predict a negative, but small effect of penalties on hits, the IV regressions fail to establish any statistically significant causal relationship between either enforcement proxy and anti-social behavior. The apparent lack of a statistically significant relationship between both fights and penalties at the aggregate game (aggregate) level raises questions about the ability of aggregated data to capture the underlying effect of both law and order on crime; the results in this empirical literature may reflect temporal aggregation problems.

## 7 Policy Implications & Conclusion

Increases in the incidence of anti-social behavior in communities has lead to policy changes that permit individuals to legally protect themselves and others (e.g. "stand your ground" laws), increases in resources devoted to specialized enforcement, and other important policy responses. Whether specialized enforcement, community enforcement, or some combination of the two would be the most effective means of reducing anti-social behavior remains an open and important question. In this work, we advance this line of research by empirically analyzing the determinants of anti-social behavior in an environment where both types of enforcement exist in the same local area.

Standard, publicly available crime data tend to suffer from a lack of the necessary granularity. Even though recent innovations have made data available at temporally disaggregated levels, we do not usually observe information on both specialized and community enforcement activities, or of undetected anti-social behavior. To overcome this, we use data on the enforcement of player conduct in NHL games, and argue that it is possible, and appropriate, to draw broad conclusions from results generated from these data. Specifically, community enforcement acts as a substitute for specialized enforcement when specialized enforcement falls below established expectations or norms, which can be viewed as a monitoring imperfection on the part of specialized enforcers. Additionally, specialized enforcement acts as a complement to community enforcement once it has been established. In the context of hockey, this is visible in the aftermath of a fight, when penalties significantly reduce hitting beyond the impact attributable to fights.

The results in this paper provide new insights in a growing area of research in law and economics. The interaction between public and private law enforcement has garnered substantial attention in the literature, motivated by widely reported declines in confidence in public enforcement of safety. The general lack of data reflecting both private enforcement and public enforcement complicates empirical analyses on the interaction between these two types of enforcement. Attempts to systematically gather information about specialized enforcement has been piecemeal, and generally based on highly aggregated data, which could generate misleading conclusions.

Although this research focuses on a specific environment, we obtain clean measures of community and specialized enforcement actions as well as anti-social behavior in this setting. When temporally aggregated, our data tell a story similar to previous research – specialized and community enforcement are complements, and not particularly effective. When temporally disaggregated, our data indicate that these enforcement efforts share a more complicated relationship that depends on the specific order of occurrence. This highlights the importance of identifying temporally and spatially disaggregated data sources that contain information on both specialized and community enforcement for use in future research.

## References


Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica: Journal of the Econometric Society*, pages 383–396.

Acemoglu, D. and Jackson, M. O. (2014). Social norms and the enforcement of laws. Technical report, National Bureau of Economic Research.

Acemoglu, D. and Wolitzky, A. (2015). Sustaining cooperation: Community enforcement vs. specialized enforcement. Technical report, National Bureau of Economic Research.

Allen, W. D. (2002). Crime, punishment, and recidivism lessons from the national hockey league. *Journal of Sports Economics*, 3(1):39–60.

Becker, G. S. (1968). Crime and punishment: An economic approach. *The Journal of Political Economy*, 76(2):169–217.

Bennett, T., Holloway, K., and Farrington, D. (2008). The effectiveness of neighborhood watch. *The Campbell Collaboration*.

Benson, B. L. and Mast, B. D. (2001). Privately Produced General Deterrence. *Journal of Law and Economics*, 44(2):725–46.

Bicchieri, C. and Muldoon, R. (2011). Social norms. *Stanford Encyclopedia of Philosophy*.

Chalfin, A. and McCrary, J. (2014). Criminal deterrence: A review of the literature. Technical report, Working paper.

Cook, P. J. and MacDonald, J. (2011). Public safety through private action: an economic assessment of bids. *The Economic Journal*, 121(552):445–462.

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):231–257.

DeAngelo, G. and Smith, L. T. (2015). Private security and the provision of international public goods. *Working Paper*.

Deb, J. and González-Dıaz, J. (2014). Enforcing social norms: Trust-building and community enforcement. Technical report.

DiTella, R. and Schargrodsky, E. (2013). Criminal recidivism after prison and electronic monitoring. *Journal of Political Economy*, 121(1):28–73.

Donaldson, L., Ashbridge, M., and Cusimano, M. D. (2013). Bodychecking rules and concussion in elite hockey. *Plos One*.

Durlauf, S. N., Navarro, S., and Rivers, D. A. (2010). Understanding aggregate crime regressions. *Journal of Econometrics*, 158(2):306–317.

Ellickson, R. C. (1991). *Order Without Law*. Harvard University Press.

Ellison, G. (1994). Cooperation in the prisoner's dilemma with anonymous random matching. *The Review of Economic Studies*, 61(3):567–588.

Gopnik, A. (2012). Hockey without rules. `http://www.newyorker.com/news/sporting-scene/hockey-without-rules`.

Greif, A. (1993). Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition. *The American economic review*, pages 525–548.

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.

Hume, D. (1793). *A Treatise of Human Nature: Being an Attempt to Instroduce Experimental Method of Reasoning into Moral Subjects*. John Noon: London.

Lebrun, P. (2013). Jarome Iginla makes case for fighting. `http://espn.go.com/boston/nhl/story/_/id/9981354/jarome-iginla-fighting-makes-hockey-safer`.

Leeson, P. T. (2007). An-arrgh-chy: The law and economics of pirate organization. *Journal of Political Economy*, 115(6):1049–1094.

Leeson, P. T. (2009). The laws of lawlessness. *The Journal of Legal Studies*, 38(2):471–503.

Levitt, S. D. (2002). Testing the Economic Model of Crime: The National Hockey League's Two-Referee Experiment. *The B.E. Journal of Economic Analysis & Policy*, 1(1):1–21.

Lochner, L. (2007). Individual perceptions of the criminal justice system. *The American Economic Review*, 97(1):444–460.

MacDonald, J., Klick, J., and Grunwald, B. (2012). The effect of privately provided police services on crime. *U of Penn, Inst for Law & Econ Research Paper*, (12-36).

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

Nagin, D. S. (2013). Deterrence: A review of the evidence by a criminologist for economists. *Annu. Rev. Econ.*, 5(1):83–105.

Paul, R. J. (2003). Variations in NHL attendance: the impact of violence, scoring, and regional rivalries. *American Journal of Economics and Sociology*, 62(2):345–364.

Peltzman, S. (1975). The effects of automobile safety regulation. *The Journal of Political Economy*, pages 677–725.

Roth, G. and Skarbek, D. (2014). Prisons gangs and the community responsibility system. *Review of Behavioral Economics*, 1:223–243.

Skarbek, D. (2012). Prison gangs, norms, and organizations. *Journal of Economic Behavior & Organization*, 82(1):96–109.

Smith, B. (2016). Enfd to fighting would not make hockey a safer game. *The Globe and Mail*.

Sobel, R. and Osoba, B. J. (2009). Youth gangs as pseudo-governments: Implications for violent crime. *Southern Economic Journal*, 75(4):996–1018.

Stone, A. (2015). Montreal Canadiens' Brandon Prust: 'the NHL needs fighting to keep the game safe'. `http://ftw.usatoday.com/2015/02/montreal-canadiens-brandon-prust-the-nhl-needs-fighting-to-keep-the-game-safe`.

Tiebout, C. M. (1956). A pure theory of local expenditures. *The journal of political economy*, pages 416–424.

Tosi, E. (2013). Jarome iginla: Hockey is better, safer with fighting in it. *Sports Illustrated*.

Vigneault, M. (2011). *La Naissance D'Un Sport Organise au Canada: Le Hockey a Montreal, 1875-1917*. PhD thesis, Universite Laval Quebec.

Table 1: Summary statistics, events per game

| Event | Observations | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Penalties | 6,859 | 8.840 | 3.839 | 0 | 49 |
| Fights | 6,859 | 0.529 | 0.805 | 0 | 8 |
| Hits | 6,859 | 43.444 | 13.390 | 0 | 104 |
| Goals | 6,859 | 5.771 | 2.368 | 0 | 17 |
| Shots | 6,859 | 54.656 | 9.030 | 0 | 102 |
| Injuries | 6,859 | 0.249 | 0.506 | 0 | 4 |
| Face-offs | 6,859 | 57.402 | 7.364 | 1 | 85 |
| All Events | 6,859 | 297.716 | 33.064 | 1 | 427 |
| Time between events (seconds) | 2,042,037 | 12.682 | 1.605 | 0 | 22.120 |

Table 2: Minute-level analysis: fights after penalties

| | (1) | (2) | (3) |
|---|---|---|---|
| | Pr(fight) | Pr(fight) | Pr(fight) |
| Penalties$_{t-1}$ | -0.00671$^{***}$ | -2.025$^{***}$ | -0.709$^{***}$ |
| | (0.000380) | (0.114) | (0.0396) |
| Penalties$_{t-2}$ | -0.00271$^{***}$ | -0.671$^{***}$ | -0.232$^{***}$ |
| | (0.000326) | (0.0695) | (0.0269) |
| Penalties$_{t-3}$ | -0.000118 | -0.0499 | -0.0155 |
| | (0.000314) | (0.0473) | (0.0206) |
| Penalties$_{t-4}$ | -0.000505 | -0.0824 | -0.0308 |
| | (0.000317) | (0.0506) | (0.0217) |
| Penalties$_{t-5}$ | 0.000274 | 0.0222 | 0.0122 |
| | (0.000317) | (0.0465) | (0.0204) |
| $N$ | 354671 | 103598 | 102576 |
| adj. $R^2$ | -0.015 | | |

Robust standard errors clustered at the rivalry level in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Minute level controls include log hits, power play indicators, injuries, differences in shots and goals between the teams, the players' average time on ice (in total, short handed, and in power play), and minute indicators. We include game fixed effects. Column 1 estimates a linear probability model, column 2 shows logit results, and column 3 describes Probit results.

Table 3: Minute-level analysis: penalties committed after a fight

| | (1) penalties | (2) log(penalties) | (3) penalties |
|---|---|---|---|
| 1st min after | 0.429*** | 1.312*** | 1.327*** |
| | (0.0285) | (0.0792) | (0.0426) |
| 2nd min after | -0.0375*** | -0.235*** | -0.226*** |
| | (0.0111) | (0.0563) | (0.0785) |
| 3rd min after | -0.00124 | 0.00587 | 0.0398 |
| | (0.0103) | (0.0549) | (0.0703) |
| 4th min after | -0.00650 | -0.0140 | -0.0181 |
| | (0.0105) | (0.0599) | (0.0679) |
| 5th min after | -0.0171 | -0.132** | -0.0906 |
| | (0.0109) | (0.0560) | (0.0786) |
| $N$ | 326179 | 326179 | 325333 |
| adj. $R^2$ | 0.083 | 0.100 | |

Robust standard errors clustered at the rivalry level in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Minute level controls include log hits, power play indicators, injuries, differences in shots and goals between the teams, the players' average time on ice (in total, short handed, and in power play), and minute indicators. We include game fixed effects and cluster standard errors at the rivalry level. Column 1 reports the number of penalties per minute in levels. We log the dependent variable in column 2 for ease of interpretation, and we estimate a Poisson regression in column 3.

Table 4: Minute-level analysis: effectiveness of penalties and fights in deterring hits

| | (1) Hits | (2) Hits | (3) log(Hits) | (4) Hits |
|---|---|---|---|---|
| 5 minutes after fight | -0.0673$^{***}$ | -0.0569$^{***}$ | -0.203$^{***}$ | -0.0820$^{***}$ |
| | (0.0104) | (0.0112) | (0.0469) | (0.0158) |
| 5 minutes after fight (post-penalty) | 0.0549$^{***}$ | 0.0458$^{***}$ | 0.147$^{**}$ | 0.0594$^{**}$ |
| | (0.0157) | (0.0166) | (0.0697) | (0.0241) |
| 5 minutes after penalty | -0.0167$^{***}$ | -0.0192$^{***}$ | -0.0778$^{***}$ | -0.0373$^{***}$ |
| | (0.00314) | (0.00455) | (0.0195) | (0.00687) |
| 5 minutes after penalty (post-fight) | -0.00758 | -0.0125 | -0.0384 | -0.0174 |
| | (0.0103) | (0.0107) | (0.0453) | (0.0164) |
| power play | -0.554$^{***}$ | -0.522$^{***}$ | -2.188$^{***}$ | -0.965$^{***}$ |
| | (0.00408) | (0.00464) | (0.0209) | (0.00889) |
| Lagged hits | | ✓ | ✓ | ✓ |
| Lagged power play | | ✓ | ✓ | ✓ |
| TOI controls | ✓ | ✓ | ✓ | ✓ |
| Injury controls | ✓ | ✓ | ✓ | ✓ |
| Score controls | ✓ | ✓ | ✓ | ✓ |
| $N$ | 386981 | 354671 | 354671 | 354671 |
| adj. $R^2$ | 0.077 | 0.075 | 0.071 | |

Robust standard errors clustered at the rivalry level in parentheses; *** p<0.01, ** p<0.05, * p<0.1. We include game fixed effects and minute indicators. Columns 1 and 2 report results for an OLS regression with levels of hits per minute as the dependent variable, Column 3 uses log-hits as the dependent variable, and column 4 estimates a Poisson regression as a robustness check.

Table 5: Player-level analysis: effectiveness of penalties and fights

| | (1) Player hits | (2) Player hits | (3) Player hits | (4) log(Player hits) | (5) Player hits |
|---|---|---|---|---|---|
| 5 min after fight | -0.000155 | -0.000973 | -0.000994 | -0.00694 | -0.0199 |
| | (0.000516) | (0.000650) | (0.000674) | (0.00447) | (0.0150) |
| 5 min after fight (prior offender) | | | 0.000184 | 0.00294 | 0.00791 |
| | | | (0.00254) | (0.0168) | (0.0555) |
| Own penalty | -0.00360*** | -0.00351*** | -0.00319*** | -0.0210*** | -0.0682*** |
| | (0.000731) | (0.000744) | (0.000757) | (0.00502) | (0.0170) |
| Teammate penalty | 0.00146*** | 0.00146*** | 0.00132*** | 0.00860*** | 0.0302*** |
| | (0.000222) | (0.000225) | (0.000233) | (0.00154) | (0.00531) |
| Teammate penalty (prior offender) | | | 0.00149** | 0.00908** | 0.0362** |
| | | | (0.000678) | (0.00449) | (0.0153) |
| Opponent penalty | -0.00162*** | -0.00164*** | -0.00151*** | -0.00975*** | -0.0327*** |
| | (0.000219) | (0.000221) | (0.000231) | (0.00153) | (0.00534) |
| Opponent penalty (prior offender) | | | -0.00112* | -0.00781** | -0.0135 |
| | | | (0.000582) | (0.00386) | (0.0141) |
| Fight after own pen. | | 0.00676 | 0.00701 | 0.0375 | 0.122 |
| | | (0.00635) | (0.00669) | (0.0434) | (0.139) |
| Fight after teammate pen. | | 0.00240* | 0.00186 | 0.0157* | 0.0403 |
| | | (0.00124) | (0.00130) | (0.00871) | (0.0299) |
| Fight after teammate pen. (prior offender) | | | 0.00504 | 0.0290 | 0.107 |
| | | | (0.00424) | (0.0281) | (0.0909) |
| Fight after opponent pen. | | 0.00105 | 0.000903 | 0.00580 | 0.0162 |
| | | (0.00122) | (0.00129) | (0.00857) | (0.0302) |
| fight after opponent pen. (prior offender) | | | 0.00117 | 0.00681 | 0.00912 |
| | | | (0.00409) | (0.0270) | (0.0910) |
| Own pen. after fight | | -0.00333 | -0.00317 | -0.0151 | -0.0717 |
| | | (0.00373) | (0.00381) | (0.0257) | (0.0872) |
| Teammate pen. after fight | | -0.000417 | -0.0000545 | -0.00118 | -0.00457 |
| | | (0.000990) | (0.00104) | (0.00688) | (0.0243) |
| Teammate pen. after fight (prior offender) | | | -0.00415 | -0.0324 | -0.0816 |
| | | | (0.00346) | (0.0225) | (0.0814) |
| Opponent pen. after fight | | 0.000479 | 0.000697 | 0.00520 | 0.0230 |
| | | (0.000927) | (0.000977) | (0.00652) | (0.0243) |
| Opponent pen. after fight (prior offender) | | | -0.00215 | -0.0123 | -0.0545 |
| | | | (0.00306) | (0.0205) | (0.0763) |
| Power play | -0.0257*** | -0.0257*** | -0.0257*** | -0.172*** | -0.999*** |
| | (0.000271) | (0.000271) | (0.000271) | (0.00182) | (0.0127) |
| Penalty kill | -0.0182*** | -0.0181*** | -0.0181*** | -0.120*** | -0.458*** |
| | (0.000350) | (0.000350) | (0.000350) | (0.00234) | (0.00984) |
| Score & shots | ✓ | ✓ | ✓ | ✓ | ✓ |
| Injuries | ✓ | ✓ | ✓ | ✓ | ✓ |
| TOI | ✓ | ✓ | ✓ | ✓ | ✓ |
| $N$ | 6213452 | 6213452 | 6213452 | 6213452 | 3853623 |
| adj. $R^2$ | 0.002 | 0.002 | 0.002 | 0.002 | |

Robust standard errors clustered at the player-game level in parentheses; *** p<0.01, ** p<0.05, * p<0.1. We include player-game fixed effects and minute indicators. Columns 1 - 3 report results for an OLS regression with levels of hits per minute as the dependent variable, Column 4 uses log-hits as the dependent variable, and column 5 estimates a Poisson regression as a robustness check.

Table 6: Minute-level analysis: effectiveness of penalties and fights, placebo tests

| | (1)<br>Hits | (2)<br>Hits | (3)<br>log(Hits) | (4)<br>Hits |
|---|---|---|---|---|
| 5 minutes after placebo fight | -0.0101 | -0.00811 | 0.00486 | -0.0113 |
| | (0.0149) | (0.0163) | (0.0676) | (0.0221) |
| 5 minutes after placebo fight (post-penalty) | 0.0167 | 0.0103 | 0.0319 | 0.0166 |
| | (0.0222) | (0.0235) | (0.0969) | (0.0323) |
| 5 minutes after any penalty | -0.0176*** | -0.0186*** | -0.0771*** | -0.0369*** |
| | (0.00385) | (0.00567) | (0.0243) | (0.00858) |
| 5 minutes after penalty (post-placebo fight) | -0.00272 | -0.00250 | -0.0264 | 0.00606 |
| | (0.0146) | (0.0150) | (0.0647) | (0.0225) |
| 5 minutes after power play | -0.553*** | -0.526*** | -2.204*** | -0.977*** |
| | (0.00501) | (0.00574) | (0.0259) | (0.0111) |
| Lagged hits | | ✓ | ✓ | ✓ |
| Lagged powerplay | | ✓ | ✓ | ✓ |
| TOI controls | ✓ | ✓ | ✓ | ✓ |
| injury controls | ✓ | ✓ | ✓ | ✓ |
| score controls | ✓ | ✓ | ✓ | ✓ |
| $N$ | 254834 | 233584 | 233584 | 233584 |
| adj. $R^2$ | 0.076 | 0.074 | 0.070 | |

Robust standard errors clustered at the rivalry level in parentheses; *** p<0.01, ** p<0.05, * p<0.1. We include game fixed effects and minute indicators. Columns 1 and 2 report results for an OLS regression with levels of hits per minute as the dependent variable, Column 3 uses log-hits as the dependent variable, and column 4 estimates a Poisson regression as a robustness check.

Table 7: Game-level effects of penalties on fights

| | OLS | IV | OLS | IV |
|---|---|---|---|---|
| | ln(fights) | ln(fights) | total fights | total fights |
| ln(penalties) | 0.973*** | 1.061 | 0.286*** | 0.248 |
| | (0.103) | (0.581) | (0.0294) | (0.127) |
| Teams in same division | 0.0809 | 0.0988 | 0.0363 | 0.0413 |
| | (0.0937) | (0.0942) | (0.0221) | (0.0221) |
| Teams in same conference | 0.0778 | 0.0677 | 0.0384 | 0.0406 |
| | (0.115) | (0.116) | (0.0251) | (0.0253) |
| Goals scored (home team) | 0.0966*** | 0.0926** | 0.0369*** | 0.0380*** |
| | (0.0255) | (0.0317) | (0.00661) | (0.00779) |
| Goals scored (away team) | 0.0281 | 0.0329 | 0.00252 | 0.00374 |
| | (0.0263) | (0.0270) | (0.00615) | (0.00639) |
| Shots (home team) | -0.00473 | -0.00571 | -0.00199 | -0.00181 |
| | (0.00651) | (0.00693) | (0.00151) | (0.00154) |
| Shots (away team) | -0.0147* | -0.0136* | -0.00324* | -0.00295* |
| | (0.00671) | (0.00666) | (0.00149) | (0.00147) |
| Overtime game | -0.374*** | -0.378*** | -0.103*** | -0.108*** |
| | (0.102) | (0.113) | (0.0217) | (0.0244) |
| $N$ | 6856 | 6799 | 6856 | 6799 |
| adj. $R^2$ | 0.083 | 0.083 | 0.106 | 0.106 |
| First stage | | | | |
| F-stat | | 26.366 | | 26.366 |
| Adj. $R^2$ | | 0.116 | | 0.116 |

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls include home and away team dummies, season and month dummies, as well as the home and away teams' standings within their conference, and the interaction between their standings and the month within the season.

Table 8: Game-level effectiveness of penalties and fights in deterring hits

| | OLS | | IV | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | log(hits) | log(hits) | log(hits) | log(hits) |
| log(penalties) | -0.0458*** | -0.0397*** | -0.0231 | 0.0321 |
| | (0.00626) | (0.00729) | (0.0331) | (0.0463) |
| 1 fight in game | 0.00589 | 0.0369 | 0.0614 | 0.373 |
| | (0.00719) | (0.0309) | (0.0385) | (0.199) |
| 2 fights in game | 0.0214 | 0.0802 | 0.00479 | 0.541 |
| | (0.0116) | (0.0527) | (0.0634) | (0.318) |
| 3 fights in game | 0.0818*** | 0.0529 | 0.258* | 0.244 |
| | (0.0220) | (0.128) | (0.124) | (1.052) |
| 1 fight×ln(penalty) | | -0.0150 | | -0.142 |
| | | (0.0145) | | (0.0911) |
| 2 fights×ln(penalty) | | -0.0132 | | -0.117 |
| | | (0.0117) | | (0.0658) |
| 3 fights×ln(penalty) | | 0.00380 | | 0.00185 |
| | | (0.0172) | | (0.147) |
| Controls | ✓ | ✓ | ✓ | ✓ |
| $N$ | 6856 | 6856 | 6799 | 6799 |
| adj. $R^2$ | 0.419 | 0.419 | - | - |

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Controls include home and away team dummies, season and month dummies, as well as the home and away teams' standings within their conference, and the interaction between their standings and the month within the season.

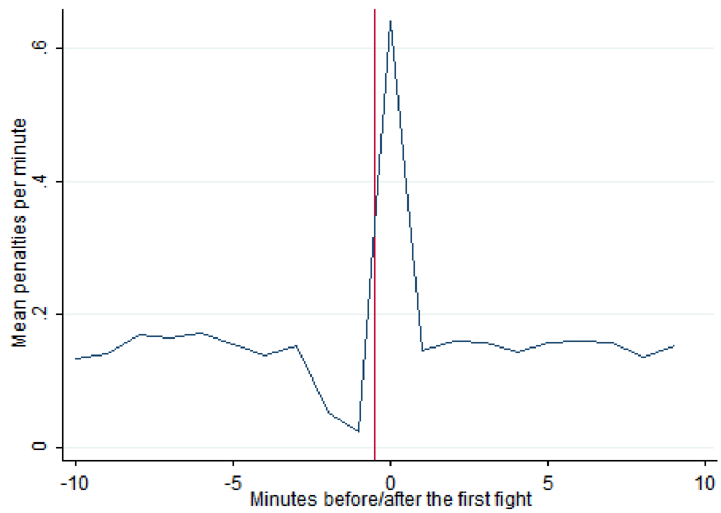


Figure 1: Average penalties called per minute before and after fights

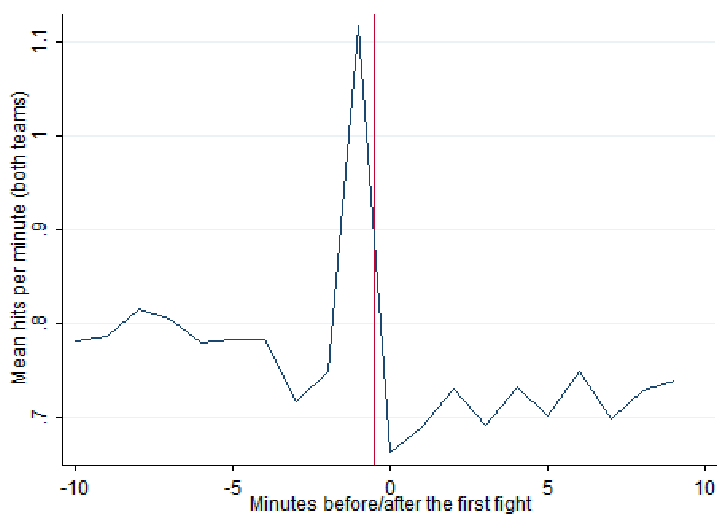


Figure 2: Average number of hits per minute before and after a fight

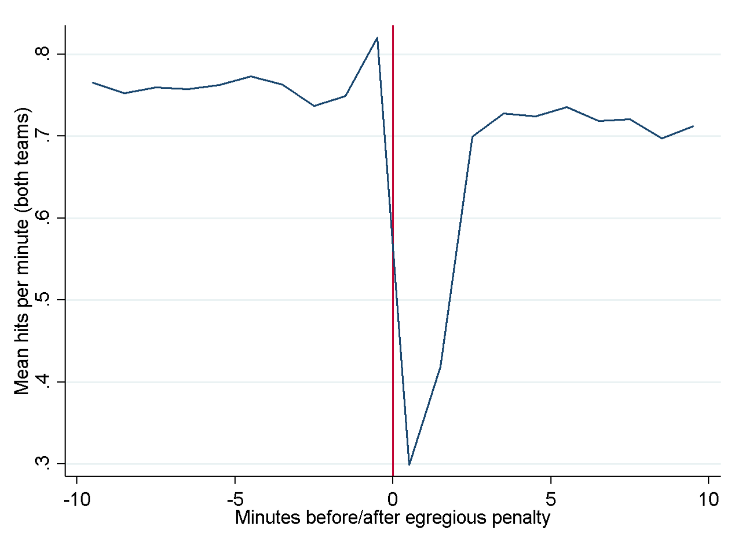


Figure 3: Average number of hits per minute before and after an egregious penalty

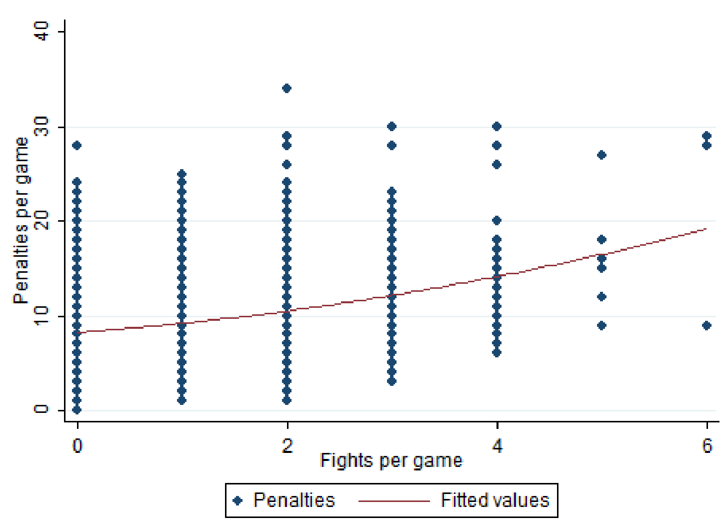


Figure 4: Game level correlation between fights and penalties