10-25-2022

A multi-site study of firearms displays by police at use of force incidents

Timothy Cubitt
University of Tasmania

Justin Nix
University of Nebraska at Omaha, jnix@unomaha.edu

Recommended Citation

Cubitt, T. & Nix, J. (2022). A multi-site study of firearms displays by police at use of force incidents. Police Quarterly.

A multi-site study of firearms displays by police at use of force incidents

Forthcoming at *Police Quarterly*[*]

Timothy Cubitt

University of Tasmania

Justin Nix

University of Nebraska Omaha

Abstract

The power to use force is a defining characteristic of policing, one that is accompanied by a responsibility to exercise these powers in the circumstances deemed necessary. This study analyzes data from four policing agencies to predict the likelihood of an officer drawing and pointing their firearm at a use of force incident. Findings suggest that situational factors were important in influencing whether an officer may draw and point their firearm. However, a priming effect, in which officers were more likely to draw their firearms when dispatched to an incident, may also be present. The rate that officers drew and pointed their firearms varied between jurisdictions, as did the nature of the incidents. Caution should be exercised in generalizing the results of single-site studies on police use of force, or introducing research into policy beyond the jurisdiction in which it was performed.

Keywords: use of force, firearms, firearm display, priming, external validity

Understanding the circumstances that give rise to police use of force (UOF) is paramount, as is understanding the tactical decisions officers make during UOF incidents. Fortunately, American policing in the 21st Century is much better at tracking uses of force and being transparent with the resulting data than in the past (Garner et al., 2018; Pate & Fridell, 1995; Shjarback, 2019). The push for more "community-oriented" policing, the diffusion of the Internet, and in some cases, federal consent decrees have encouraged agencies to post incident-level UOF data to their websites, where researchers are free to download and test their hypotheses (see e.g., Chanin & Courts, 2017; Matusiak et al., 2022; Rosenbaum et al., 2011). These, alongside increased interest in police UOF after Michael Brown was killed in Ferguson (President's Task Force on 21st Century Policing, 2015), have led to an explosion of research on the topic in the last decade.

Police are authorized to use the amount of force necessary to compel compliance or to defend themselves. This is in fact the defining characteristic of the police function (Bittner, 1970). Consequently, the public expects agencies to document when officers exercise this authority. Unfortunately, use of force definitions vary considerably across states and local jurisdictions, as do the behaviors considered a "reportable" use of force (Stoughton et al., 2020). For instance, according to the most recent estimates, only 54% of all local agencies require officers to document when they point their firearms at people (Brooks, 2020). While pointing a firearm is a UOF, it stands apart from other tactics (e.g., control techniques, TASERs, OC spray) because officers are neither trained nor expected to point their firearms to compel compliance or prevent suspects from fleeing. Instead, firearms are to be drawn when officers reasonably believe they might need to discharge them to respond to or prevent an imminent, deadly attack.[1]

[1] For example, the Los Angeles Police Department's policy (revised June 29, 2020) states: "Officers shall not draw or exhibit a firearm unless the circumstances surrounding the incident create a reasonable belief that it may be necessary to use the firearm. When an officer has determined that the use of deadly force is not necessary, the

Decades of research and dozens of studies have examined the correlates and causes of UOF in a general sense (Bolger, 2015), as well as specific types of force such as the TASER (Bishopp & Klinger, 2015; White & Ready, 2007) or OC spray (Kaminski et al., 1998; Smith & Alpert, 2000). This scholarship suggests that various situational, environmental, suspect, and officer characteristics may influence the likelihood of police using force. However, far less is known about the circumstances that give rise to when officers point *but don't shoot* their firearms. For example, officers' preoccupation with danger (Skolnick, 1966; Sierra-Arévalo, 2021) might make them more prone to draw and point their firearms when performing certain activities (e.g., serving warrants, making arrests) or being dispatched to certain calls for service (e.g., domestic disputes; see Nix et al., 2021). Meanwhile, the science of implicit biases and stereotype threats raises concerns that a suspect's race may influence officers' decisions to draw and point their firearms (James et al., 2016; Kahn et al., 2016; Smith & Alpert, 2007; Trinkner et al., 2019). And finally, the "policing as a craft" argument (Bayley & Bittner, 1984) suggests that with experience, officers become more adept at dealing with people and resolving conflicts without relying on coercion (Bayley & Garofalo, 1989; Paoline & Terrill, 2007). Thus, inexperienced officers might be more likely to draw and point their firearms than those with more years on the job.

Comparatively less empirical attention has been given to the factors influencing officers' decisions to draw and point their firearms, but the studies we do have focus primarily on suspect and officer characteristics, such as race (Stansfield et al., 2021; Ridell & Worrall, 2021; Worrall et al., 2021) or the decision to pull the trigger once the firearm has been drawn (Wheeler et al., 2017; Worrall et al., 2018). Two studies also suggest that requiring officers to report when they

officer shall, as soon as practicable, secure or holster the firearm…Moreover, any intentional pointing of a firearm at a person by an officer shall be reported." See "Drawing or Exhibiting Firearms" at https://www.lapdonline.org/newsroom/policy-on-the-use-of-force-revised/.

display their firearms is associated with a reduction in police shooting rates (Jennings & Rubado, 2017; Shjarback et al., 2021).

One thing clear from these studies is that officers frequently draw and point *but don't shoot* their firearms. For example, in New Orleans between 2016 and 2019, Riddell & Worrall (2021) found that officers pointed their firearms during 2,116 "response to resistance" incidents – 50% of the total (N=4243). In comparison, officers pointed their conducted energy weapon (CEW) in only 5% of such incidents. However, less than 1% of officers who pointed their firearms ultimately discharged them; meanwhile, 44% of those who pointed their CEW ultimately discharged them. In a separate study, Wheeler et al. (2017) found that Dallas police officers were involved in 207 shootings over a 14-year period (~15 per year on average) compared to 1,702 incidents where they pointed their guns but did not shoot in the first 4 years after a reporting requirement was instituted (~425 per year on average). This suggests Dallas cops pulled the trigger upon drawing and pointing their firearms about 3% of the time. Granted, these are just two departments among thousands, but for every 100 times officers in these agencies draw and point their guns, they ultimately *do not* pull the trigger somewhere between 97 and 99 times. This suggests there may be a glaring hole in our collective knowledge with respect to police use of deadly force – and in particular, understanding the factors associated with officers' *restraint* in the use of deadly force.[2]Another thing clear from emerging scholarship is that studies usually focus on single agencies. Yet, it is well-established that agencies differ in terms of their organizational culture (Chan, 1996; Wilson, 1967). Furthermore, studies consistently reveal that administrative policies vary across

---

[2] For exceptions, see a series of simulated experiments by Lois James and colleagues (James et al., 2013, 2014, 2016, 2018a, 2018b). See also Fryer (2019). To better understand the factors associated with police shootings in Houston, Fryer compiled a random sample of arrest codes "in which lethal force is more likely to be justified: attempted capital murder of a public safety officer, aggravated assault on a public safety officer, resisting arrest, evading arrest, and interfering in an arrest" (p. 1213). This sample is assumed by the author to reflect reasonably the universe of incidents wherein officers exercised restraint in using deadly force.

organizations, and significantly influence officer decision-making in the field (Fyfe, 1988; Terrill & Paoline, 2017; White, 2001). In their study of three police departments, Terrill and Paoline (2017) demonstrated "that officers working within the most restrictive policy framework used force less readily than officers who operated within more permissive policy environments" (p. 193). In a separate analysis of use of force behaviors among officers in six departments, Paoline et al. (2021) documented a great deal of variation in use of force rates – irrespective of the benchmark used (contacts, arrests, or Part I crimes). And here again, agencies with more restrictive UOF policies exhibited lower UOF rates. Thus, there is good reason to expect significant variation across agencies in the rate at which officers draw and point their firearms.

The purpose of our study is to shed additional light on the factors associated with officers' decisions to draw and point their firearms. To do so, we build on the existing literature by compiling open-source data from four urban police departments – Austin (Texas), Baltimore (Maryland), Dallas (Texas), and Portland (Oregon) – and use machine learning techniques, alongside more traditional statistical modelling, to predict the likelihood of any UOF incident involving an officer drawing and pointing their firearm. The use of multiple agencies' data is a contribution in and of itself, as most prior work on this topic has been focused on a single department (often Dallas). We note that, to our knowledge, no published study has analyzed data from Austin, Baltimore, or Portland. Likewise, our use of machine learning techniques makes a methodological contribution to the police UOF literature. Before we elaborate on our data and methods, we briefly discuss the use of machine learning in policing.

Machine Learning in Policing

Machine learning analytics may be better matched to the complexity of the decision-making processes reflected in criminal justice data, and complexity of data ensemble in administrative data, than generalized linear modelling (Brennan & Oliver, 2013). Although machine learning techniques have been used to predict criminal offending, and even rare violent events, with reasonable success (Berk, 2017; Berk et al., 2016; Berk & Sorenson, 2020), there have been very few applications to police behavior. Emerging literature proposes that machine learning is a powerful tool for developing risk assessments (Berk, 2021), but it remains that stratifying behavior for analysis provides considerable insight into those specific behaviors. In this case, prediction of drawing and pointing firearms may contribute another important part of the deadly force story.

Machine learning analytics have been used in recent years to interrogate policing data with considerable accuracy (Berk et al., 2009; Berk et al., 2016; Cubitt et al., 2020). For example, analytical processes of this type have been used to forecast domestic violence (Berk, 2019; Berk et al., 2016; Grogger et al., 2021) and high-harm offense types (Berk et al., 2009; Cubitt & Morgan, 2022). Machine learning, while remaining an underutilized analytical methodology in the field of policing, offers a viable alternative to generalized linear modelling, by allowing data to be interrogated with greater granularity.

While this remains an emerging analytical approach, there are some concerns relating development of prediction models, including the importance of data quality, and the need to be discerning in training models (Bennett Moses & Chan, 2018). Importantly, ethical and jurisprudential concerns have been identified (Berk et al., 2018; Coglianese & Lehr, 2017), largely centering on the suggestion of mindfulness regarding the decisions computed by these models, particularly in complex systems like criminal justice. However, these techniques have proliferated in recent years and are increasingly applied to policing data, often outperforming traditional analytics (Couronné et al., 2018; Grogger et al., 2021). While transparency in the decisions computed by these models is essential (McKay, 2019), the primary risks associated with machine learning approaches and applications appear to be misclassification of individuals and blind adherence to computed models (Ridgeway, 2013), although the implementation process also bears consideration. Regardless of the quality of the data, or strength of the modelling processes, poor implementation strategies may undermine the analytics, a factor which requires consideration in any analytical technique (Stevenson & Doleac, 2021).

The Present Study

An array of literature considers UOF (Bolger, 2015), and there is considerable research focused on the UOF involving firearms (Nix et al., 2017; Nix & Shjarback, 2021; VerBruggen, 2022). Comparatively less research considers the factors associated with officers drawing and pointing their firearms, but not subsequently discharging them (Ridell & Worrall, 2021; Wheeler et al., 2017; Worrall et al., 2018, 2021). This is a serious use of police authority, and may offer notable insight into the behavior of officers at UOF incidents. As a result, in undertaking this analysis we intend to answer two key questions:

1. Is it possible to predict which use of force incidents will feature a patrol officer drawing and pointing, but not discharging their firearm?
2. What are the contributory factors to patrol officers drawing and pointing, but not discharging their firearms at use of force events among these data?

Data Description

As a feature of operational procedures, and intending to improve transparency in policing, departments frequently collect and publish data relating to UOF incidents. These data consist of a range of features, commonly including the reason provided by officers for the UOF, their assignment, the service type undertaken at the time that force was used, whether an arrest was made, whether an offense occurred and the offense type, whether an officer was injured and severity of the injury, demographics of the individual subjected to UOF, and certain demographics of the officer involved. Pivotally, some police departments (PD) make the types of force used by officers available among their data. To improve the validity and generalizability of findings, it was important to consider more than one PD. In light of prior work documenting organizational differences in police culture (Chan, 1996; Wilson, 1967) and the effects of administrative policies on UOF behaviors (Fyfe, 1988; Terrill & Paoline, 2017; White, 2001), focusing on a single jurisdiction may provide insight relative to the jurisdiction itself, while limiting the generalized insight in relation to UOF incidents,

The selection of PDs for analysis revealed important features of UOF reporting, and important barriers to the analysis of these data. It was immediately evident that PDs did not commonly report the same set of information, and when they did, it was often structured differently between jurisdictions. As a result, this research required data reported through similar channels (open data portals), from PDs that featured similar reporting protocols. Four PDs were selected, and while these departments reported an array of data, there were several essential variables that were available for each jurisdiction. This research therefore considers only the variables that were consistent between these four PDs, and their contribution to the likelihood that a patrol officer may draw and point their firearm as a means of control at a UOF event.

Data were obtained for the Austin PD (City of Austin Open Data Portal, 2022), Baltimore PD (Project Comport, 2019), Dallas PD (Dallas Open Data, 2022), and the Portland Police Bureau (Portland Police Bureau, 2022). Prior to aggregating and analyzing these data, the policy documents for the use of force in each agency were reviewed to consider (1) whether pointing a firearm was considered a use of force, and (2) whether reporting requirements were similar between agencies (see Austin PD (2020); Baltimore PD (2019a; 2019b); Dallas PD (2021); Portland PB (2020)). Use of force policy in each of these agencies was markedly similar, while there were only marginal differences in reporting requirements, primarily relating to officers that observed the use of force. Drawing and pointing firearms was considered reasonable in each agency when an objective and reasonable threat that may escalate to require deadly force was identified. However, each agency considered the drawing and pointing of a firearm alone as a serious use of force. Further, in each agency, warning shots were explicitly prohibited, meaning there was a clear distinction between the reporting on drawing and pointing a firearm at a citizen as a use of force, and discharging a firearm.

All four agencies required the officer that drew and pointed their firearm to formally report their use of force. However, policy in Portland PB and Baltimore PD also required all officers that observed a use of force to submit an independent report. Austin PD required observing officers to make a supplemental report to the primary reporting officer that drew and pointed their firearm, while Dallas PD broadly stated that any use of force above a certain threshold, that includes drawing and pointing a firearm, must be reported. Notably, the Baltimore and Dallas PDs suggested that where firearms were drawn while entering a building during the service of a warrant, without being pointed in response to a specific threat, these need not be reported as a use of force.

Given that these jurisdictions featured markedly similar use of force and reporting policies, data from each jurisdiction were aggregated for an overlapping three-year period, from 1st January 2017 to 31 December 2019. Each agency featured substantial data relating to UOF incidents during this three-year period, with Austin reporting 11,703, Baltimore reporting 14,464, Dallas reporting 8,367, and Portland reporting 4,079 UOF incidents. Importantly, there were variables reflecting the same content reported among each jurisdiction. These included the motivation for the UOF as reported by officers, the service type that officers were engaged in at the time of using force, the length of service of the officer (tenure), whether the incident involved an arrest, whether an officer was injured during the incident, the race of the citizen, and the gender of the citizen.

However, among these data, there were varying degrees of missingness. Relating to variables considered here, data from Austin featured 991 records with missing data (8.5%), while Baltimore featured 1,855 (12.8%), Dallas featured 643 (7.7%), and Portland only featured 36 (0.9%). Riddell & Worrall (2021), in considering UOF incidents in New Orleans, identified that records were commonly duplicated at UOF incidents in which officers drew their firearms. To account for this, we flagged any records that featured the same date, time, and officer identification number as duplicates. This resulted in the identification of 405 duplicate records in Austin, 4,990 in Baltimore, 459 in Dallas, and 6 in Portland that were removed from the analysis. After removing missing and duplicate records we retained 10,307 records from Austin PD, 7,589 from Baltimore PD, 7,265 from Dallas PD, and 4,033 records from Portland Police Bureau that described UOF incidents between the January of 2017, and December of 2019.

This research sought to consider whether, using these variables, it was possible to predict whether patrol officers would draw and point their firearm at an interaction that featured the UOF. Patrol officers were selected based on their duties at the time of the use of force. In particular, we sought to understand which variables, and which features within those variables were most contributory to the likelihood of officers drawing and pointing their firearms as a means of control.

We restricted our analysis to patrol officers due to the variety of interactions they have, in opposition to specialist officers for whom drawing firearms may be a function of their specialist duties[3]. For example, an officer attached to a tactical response unit may draw their firearm at every call out, where a fraud investigator may draw their firearm at very few if any incidents, meaning the duty type variable likely explains the vast majority of variance in incidents at which these officers draw their firearm. Subsequently, this research sets aside specialist officers to focus on the decisions of patrol officers.

Methodology

Analytical Procedure

It was first important to compare the available data between each jurisdiction. Summary statistics were provided for each common variable between agencies and a simple correlation analysis was then undertaken. Correlation is important prior to modelling, primarily to rule out collinearity. Although bootstrap aggregation of the random forest typically means the influence of collinear variables is limited, individual feature importance may be artificially inflated (Cubitt et al., 2020). These variables featured a heterogeneous structure, and as such required a heterogeneous correlation matrix (McGrath & Meyer, 2006). Because the data featured a mix of binary, categorical and continuous variables, the correlation matrix produced here employed a mix of tetrachoric (Brown & Benedetti, 1977), polychoric (Brown, 2006; Babchishin & Helmus, 2016), and Pearson correlations (Holgado-Tello et al., 2010; Babchishin & Helmus, 2016).

Random Forest and Logistic Regression

The volume of data considered was substantial, and the behaviors described by these data were complex. Accordingly, a machine learning analysis was employed. Machine learning

[3] Random forest modelling including specialist officers was undertaken separately. Model accuracy was particularly high (AUROC = 0.9355), with the increased model accuracy largely attributable to tactical response units drawing and pointing firearms in response to citizens that refused to comply.

analytics are particularly useful in discerning non-linear interactions between variables (Berk, 2013), a common feature among data emerging from complex behavioral decisions. As a result, the random forest algorithm was implemented. A key aspect of the random forest is its transparency, allowing for insight into the most contributory features of the modelling, including where effect size fluctuated within the range of covariates. The random forest has been consistently found to outperform generalized linear modelling, such as logistic regression (Couronné et al, 2018), however it is still considered to be an emerging methodology, and as such it is important to benchmark against logistic regression to ensure the more accurate model is employed. Logistic regression is a common feature of classification exercises using policing data (Kane & White, 2009; Gaub, 2020), and has also been established as a strong comparator for considering the efficacy of the random forest (Cubitt et al., 2020).

To compute the random forest, data were randomized and partitioned into a 70% training set and a 30% test set. The random forest algorithm was trained on the larger set, with the model tested using the partitioned test set (Hyndman & Anthanasopoulos, 2014). Modelling was performed through application of preprocess design matrices, with analysis undertaken using the statistical analysis software, R version 4.2.0, and the ‘randomForest’, ‘dplyr’, ‘pROC’, ‘pdp’, ‘PerformanceAnalytics’, and ‘ggplot2’ packages. The initial model was trained on the binary variable representing the decision by officers to draw and point, but not discharge their firearm at a UOF incident. The model was then tested against the partitioned test set, to identify the effectiveness of the classification model developed. A logistic regression was then estimated employing the same data structure, undertaking the same classification task.

A Receiver Operating Characteristic (ROC) curve was employed to identify the accuracy of each random forest and logistic regression, through the Area Under the Receiver Operating

Characteristic (AUROC) curve. The ROC curve identifies the true positive rate of classification (y-axis), compared with the false positive rate (x-axis) at any threshold value. The AUROC, which we refer to in simple terms as the accuracy of the model, represents the probability that a randomly selected case will be accurately classified.

To find the most robust modelling approach, we tune the hyperparameters of the random forest model to optimize the number of iterations, and variables randomly considered at each split. We then report the out-of-bag error estimate, which describes the aggregate error of the random forest on the training set (Schonlau & Zou, 2020). The random forest performs a type of cross validation, using out-of-bag samples, as a component of the training step of the modelling process. We therefore report the out-of-bag error estimate, alongside the confusion matrix for prediction error on the test set, and the AUROC, to describe accuracy of the modelling processes (Svetnik et al, 2003; Couronné et al., 2018; Schonlau & Zou, 2020).

To support decisions regarding modelling, we then consider whether the difference between the area under the ROC curve for each model was statistically significant. To do this, we implement the bootstrap test for statistical significance between Receiver Operating Characteristic curves.

The results of each random forest model were interpreted through Mean Decrease Gini (MDG) (Hong et al., 2016). The Gini coefficient is a measure of statistical dispersion, in which results attributed to variables are interpreted as a proportion of the overall random forest model, relative to the AUROC produced by ROC curve. In simple terms, the AUROC identifies the accuracy of the model, while the variables are attributed an MDG coefficient identifying their importance in the accuracy of those predictions. To supplement these analyses, a confusion matrix was produced for the test set of each model. The confusion matrix measures the performance of the trained model on the test set, providing a measure of how often the model successfully or unsuccessfully made predictions (Barnes & Hyatt, 2012).

Partial Dependence

A post-hoc analysis was undertaken of covariates that the random forest indicated as having a noteworthy interaction with the likelihood of patrol officers drawing and pointing their firearms. Partial Dependence Plots (PDP) were employed to consider the nature of the relationship between covariates in these models and the outcome variable. Functionally, PDPs provide the logit contribution of a given variable to the probability of classification to the dependent variable, at different points within the range of that variable, relative to the Gini coefficient emerging from the random forest. Put simply, PDPs show the association of individual variables with the outcome variable, at different points within the range of those variables.

PDPs have been used in the analysis of policing data, relating to domestic violence (Grogger et al., 2021), police misconduct (Cubitt., 2020; Cubitt & Birch, 2021) and high harm offending by outlaw motorcycle gang members (Cubitt & Morgan, 2022). In each of these analyses, PDPs were employed to consider the direct relationship between single features, and the outcome variable. However, PDPs may also be used to consider the interaction effect of any given number of variables, to assess their joint impact on the outcome variable. An interaction effect refers to the simultaneous effect of two or more independent variables on a dependent variable, in which their joint effect is different, greater or lesser, than the sum of their parts (Lavrakas, 2008). Applying this approach is important, and novel to this area of study, as it helps provide greater insight, in this circumstance, into the correlates of drawing firearms under different conditions.

Applying an interaction effect approach to joint variables is a common approach in biomedical research (Kang et al., 2021). However, this approach has not previously been applied as a post-hoc analysis to elucidate interactions resulting from a machine learning analysis in policing, or policing literature. Here, we apply this approach to measure the joint effect of several key variables on the likelihood of a patrol officer drawing and pointing their firearm.

Results

Summary Statistics and Correlation

There were 29,240 UOF incidents in these jurisdictions for consideration between the 1st of January 2017 and the 31st of December 2019. There was some variance in the rate of drawing firearms at UOF incidents between agencies, with Dallas featuring the highest rate, and Austin the lowest. Table 1 suggests that officers most commonly used force when responding to incidents that they observed, or to which they were called out. However, it was notable that officers in Baltimore more commonly used force when serving a warrant than other jurisdictions, while officers in Dallas featured an elevated rate of responding to incidents, and using force, while off duty. The reported motivation for the UOF varied between jurisdictions. In Austin and Dallas, the predominant motivation was to make an arrest, while in Baltimore and Portland, it was where a citizen did not comply with directions issued by an officer. Importantly, with the exception of Austin, a notable proportion of officers reported self-defense, or the defense of others as a motivation for UOF.

[TABLE 1 HERE]

Not all instances of UOF involved attempting an arrest, and not all instances in which officers were motivated by making an arrest, resulted in an arrest. Dallas reported a higher rate of arrests at UOF incidents than other jurisdictions, but overall, the majority of UOF incidents featured an arrest. Force was most commonly used against male citizens. In two jurisdictions, officers most commonly used force against Black citizens, while in two, officers most commonly used force against White citizens. On average, Austin featured the least experienced officers at UOF events, while injuries occurred to officers at a roughly equitable rate, with Baltimore reporting an elevated rate of injuries.

In considering correlation between variables, Table 2 suggests that the available variables featured little relationship. The most notable correlation coefficient was between the motivation for UOF reported by officers, and the UOF at an arrest ($r=-0.390$, $p=<0.01$). The only other coefficients worthy of note described the correlation between the motivation for UOF and the policing jurisdiction ($r=0.240$, $p<0.01$), and officer tenure and the policing jurisdiction ($r=0.210$, $p<0.01$). These findings suggest that the likelihood of collinearity influencing subsequent modelling was low.

[TABLE 2 HERE]

Random Forest and Logistic Regression

The random forest was computed alongside a logistic regression for comparison. These models attempted to predict whether patrol officers would draw, but not discharge, their firearms at a UOF incident. As suggested by Figure 1, both the random forest and the logistic regression featured success in predicting this outcome, with the random forest (AUROC = 0.7662) marginally outperforming the logistic regression (AUROC = 0.7451). An AUROC of greater than 0.7 is considered to represent a noteworthy prediction rate (Grogger et al, 2021), in this circumstance each model met this criterion.

[FIGURE 1 HERE]

The difference between the area under the ROC curve for these models was statistically significant ($p<0.01$). However, while the difference between modelling approaches was statistically significant, to fully explore the findings we report the outcomes of both modeling procedures.

A confusion matrix was produced to consider the distribution of correct, and incorrect decisions made by the random forest (see Table 3). Aggregate misclassification of UOF incidents in the test-set was 10.47%, which closely adhered to the out-of-bag estimate of error on the training sample (11.07%). The model featured substantial success at predicting incidents in which officers would not draw their firearms, while featuring less success predicting instances in which they would.

[TABLE 3 HERE]

The random forest suggested that the relative experience level of officers was strongly associated with likelihood of drawing and pointing firearms as a means of control. Table 4 details the relative importance of each variable in this model, with officer tenure demonstrating the strongest association, notably outperforming other variables. The jurisdiction in which the incident occurred was also important, suggesting differences in likelihood of officers drawing and pointing firearms between departments. While the motivation for the UOF reported by officers, citizen race and the service type engaged in by officers held some association with likelihood of drawing firearms, the remaining variables did not feature notable association.

[TABLE 4 HERE]

The findings of the logistic regression, computed using the same data and structure as the random forest, largely supported the findings of the random forest (see Table 5). As was reflected in the summary statistics, officers in Baltimore, Dallas, and Portland appeared more likely to draw firearms than those in Austin. Firearms were likely to be drawn when serving warrants, or in response to incidents to which officers were dispatched, rather than incidents that they observed. Officers were significantly more likely to draw firearms on male citizens than female, and on Hispanic and Asian citizens than White citizens. However, officers were no more or less likely to draw firearms on Black citizens than White citizens during use of force incidents.

[TABLE 5 HERE]

Partial Dependence and Interaction Effects

Partial dependence plots were employed to consider the interaction, at different points within variables, with patrol officers drawing their firearm at UOF incidents. The variable featuring the strongest interaction with this phenomenon was the tenure of officers, a variable that represented the experience level of officers. Figure 2 suggests that drawing firearms was associated with officers around six years into their career; however, it was not exclusively associated with junior officers. Officers with around 20 years' experience were also associated with drawing firearms.

[FIGURE 2 HERE]

Partial dependence plots were then produced to consider an interaction effect between key variables, with the interaction effect referring to the simultaneous effect of two or more independent variables on a dependent variable. These were particularly employed to consider the interaction effect of (1) the motivation for the UOF and the jurisdiction, (2) the service type undertaken at the time force was used and the jurisdiction, and (3) the race of a citizen subject to UOF and the jurisdiction. The intention was to consider whether the motivation for UOF, or the service type that officers were engaged in, influenced the decision to draw and point firearms as a means of control, and whether this differed between policing jurisdictions. Figure 3 considered the interaction effect of motivation for UOF and jurisdiction on drawing and pointing firearms. In all jurisdictions, citizens refusing to comply with officer directions was the motivation most associated with officers drawing firearms. Figure 4 considered the interaction effect of the service type officers were engaged in prior to the UOF, and the jurisdiction of officers, on the likelihood of officers drawing firearms. In the majority of jurisdictions being dispatched to an incident was most associated with drawing firearms, with the exception of Dallas, in which observing an incident was most associated with drawing firearms.

[FIGURE 3 HERE]

[FIGURE 4 HERE]

A partial dependence plot was computed to consider the interaction effect between the motivation for UOF, the service type engaged in, and the likelihood of drawing firearms. Figure 5 suggested drawing firearms was most associated with officers being dispatched to an incident, except when the intention was to make an arrest, where the incident was most likely observed by officers.

[FIGURE 5 HERE]

Finally, a partial dependence plot was computed for the interaction effect between the race of citizens, the policing jurisdiction, and the likelihood of officers drawing their firearms. Figure 6 suggests that in two jurisdictions, Baltimore and Dallas, Black citizens were most likely to have firearms drawn on them by officers. However, in Austin and Portland, White citizens were most likely to have firearms drawn on them by officers.

[FIGURE 6 HERE]

Discussion

Pointing a firearm at a person is a significant use of police authority – and unfortunately, one that is less reliably tracked across agencies and far less empirically scrutinized than when officers discharge their firearms (see, e.g., Garner et al., 2018, Table 4; VerBruggen, 2022). Prior work concerned with this behavior has primarily considered the association between suspect and officer characteristics, alongside situational variables, and whether they influenced the decision to draw but not discharge a firearm at UOF events (Ridell & Worrall, 2021; Worrall et al., 2021). Our findings suggest that there may be a complex interplay of situational and characteristic features that relate to the likelihood of officers drawing firearms, and pivotally, there may be jurisdictional differences.

In New Orleans, Riddell and Worrall (2021) found that citizen race was not a significant predictor of firearms displays in incident-level analyses – instead, situational characteristics were more important. Meanwhile, in Dallas, Worrall and colleagues (2021) found that Black suspects were significantly less likely than White suspects to have guns drawn on them during UOF incidents, but the difference was not statistically significant when restricted to the subset of UOF incidents involving an arrest. In the present study, we observed that situational variables were more important in model accuracy than race, but we also find that the extent to which race was associated with drawing firearms varied across jurisdictions. For example, Figure 6 showed that in some jurisdictions, drawing firearms was most associated with Black citizens, while in others it was more associated with White citizens. These findings demonstrate the nuance in analyzing UOF incidents between jurisdictions, and the importance of benchmarking findings with other research and other agencies.

In comparing findings, the rate at which officers drew their firearms in different jurisdictions differed considerably. Riddell and Worrall (2021) reported that almost half of UOF incidents in New Orleans between 2016 and 2019 resulted in a firearm being drawn. In contrast, Stansfield and colleagues (2021) reported that in New Jersey between 2012 and 2016, only 1,425 out of more than 70,000, or 0.02% of UOF incidents resulted in officers drawing or discharging their firearms. This discrepancy in rates of drawing firearms between jurisdictions should raise concerns about generalizing from single-site UOF studies.

Priming the decision to draw firearms

The findings from interaction effects suggested that a priming effect may be present in the decision to draw firearms at a UOF event. Priming is a phenomenon in which exposure to an initial stimulus unconsciously influences a response to a subsequent stimulus (Bargh et al., 1996; Shanks et al., 2013). Here, Figure 4 suggested that firearms displays were most likely in response to a call for service from dispatch. Logistic regression findings supported this effect, also noting an association between drawing firearms and the service of warrants. Both of these are duties in which an officer is exposed to stimulus, either from dispatch or content relating to the warrant, before a gap in time prior to the UOF incident. We contrast this with instances where, for example, and officer observed an offense and responded immediately. At UOF incidents such as these, drawing firearms was less likely.

Taylor (2019) employed an experimental methodology to test the likelihood that officers would engage in deadly UOF, depending on the type of information provided to them by dispatch when called out to an incident. This research intended to consider whether officers were primed by the information provided to them prior to the incident, with a significant increase in fatal shootings associated with incidents in which erroneous information was provided by dispatch (Taylor, 2019). The present findings suggest that officers who were dispatched to incidents, and those that were serving a warrant, were more likely to draw their firearms at UOF incidents. We suggest that a similar priming effect, to that identified by Taylor (2019), may be operating here. Officers who experienced a delay between the call out and arriving at an incident, either from dispatch or from duties relating to serving a warrant, held a greater association with drawing firearms, when compared with officers that observed and responded to an incident immediately. While we suggest that the effect observed in the present study may be, at least in part, attributable to priming, this is an under-studied phenomenon in policing. The influence of priming on perceived risk in use of force scenarios requires greater consideration, and may be an important area for future research.

The generalizability of single site studies

Findings from both summary statistics (Table 1) and the logistic regression (Table 5) demonstrated the notable difference in the rate at which officers drew their firearms. Officers in Austin drew their firearms least, with officers in Baltimore, Dallas, and Portland drawing firearms at least three times as frequently, a finding that was also reflected in the estimated odds of an officer drawing their firearm in each jurisdiction. Further, the differences in types of incidents that resulted in police using force, between jurisdictions, appeared noteworthy. This was reflected in the service type engaged in at time of UOF and the reported motivations for UOF.

These findings suggest that there may be some concerns relating to the generalizability of single site studies on firearms displays. The considerable disparity between the rate of displaying firearms between jurisdictions in our study reinforces the notion that caution should be exercised in suggesting generalized policy change from analyses that consider only one jurisdiction.

Although we cannot comment on the discharge of firearms using these data, the data used in this research, particularly when compared with other research on the same subject matter, suggest that single site studies may not provide sufficient context to generalize findings.

Generalizability is not only important to consider in prediction studies. Among studies employing network analytics, there is emerging evidence of the social transmission of firearms use among peer groups of officers. Pivotally, this research does not always point to the same effect between jurisdictions. For example, Ouellet and colleagues (2022), analyzing data from New Jersey, suggested that officers with greater exposure to colleagues that have a history of discharging their firearm were less likely to discharge their own. However, in a similar study using data from Chicago, Zhao and Papachristos (2020) found that officers located in brokerage positions in a network, a position in between other officers, or groupings of officers, were more likely to discharge their firearm. While the body of research employing network analytics to firearms use by police is emerging, comparing findings between jurisdictions also appears important.

Implications

This research features implications in both applied and research domains. Initially, it appears that the officers associated with drawing firearms were not necessarily those expected. For example, the tenure of officers was an important predictor, but it was not only junior officers that were associated with drawing firearms – this effect peaked again among later-career officers. There were also implications for the duties undertaken by officers. In particular, duties relating to the service of a warrant, or when dispatched to an incident, featured an elevated likelihood of drawing firearms. The implications of these findings relate to the way in which we view settings in which officers exercise force of this type. It is more likely that officers would draw firearms at events in which there was a temporal gap between being notified of the duties, and the incident. While we have suggested that this may be attributable to a priming effect, given that officers observing an incident did not demonstrate a similar response, we cannot rule out observed incidents featuring less inherent risk for police – certainly it is possible that warrant service duties feature greater risk. Although there is some way to go before identifying causation in this effect, if it is indeed attributable to a priming effect, there are important implications for the ways that information is distributed to officers, and the methods of engagement in response to dispatch call-outs and warrant service duties (Simpson, 2021; Simpson & Orosco, 2021).

While we have already discussed the difficulty generalizing single-site studies, the implications of this finding bears note for research, and the extent to which findings may be used to inform policy or practice. In this study, we analyze four policing agencies, and within those four agencies there is some meaningful variation in the situational, motivational, and characteristic traits of UOF incidents. Whether this is a function of the jurisdictional context, or the training or culture of the individual agencies (Ingram et al., 2018), we cannot be certain. However, it is clear that the decision of an officer to draw their firearm is a complex one, subject to considerable influence from situational factors alongside individual characteristics. The variation in rates at which officers draw their firearms across jurisdictions means researchers should be hesitant in generalizing their conclusions or suggesting that their findings be drawn into policy or practice beyond the jurisdiction that they are analyzing. Research focusing on officers drawing, but not discharging their firearms is relatively limited. This knowledge area, alongside UOF more broadly, may benefit from multi-site studies to more closely consider the generalizability of findings.

Limitations

Data entry is an important feature of the policing environment, typically performed as a non-priority by time-poor staff. Given the primary intention of this data was not research, but rather an administrative process by each of these agencies, this is considered to be naturally occurring data (Lester et al, 2017). While we consulted the policy documents for when firearms should be drawn and pointed at civilians, and the reporting requirements for this use of force type to ensure similarity, there were also some minor differences. It is possible that the marginal differences in reporting requirements for officers that observed use of force incidents may have impacted upon reported rates of use of force in these data. For example, while the reported rate of use of force at warrant duties was low, it differed between agencies. The similarities in the policy statements of these agencies does not ultimately guarantee similar data quality, or that the same incidents will be captured between agencies; there may still be differences in officers reporting, or quality of administrative data entry. Naturally occurring data commonly features flaws, such as missing data and inaccuracy of input. While we attempt to mitigate this limitation through removing records that were missing data, it remains possible that the quality of the data varied between these four agencies.

These data represent a large volume of records for analysis, however there are almost certainly unreported instances of UOF. Data here does not include instances that involved minor acts of force. Analytically, this is a within groups analysis, meaning findings should be applied to instances of UOF only, and should not be generalized to all police interactions with citizens. Indeed, it is important to note we cannot rule out the possibility of collider bias, whereby sampling on UOF incidents blocked out any biases that might have influenced officers' decisions to stop and/or use force on citizens (Knox et al., 2020; Neil & Winship, 2019). In other words, although we observed that Black citizens were no more or less likely than White citizens to have guns drawn on them during UOF incidents, we cannot say with certainty that Black people are equally likely *overall* to have an officer draw a gun on them, especially given that studies routinely show police are more likely to stop racial minorities in the first place (Epp et al., 2014; Gelman et al., 2007; Pierson et al., 2020). Thus, one opportunity for future research in this area is to employ a matched sample of citizen interactions with police that did not result in a UOF. Unfortunately, the present research could not consider the prevalence of UOF among all police interactions.

As previously mentioned, collinearity is a risk among administrative and procedural datasets. In particular, collinear variables are an important risk in random forest modelling. Highly correlated variables increase the possibility that they may artificially inflate the rate of classification accuracy of a given model. However, bootstrap aggregation, or bagging, in the random forest mitigates the likelihood that accuracy of models presented here are impacted upon by collinear variables, although individual variable importance may be marginally impacted (Cubitt et al., 2020). To identify the likelihood of this influence, we computed and presented a correlation matrix as Table 2. This matrix did not suggest that there was sufficient correlation between variables to influence variable importance, in particular it did not reveal variable correlation that may overcome the bagging technique. While it was important to note this as a limitation, there was no meaningful evidence of collinearity among these data. Further, the confusion matrix for the random forest closely adhered to the out-of-bag error estimate on the training set, and the area under the ROC curve suggesting that collinearity was not a noteworthy aspect of this analysis.

Finally, it is noteworthy that certain variables performed differently in the logistic regression analysis than they did in the random forest. For example, in the logistic regression the officer tenure variable suggested that longer tenure was associated with decreased odds of an officer drawing their firearm. However, while the partial dependence plot in Figure 2 suggested a similar aggregate effect as the logistic regression, there were certain points at which the association with drawing and pointing firearms increased. In this instance the findings are explainable, as the logistic regression describes the aggregate effect, while the partial dependence plot describes the effect at each point within officer tenure. However, it is important to note that there is nothing preventing a variable with a small linear effect size, as estimated in a logistic regression, from featuring high importance in a random forest model. The random forest relies on the way covariates interact to discriminate between the binary outcome variable, it then describes the differential importance of the covariates in undertaking this task. As a result, while this is not entirely the case here, it is possible that the logistic regression, and the random forest take different paths to predict the outcome variable, and may therefore produce divergent results.

Conclusion

We close by emphasizing that this sort of research is only possible when agencies are transparent with their data. We also implore the thousands of agencies that do not currently require officers to document when they point their firearms at citizens (Brooks, 2020) to begin doing so immediately. Not only does such a reporting requirement appear to be associated with a reduction in police shootings (Jennings & Rubado, 2017) without jeopardizing officer safety (Shjarback et al., 2021), it also enables agencies and researchers to better review and understand "near misses" – those instances in which officers averted the use of deadly force. But even in the absence of these two justifications, there is a third reason to do so: it is the morally appropriate thing for police in a democratic society to do. We hope that our work will spark additional research on this specific use of police coercion.

References

Austin Police Department. (2020). General Orders. https://www.austintexas.gov/sites/default/files/files/Police/General_Orders.pdf

Babchishin, K.M. & Helmus, L.M. (2016). The influence of base rates on correlations: An evaluation of proposed alternative effect sizes with real-world data. *Behavior Research Methods*, 48, 1021-1031.

Barnes, G.C. & Hyatt, J.M. (2012). *Classifying adult probationers by forecasting future offending.* Washington DC: US Department of Justice.

Bayley, D. H. & Bittner, E. (1984). Learning the skills of policing. *Law and Contemporary Problems*, 47(4), 35-59.

Bayley, D. H., & Garofalo, J. (1989). The management of violence by police patrol officers. *Criminology*, 27, 1-25.

Baltimore Police Department. (2019a). Use of force. https://public.powerdms.com/BALTIMOREMD/documents/51042

Baltimore Police Department. (2019b). Use of force reporting, review, and assessment. https://public.powerdms.com/BALTIMOREMD/documents/51042

Baumgartner, F. R., Epp, D. A., & Shoub, K. (2018). *Suspect citizens: What 20 million traffic stops tell us about policing and race*. Cambridge University Press.

Bennett Moses, L, & Chan J. (2018). Algorithmic prediction in policing: Assumptions, evaluation, and accountability. *Policing and Society*, 7, 806–822. https://doi.org/10.1080/10439463.2016.1253695

Berk, R., Sherman, L., Barnes, G., Kurtz, E., and Ahlman, L. (2009). Forecasting murder within a population of probationers and parolees: A high stakes application of statistical learning. *Statistics in Society*, 172(1), 191–211. https://doi.org/10.1111/j.1467-985x.2008.00556.x

Berk, R. (2013). Algorithmic criminology. *Security Informatics*, 2(5), 1-14.

Berk, R., Sorenson, S.B. & Barnes, G. (2016). Forecasting domestic violence: A machine learning approach to help inform arraignment decisions. *Journal of Empirical Legal Studies*, 13(1), 94-115.

Berk, R., Heidari, H., Jabbari, S., Kearns, M. & Roth, A. (2018). Fairness in criminal justice risk assessment: The state of the art. *Sociological Methods & Research*, 50(1), 3–44. https://doi.org/10.1177/0049124118782533

Berk, R.A. (2019). *Machine Learning Risk Assessments in Criminal Justice Settings*. Springer: New York.

Bittner, E. (1970). The functions of the police in modern society: A review of background factors, current practices, and possible role models.

Bishopp, S. A., Klinger, D. A., & Morris, R. G. (2015). An examination of the effect of a policy change on police use of TASERs. *Criminal Justice Policy Review*, *26*(7), 727-746.

Bolger, P. C. (2015). Just following orders: A meta-analysis of the correlates of American police officer use of force decisions. *American journal of criminal justice*, *40*(3), 466-492.

Brooks, C. (2020). L*ocal police departments: Policies and procedures, 2016* (NCJ 254826). US Department of Justice: Bureau of Justice Statistics.

Brown, M. B., & Benedetti, J. K. (1977). On the mean and variance of the tetrachoric correlation coefficient. *Psychometrika*, *42*, 347–355. doi:10.1007/BF02293655

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford.

Chan, J. (1996). Changing police culture. *British Journal of Criminology*, *36*(1), 109-134.

Chanin, J., & Courts, J. (2017). Examining the determinants of police department online transparency. *Criminology, Criminal Justice, Law & Society*, *18*, 52.

City of Austin Open Data Portal. (2022). 2019 Response to Resistance Data. https://data.austintexas.gov/dataset/2019-Response-to-Resistance-Data/3bfz-mri4

Couronne, R., Probst, P. & Boulesteix, A.L. (2018). Random forest versus logistic regression: A large-scale benchmark experiment. *BMC Bioinformatics*, 19(270).

Cubitt, T.I.C., Wooden, K. & Roberts, K. (2021). A machine learning analysis of serious misconduct among police. *Crime Science*, 9(22).

Cubitt, T.I.C. & Birch, P. (2021). A machine learning analysis of misconduct in the New York Police Department. *Policing: An International Journal,* 44(5), 800-817.

Cubitt, T.I.C. & Morgan, A. (2022). Predicting high harm offending using machine learning: An application to outlaw motorcycle gangs. *Trends and Issues in Crime and Criminal Justice* no. 646. Canberra: Australian Institute of Criminology.

Dallas Open Data. (2022). Police Response to Resistance. https://www.dallasopendata.com/Public-Safety/Police-Response-to-Resistance-2019/46zb-7qgj

Dallas Police Department. (2021). General Orders. https://dallaspolice.net/resources/Shared%20Documents/General-Orders.pdf

Epp, C. R., Maynard-Moody, S., & Haider-Markel, D. P. (2014). *Pulled over: How police stops define race and citizenship*. University of Chicago Press.

Fryer Jr, R. G. (2019). An empirical analysis of racial differences in police use of force. *Journal of Political Economy*, *127*(3), 1210-1261.

Fyfe, J. J. (1988). Police use of deadly force: Research and reform. *Justice Quarterly*, *5*(2), 165-205.

Garner, J. H., Hickman, M. J., Malega, R. W., & Maxwell, C. D. (2018). Progress toward national estimates of police use of force. *PloS one, 13*(2), e0192932.

Gelman, A., Fagan, J., & Kiss, A. (2007). An analysis of the New York City police department's "stop-and-frisk" policy in the context of claims of racial bias. *Journal of the American Statistical Association*, *102*(479), 813-823.

Grogger, J., Gupta, S., Ivandic, R. & Kirchmaier, T. (2021). Comparing conventional and machine-learning approaches to risk assessment in domestic abuse cases. *Journal of Empirical Legal Studies*, 18(1), 90–130. https://doi.org/10.1111/jels.12276

Holgado-Tello, F. P., Chacón-Moscoso, S., Barbero-García, I., & Vila-Abad, E. (2010). Polychoric versus Pearsons correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality and Quantity*, *44*, 153–166. doi:10.1007/s11135-008-9190-y

Hong, H., Xioling, G. & Hua, Y. (2016, 26-18 August 2016). Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest. Paper presented at the 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS).

Hyndman, R. J., and Anthanasopoulos, G. (2014). *Forecasting: Principles and practice*. Melbourne, Victoria: Monash University.

Ingram, J. R., Terrill, W., & Paoline III, E. A. (2018). Police culture and officer behavior: Application of a multilevel framework. *Criminology*, *56*(4), 780-811.

James, L., Vila, B., & Daratha, K. (2013). Results from experimental trials testing participant responses to White, Hispanic and Black suspects in high-fidelity deadly force judgment and decision-making simulations. *Journal of Experimental Criminology*, *9*(2), 189-212.

James, L., Klinger, D., & Vila, B. (2014). Racial and ethnic bias in decisions to shoot seen through a stronger lens: Experimental results from high-fidelity laboratory simulations. *Journal of Experimental Criminology*, *10*(3), 323-340.

James, L., James, S. M., & Vila, B. J. (2016). The reverse racism effect: Are cops more hesitant to shoot black than white suspects? *Criminology & Public Policy*, *15*(2), 457-479.

James, L., James, S., & Vila, B. (2018a). The impact of work shift and fatigue on police officer response in simulated interactions with citizens. *Journal of Experimental Criminology*, *14*(1), 111-120.

James, L., James, S., & Vila, B. (2018b). Testing the impact of citizen characteristics and demeanor on police officer behavior in potentially violent encounters. *Policing: An International Journal*, 41(1), 24-40.

Jennings, J. T., & Rubado, M. E. (2017). Preventing the use of deadly force: The relationship between police agency policies and rates of officer-involved gun deaths. *Public Administration Review*, *77*(2), 217-226.

Kahn, K. B., Goff, P. A., Lee, J. K., & Motamed, D. (2016). Protecting whiteness: White phenotypic racial stereotypicality reduces police use of force. *Social Psychological and Personality Science*, *7*(5), 403-411.

Kaminski, R. J., Edwards, S. M., & Johnson, J. W. (1998). The deterrent effects of oleoresin capsicum on assaults against police: Testing the velcro-effect hypothesis. *Police Quarterly*, *1*(2), 1-20.

Kane, R.J. & White, M.D. (2009). Bad cops: A study of career-ending misconduct among New York City police officers. *Criminology and Public Policy*, 8(4), 737-769.

Kang, J.W.M., Mor, D. & Keay, K.M. (2021). Nerve injury alters restraint-induced activation of the basolateral amygdala in male rats. *Brain Structure and Function*, 226(4), 1209-1227.

Knox, D., Lowe, W., & Mummolo, J. (2020). Administrative records mask racially biased policing. *American Political Science Review*, *114*(3), 619-637.

Lavrakis, P.J. (2008). Encyclopedia of Survey Research Methods. *SAGE Research Methods*. https://dx.doi.org/10.4135/9781412963947

Matusiak, M. C., Cavanaugh, M. R., & Stephenson, M. (2022). An assessment of officer-involved shooting data transparency in the United States. *Journal of interpersonal violence*, *37*(1-2), 472-496.

McGrath, R. E., & Meyer, G. J. (2006). When effect sizes disagree: The case of the *r* and *d*. *Psychological Methods*, *11*, 386–401.

McKay, C. (2019). Predicting risk in criminal justice procedure: Actuarial tools, algorithms, AI and judicial decision-making. *Current Issues in Criminal Justice*, 32(1), 22–39. https://doi.org/10.1080/10345329.2 019.1658694

Molden, D. C. (2014). Understanding priming effects in social psychology: What is “social priming” and how does it occur? In Molden, D. (Ed.), Understanding priming effects in social psychology (pp. 3–13). New York, NY: The Guilford Press.

New Orleans Open Data. (2022). NOPD Use of Force Incidents. https://data.nola.gov/Public-Safety-and-Preparedness/NOPD-Use-of-Force-Incidents/9mnw-mbde

Neil, R., & Winship, C. (2019). Methodological challenges and opportunities in testing for racial discrimination in policing. *Annual Review of Criminology*, *2*, 73-98.

Nix, J., Campbell, B. A., Byers, E. H., & Alpert, G. P. (2017). A bird's eye view of civilians killed by police in 2015: Further evidence of implicit bias. *Criminology & Public Policy*, *16*(1), 309-340.

Nix, J., Richards, T. N., Pinchevsky, G. M., & Wright, E. M. (2021). Are domestic incidents really more dangerous to police? Findings from the 2016 National Incident Based Reporting System. *Justice Quarterly*, *38*(7), 1405-1427.

Nix, J., & Shjarback, J. A. (2021). Factors associated with police shooting mortality: A focus on race and a plea for more comprehensive data. *PLoS one*, *16*(11), e0259024.

Ouellet, M., Hashimi, S. & Vega Yon, G.G. (2022). Officer network and firearm behaviours: Assessing the social transmission of weapon use. *Journal of Quantitative Criminology*, ahead of print.

Paoline III, E. A., & Terrill, W. (2007). Police education, experience, and the use of force. *Criminal justice and behavior*, *34*(2), 179-196.

Paoline III, E. A., Terrill, W., & Somers, L. J. (2021). Police officer use of force mindset and street-level behavior. *Police Quarterly*, *24*(4), 547-577.

Pate, A. M., & Fridell, L. A. (1995). Toward the uniform reporting of police use of force: Results of a national survey. *Criminal Justice Review*, *20*(2), 123-145.

Pierson, E., Simoiu, C., Overgoor, J., Corbett-Davies, S., Jenson, D., Shoemaker, A., ... & Goel, S. (2020). A large-scale analysis of racial disparities in police stops across the United States. *Nature Human Behaviour*, *4*(7), 736-745.

Portland Police Bureau. (2020). Use of Force. https://www.portlandoregon.gov/police/article/751998

Portland Police Bureau. (2022). Use of Force Report. https://www.portlandoregon.gov/police/76875Project Comport, (2019). Baltimore Police Department: Use of Force Incidents. https://www.projectcomport.org/department/BPD/useofforce/#:~:text=Use%20of%20Force%20Incidents,is%20reported%2C%20tracked%20and%20investigated.

Riddell, J. R., & Worrall, J. L. (2021). Predicting firearm and CEW displays as police officers' response to resistance. *Journal of Criminal Justice*, *72*, 101775.

Ridgeway, G. (2013). The pitfalls of prediction. *National Institute of Justice Journal*, 271, 34–40. https://nij.ojp.gov/topics/articles/pitfalls-prediction

Rosenbaum, D. P., Graziano, L. M., Stephens, C. D., & Schuck, A. M. (2011). Understanding community policing and legitimacy-seeking behavior in virtual reality: A national study of municipal police websites. *Police Quarterly*, *14*(1), 25-47.

Schonlau, M. & Zou, R.Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal: Promoting communications on statistics and Stata,* 20(1), 3-29. https://doi.org/10.1177%2F1536867X20909688

Shanks, D.R., Newell, B.R., Lee, E.H., Balakrishnan, D., Ekelund, L., Cenac, Z., Kavvadia, F. & Moore, C. (2013). Priming intelligent behavior: An elusive phenomenon. *PLOS ONE, 8(4).*

Shjarback, J. A. (2019). State-mandated transparency: A discussion and examination of deadly force data among law enforcement agencies in Texas. *Journal of Crime and Justice*, *42*(1), 3-17.

Shjarback, J. A., White, M. D., & Bishopp, S. A. (2021). Can police shootings be reduced by requiring officers to document when they point firearms at citizens? *Injury Prevention*, *27*(6), 508-513.

Sierra-Arévalo, M. (2021). American policing and the danger imperative. *Law & Society Review*, *55*(1), 70-103.

Simpson, R. (2021). Calling the police: Dispatchers as important interpreters and manufacturers of calls for service data. *Policing: A Journal of Policy and Practice*, *15*(2), 1537-1545.

Simpson, R., & Orosco, C. (2021). Re-assessing measurement error in police calls for service: Classifications of events by dispatchers and officers. *PLOS ONE*, *16*(12), e0260365.

Skolnick, J. (1966). *Justice without trial: Law enforcement in democratic society*. John Wiley.

Smith, M. R., & Alpert, G. P. (2000). Pepper spray–A safe and reasonable response to suspect verbal resistance. *Policing: An International Journal of Police Strategies & Management.*

Smith, M. R., & Alpert, G. P. (2007). Explaining police bias: A theory of social conditioning and illusory correlation. Criminal Justice and Behavior, 34(10), 1262-1283.

Stansfield, R., Aaronson, E., & Okulicz-Kozaryn, A. (2021). Police use of firearms: Exploring citizen, officer, and incident characteristics in a statewide sample. *Journal of Criminal Justice*, *75*, 101828.

Stevenson, M. & Doleac, J.L. (2021). Algorithmic risk assessment in the hands of humans. https://doi.org/10.2139/ssrn.3489440

Svetnik, V., Liaw, A., Tong, C., Culberson, C., Sheridan, R.P. & Feuston, B.P. (2003). Random Forest: A classification and regression tool for compound classification and QSAR Modeling. *Journal of Chemical Information and Computer Sciences,* 43(6), 1947-1958. https://doi.org/10.1021/ci034160g

Taylor, P.L. (2019). Dispatch priming and the police decision to use deadly force. *Police Quarterly,* 23(3), 311-332.

Terrill, W., & Paoline III, E. A. (2017). Police use of less lethal force: does administrative policy matter? *Justice Quarterly*, *34*(2), 193-216.

Trinkner, R., Kerrison, E. M., & Goff, P. A. (2019). The force of fear: Police stereotype threat, self-legitimacy, and support for excessive force. *Law and Human Behavior*, *43*(5), 421.

VerBruggen, R. (2022). Fatal Police Shootings and Race: A Review of the Evidence and Suggestions for Future Research. *Manhattan Institute Report*, 1-32.

Wheeler, A. P., Phillips, S. W., Worrall, J. L., & Bishopp, S. A. (2017). What factors influence an officer's decision to shoot? The promise and limitations of using public data. *Justice Research and Policy*, *18*(1), 48-76.

White, M. D. (2001). Controlling police decisions to use deadly force: Reexamining the importance of administrative policy. *Crime & Delinquency*, *47*(1), 131-151.

White, M. D., & Ready, J. (2007). The TASER as a less lethal force alternative: Findings on use and effectiveness in a large metropolitan police agency. *Police Quarterly*, *10*(2), 170-191.

Wilson, J. Q. (2009). *Varieties of police behavior: The management of law and order in eight communities*. Harvard University Press.

Worrall, J. L., Bishopp, S. A., Zinser, S. C., Wheeler, A. P., & Phillips, S. W. (2018). Exploring bias in police shooting decisions with real shoot/don't shoot cases. *Crime & Delinquency*, *64*(9), 1171-1192.

Worrall, J. L., Bishopp, S. A., & Terrill, W. (2021). The effect of suspect race on police officers' decisions to draw their weapons. *Justice Quarterly*, *38*(7), 1428-1447.

Zhao, L. & Papachristos, A.V. (2020). Netowrk position and police who shoot. *The ANNALS of the American Academy of Political and Social Science*, *687*(1), 89-112.

Table 1. Characteristics of use of force incidents by patrol officers between 1 January 2017 and 31 December 2019 among 4 Police Departments.

	Agency			
	Austin	Baltimore	Dallas	Portland
Complete records (*n*)	10,307	7,589	7,265	4,033
Situational characteristics				
Drew firearm (%)	3.75	14.88	16.48	11.70
Service type at time of use of force (%)				
Observed incident	25.17	50.62	60.67	20.42
Dispatched to incident	72.24	45.21	34.10	79.47
Serving a warrant	0.39	3.94	0.67	0.00
Off duty	0.00	0.23	2.63	0.11
Other	2.20	0.00	1.93	0.00
Reported motivation for use of force (%)				
Making an arrest	64.35	17.68	52.48	13.92
Citizen failed to comply	27.77	67.64	26.90	67.24
Defense of self or others	6.70	14.68	20.62	18.84
Other	1.18	0.00	0.00	0.02
UOF that resulted in an arrest (%)	64.29	68.96	85.33	72.78
Citizen characteristics				
Citizen gender				
Male	74.99	82.87	82.55	76.23
Female	25.01	17.13	17.45	23.77
Citizen race				
Asian	0.59	0.17	0.55	2.84
Black	20.85	84.87	54.89	29.03
Hispanic	32.13	1.37	21.06	8.78
Indigenous American	0.11	0.01	0.29	1.34
White	45.94	6.90	22.11	58.01
Unknown	0.38	4.71	1.10	0.00
Officer characteristics				
Average officer tenure (years)	5.24	6.36	7.32	9.43
Officer injured (%)	9.52	18.80	9.24	5.08

Table 2. Correlation matrix for data on use of force incidents by patrol officers

Draw and point firearm	.059**	-.026**	-.059**	-.091**	.018*	-.077**	.001	.120**
	Motivation	.046**	-.049**	.022**	.094**	.025**	-.390**	.240**
		Service type	.134**	.069**	.032**	-.020**	-.160**	.079**
			Citizen race	.068**	-.012**	-.065**	-.072**	-.055**
				Citizen gender	.006	-.034**	-.086**	.005**
					Officer tenure	.007	.006	.210**
						Officer injured	.053**	-.047**
							Arrest	.160**
								Police jurisdiction

*<0.05, **<0.01

Table 3. Confusion matrix for patrol officers drawing but not discharging firearms in the test set

	Predicted negative	Predicted positive	Total	Misclassification rate
True negative	7,527	649	8,176	7.95%
True positive	268	313	581	46.13%
Total	7,795	962	8,757	
Misclassification rate	3.56%	67.46%		

Table 4. Variable importance for random forest trained on officers drawing their firearms.

Variable	Mean Decrease Gini
Officer tenure	50.74
Motivation for use of force	20.56
Policing jurisdiction	9.45
Citizen race	6.08
Service type at use of force incident	4.41
Arrest during use of force incident	3.36
Officer injured	2.75
Citizen gender	2.65

Table 5. Logistic regression estimates and odds ratios

Variable	Log-odds	SE	OR
Service type at use of force incident			
Observed incident (ref.)			
Dispatched to incident	.094	.043	1.098*
Off duty	-.746	.277	.047**
Warrant service	-.608	.145	1.837**
Other	-.006	.201	.993
Motivation for use of force			
Refused to comply (ref.)			
Make an arrest	1.288	.055	3.624**
Defense of self or others	1.559	.053	4.757**
Other	1.247	.375	3.478**
Arrest during use of force incident			
No (ref.)			
Yes	-.427	.051	.652**
Citizen gender			
Female (ref.)			
Male	.903	.062	2.468**
Citizen race			
White (ref.)			
Asian	.454	0.205	1.574*
Black	.033	.054	1.033
Hispanic	.311	.065	1.364**
Indigenous American	-.096	.471	.382*
Unknown	.021	.142	1.021
Officer tenure	-.008	.003	.991*
Officer injured			
No (ref.)			
Yes	-1.089	.092	.336**
Policing jurisdiction			
Baltimore (ref.)			
Austin	-2.033	.076	0.131**
Dallas	-0.341	.056	0.711**
Portland	-0.332	.072	0.717**

NOTE: Categorical variables employ the first level as the reference variable for analysis. The findings for each subsequent category within that variable are respective to the baseline of the first variable.
*<0.05, **<0.01

Figure 1. Receiver Operating Characteristics Curve random forest (black) and logistic regression (red) for officers drawing, but not discharging their firearms at a use of force incident.

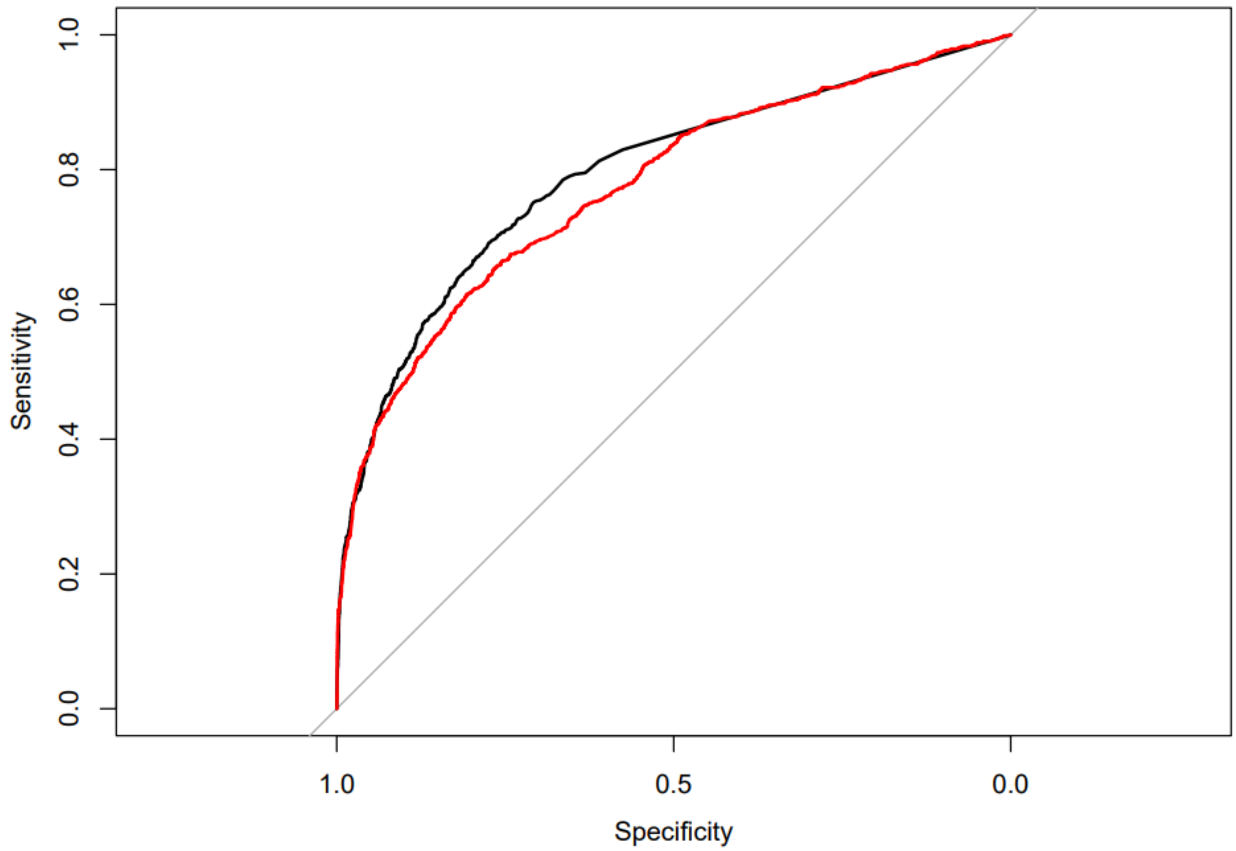

Figure 2. Partial dependence plot for officer tenure and likelihood of drawing firearms as a means of control at use of force incidents

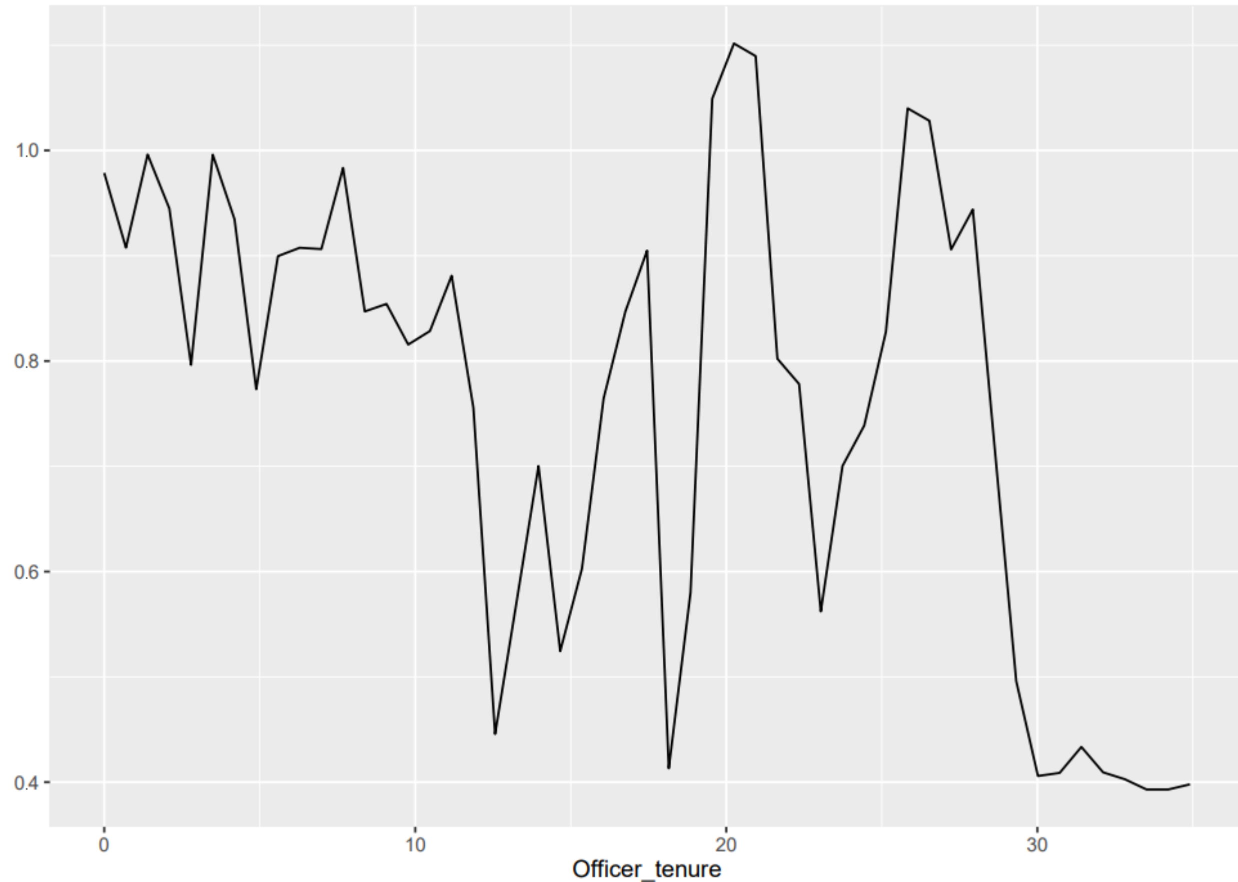

Figure 3. Partial dependence plot for the interaction effect between officer motivation, and the jurisdiction of officers, and likelihood of drawing firearms as a means of control at use of force incidents.

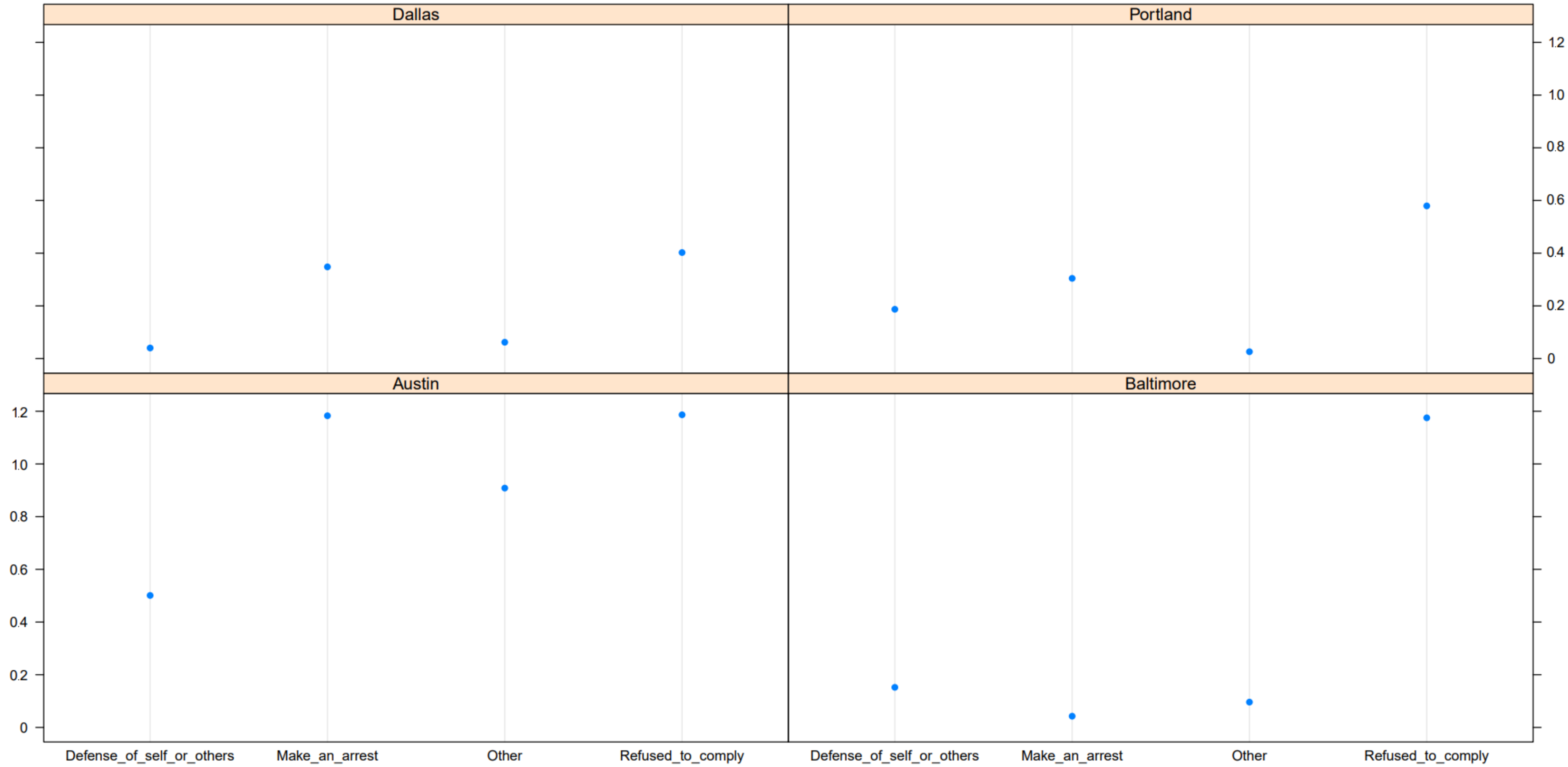

Figure 4. Partial dependence plot for the interaction effect between service type, and the jurisdiction of officers, and likelihood of drawing firearms as a means of control at use of force incidents.

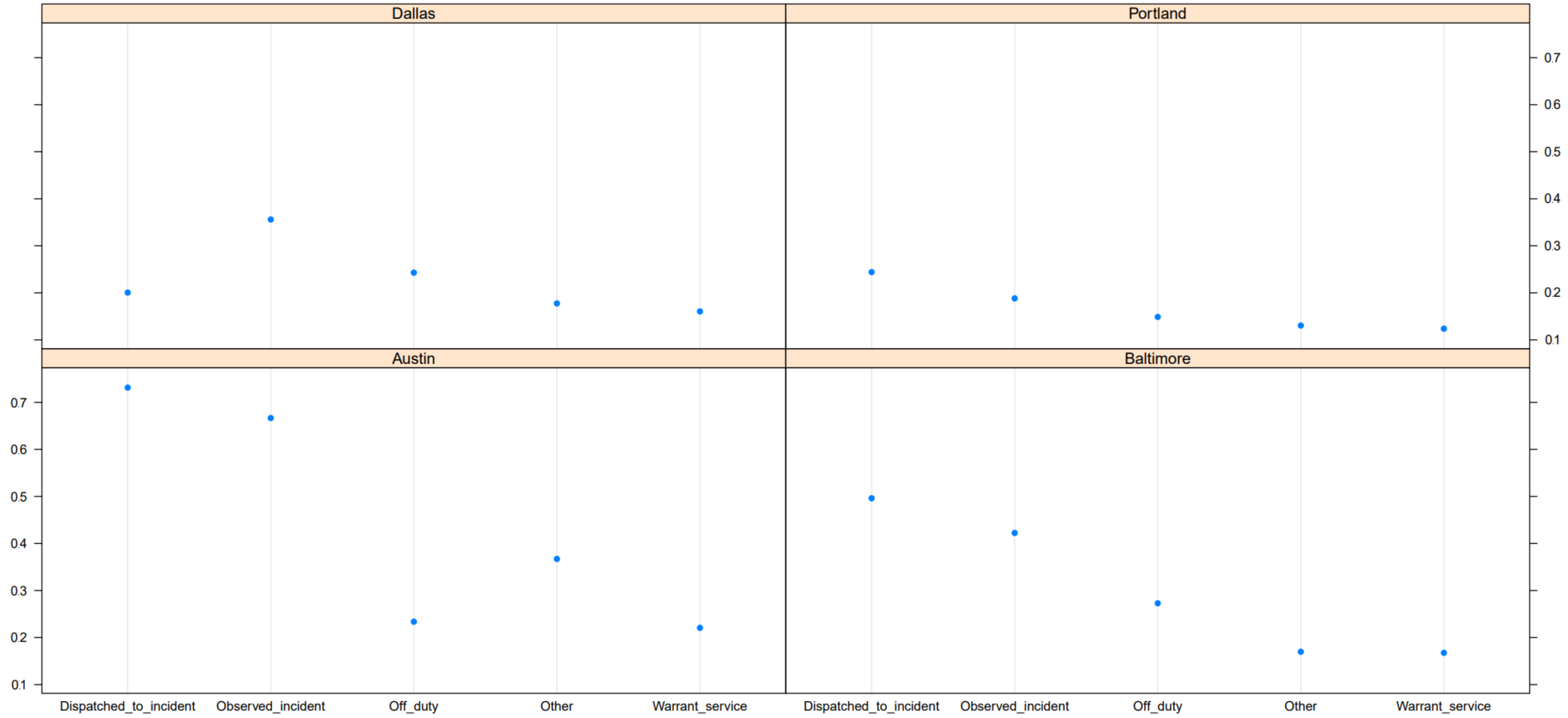

Figure 5. Partial dependence plot for the interaction effect between officer motivation, and service type, and likelihood of drawing firearms as a means of control at use of force incidents

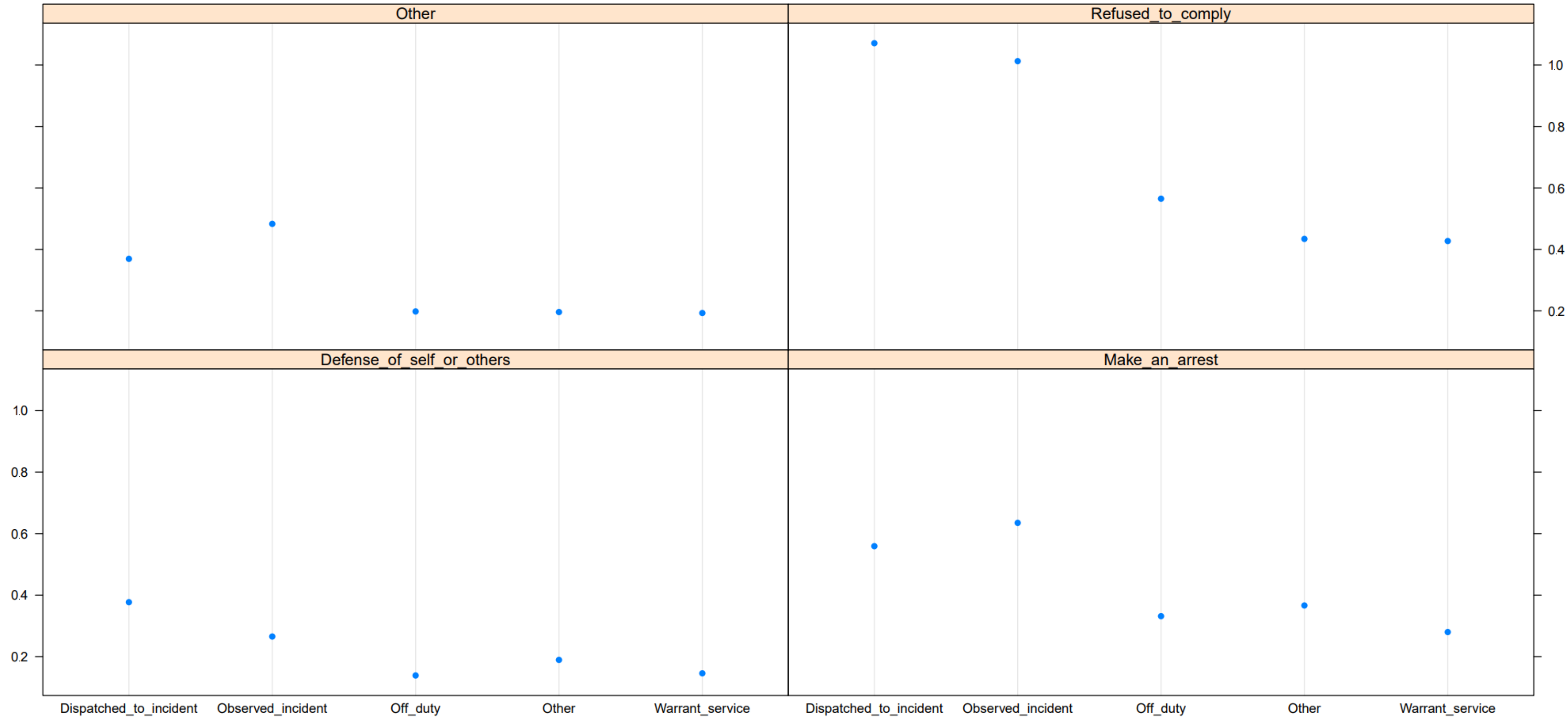

Figure 6. Partial dependence plot for the interaction effect between race of the citizen, and jurisdiction, and likelihood of drawing firearms as a means of control at use of force incidents